Advances in Spatial Science

Riccardo Crescenzi
Marco Percoco *Editors*

Geography, Institutions and Regional Economic Performance

Springer

Advances in Spatial Science

For further volumes:
http://www.springer.com/series/3302

Riccardo Crescenzi • Marco Percoco
Editors

Geography, Institutions and Regional Economic Performance

Springer

Editors
Riccardo Crescenzi
London School of Economics
Geography and Environment
Houghton Street
WC2A 2AE London
United Kingdom

Marco Percoco
Department of Policy Analysis
and Public Management
Università Bocconi
via Rontgen 1
20121, Milano
Italy

ISSN 1430-9602
ISBN 978-3-642-33394-1 ISBN 978-3-642-33395-8 (eBook)
DOI 10.1007/978-3-642-33395-8
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012950140

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Introduction

Riccardo Crescenzi and Marco Percoco

Abstract This Introduction discusses aims, objectives, rationale and structure of the book and presents some general conclusions on the role of geography, institutions and culture in shaping regional economic dynamics. An agenda for future research in this multi-disciplinary field is also presented. Regional development is a complex multifaceted phenomenon whose in-depth understanding calls for the joint consideration of a variety of factors and structural characteristics of places and agents. The understanding of regional economic performance hence calls for an explicit consideration of both "hard" and "soft" factors of development, especially in terms of geography, culture and institutions.

Keywords Geography • Institutions • Regional development

An in-depth understanding of the determinants of economic growth is crucial in order to shed light on the evolution of competitive advantage and economic performance at the national, regional and local level. In the 'standard' neoclassical approach to economic growth (Solow 1956; Swan 1956) capital abundant countries (or regions[1]) show a lower growth rate than those with a lower capital intensity due to the

[1] Assuming the same "structural" features (saving rates, population growth rate, level of technological development).

JEL Classification Numbers: O18, R12

R. Crescenzi (✉)
Department of Geography and Environment and SERC, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom
e-mail: r.crescenzi@lse.ac.uk

M. Percoco
Department of Policy Analysis and Public Management, PAM and CERTeT, Università Bocconi, via Rontgen 1, Milano 20121, Italy
e-mail: marco.percoco@unibocconi.it

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_1,

assumptions of decreasing returns to capital and technology as a free good. This implies a long-run process whereby all countries (and regions) converge upon the same steady-state (De La Fuente 2000) and “labour and capital are allocated in a manner equalizing returns among sectors and regions” (Chisholm 1995, p. 77). Conversely, in the ‘Endogenous Growth’ approach (Romer 1986), the endogenously determined rate of technical progress (persistently divergent between leader systems and late comers) and constant (Lucas 1988) or increasing (as in Romer 1990 and Grossman and Helpman 1991 in a monopolistic competition framework) returns on capital (including human capital) determine the opposite outcome: potentially persistent divergence in the economic performance of countries and regions. Similar conclusions – although based on a completely different conceptual framework – are arrived at by the Neo – Keynesian approach (Myrdal 1957; Hirschman 1958; Kaldor 1961): persistent divergence in growth rates is supported by cumulative causation processes and by the interaction of the Keynesian concepts of accelerator and multiplier which enhance and maintain divergent patterns of regional growth.

The ‘convergence/divergence’ dichotomy dominated the theoretical and empirical literature on the determinants of economic performance for almost two decades. The analysis of regional and local economic dynamics has been strongly influenced by this debate. Blanchard (1991) suggests that macro-economists have rediscovered regional economics because there are many more ‘cities, states and regions’ and more degrees of freedom. However, viewing different spatial units as interchangeable concepts reflects an a-spatial approach to the analysis of (regional) economic dynamics in which space and geography play no active role in the understanding of socio-economic phenomena.

The ‘a-spatiality’ of the macro-economic approach to regional and local economic dynamics has been progressively overcome by the development of geographical analysis in ‘mainstream’ economics and by the evolution of economic geography towards more ‘general’ conceptualisations of local processes.

The intrinsically spatial mechanisms of geographical adjustment highlighted by the New Economic Geography (Krugman 1991; Fujita et al. 1999) are of fundamental importance for an understanding of regional economic performance: divergence and convergence trends are explicitly reconciled into a homogeneous theoretical framework where the equilibrium depends on interactions between agglomeration (economies of scale, home market effect, backward and forward linkages, labour pooling) and dispersion (prices for intermediates, wages, competition) forces. Changes in transaction and transport costs (due to economic integration and “globalisation”) modify the balance between these forces thus eventually generating new core-periphery patterns.

In addition, the methodological progression – and in some sense the increasing conceptual and methodological independence from ‘standard’ macro-economic analysis – has allowed the most recent literature in regional economics and economic geography to shed light on the relevance not only of ‘proximity/geography, local synergy, and interactions’ (Camagni 1995, p. 317) but also of ‘inter-organization networks, financial and legal institutions, technical agencies and research infrastructures, education and training systems, governance structures and innovation policies' (Iammarino 2005, p. 499) in shaping innovation and, ultimately, regional economic performance. This literature has analysed the role of the geographical clustering of economic activities in the genesis of synergies favouring the process of innovation (Becattini 1987; Camagni 1995) and the impact of socio-institutional factors[2] on regional and local economic performance. In a similar vein, other theoretical contributions have investigated the exposure to 'localised knowledge spillovers' as a source of the differential 'productivity' of innovative activities pursued in different localities (Audretsch and Feldman 1996; Cantwell and Iammarino 2003; Jaffe 1986; Sonn and Storper 2005).

Regional developmental patterns are also influenced by the geography of trade flows both at local (interregional) and global (international) level. In addition, recent theoretical and empirical contributions seem to confirm this view for cities, whose economic performance is considered to depend on the flows of intra-urban (as well as inter-urban) trade and on the spatial structure of production in terms of delocalization of firms and hollowing out (Hewings et al. 1998; Rossi-Hansberg et al. 2009). However, regions interact not only on the basis of market transactions for goods and services, but also through the movement of people, mainly in terms of commuting patterns and migration flows. Individuals tend to interact more frequently and to establish long term relationships with individuals sharing the same cultural background, lowering transaction costs (Tabellini 2008). Similarly, the establishment of trade relations between regions and countries with similar cultural characteristics is also easier.

Since the work of Putnam (1993), economics literature has focused on a specific trait of culture in terms of trust as a determinant of better economic performance and institutions. People with a high level of trust seem to make organizations more efficient since the interactions within firms and between firms or agents entail lower costs. In other words, "good culture" is a factor facilitating the movement of goods and people both within and between regions because of the reduction in transaction costs. Furthermore, 'trust' promotes the spread of innovation through the diffusion of knowledge.

All these factors – largely overlooked by 'traditional' growth theories – play an important role in explaining differential regional economic dynamics and regions' differentiated capabilities to benefit from the process of globalisation. Different streams of literature have now reached a consensus on the relevance of relational, institutional and geographical factors as predictors of regional and local economic performance. However, both theoretical and empirical contributions in considering these factors have remained somewhat separated due to the underlying differences in their background literature and approach to regional economic analysis.

[2] From the "institutionalist" perspective, where innovation is not directly the outcome of a linear production function and the institutional environment acts directly as the generator of creative synergies and externalities (Dosi et al. 1988, Freeman and Soete 1997 among others) to the "evolutionary approach" (e.g. Nelson and Winter 1982) and the (regional) system of innovation approach (Freeman 1994; Lundvall 1992; Edquist 1997).

The aim of this book is to (re)consider jointly and systematically various research streams that – from different perspectives – address the role of geographical, institutional and cultural factors in the genesis of economic dynamics at the regional and local level. The various chapters of the book will cover different – though complementary – streams of literature on regional economic performance in order to emphasize their points of contact, areas of consensus (or disagreement) on the geographical, institutional and cultural determinants of regional economic dynamism. In other words, the objective of the book is not to present a single homogeneous view on this complex set of concepts. On the contrary it aims to do full justice to the heterogeneity of the existing streams of literature while shedding new light on the transmission channels (and their relative importance) that link regional growth to its 'immaterial' determinants with a special reference to the geographical, technological and institutional factors discussed above. Overall the book adopts an eclectic approach to the territorial genesis of regional growth: it combines different theoretical strands, covers the whole European Union with some chapters focused on specific European countries (i.e. Italy, Spain and the UK) but also considers non-European (New Zealand) and emerging economies (Brazil). Furthermore, the book also combines a variety of empirical approaches to the analysis of territorial economic dynamism: the macro-aggregated approach taking regions as unit of observation is complemented by the micro firm-level perspective. The evidence produced in the book is extensively applied to the analysis of regional and local development policies with substantial suggestions for future improvement.

The book is organized into four parts. The first section of the book develops the broad conceptual framework for the analysis of the relationship between geography, institutional and cultural factors and economic development. It examines regional growth and innovation dynamics from different theoretical perspectives that incorporate geographical and institutional factors into the analysis of regional economic performance in different (complementary) ways. Building upon this conceptual framework the second section explores the role of institutional and cultural factors in shaping regional economic dynamics. The focus of the third section is, instead, on geographical processes and here different chapters explore the impact of clusters, accessibility, urbanization processes and localized inter-firm linkages on economic performance. The dynamic interactions of economic agents across space are covered in the fourth part of the book that examines the geography of trade flows, labour and capital mobility.

The Chapter "Globalisation and Endogenous regional growth" by Capello and Fratesi takes a meso-level perspective and analyses the endogenous factors that allow regions to compete successfully in a globalized economy. These factors vary across different groups of regions, due to their different degrees of specialization and heterogeneous endowment in terms of connectivity with the rest of the world. The empirical analysis covers all European regions and looks at the performance of specific sub-groups of regions by comparing regions specialised in the best performing sectors and well-endowed in terms of external connections ('global players') with regions specialised in the best performing sectors but insufficiently endowed in terms of 'global' connections ('regional players'). The empirical results confirm the importance of globalization and FDI attractiveness for regional economic performance: 'global players' benefit from a persistent source of structural advantage.

In Chapter "*TFP convergence across European regions: a comparative spatial dynamics analysis*", Di Liberto and Usai address regional dynamics in terms of not only the traditional neoclassical-type of convergence but also Total Factor Productivity (TFP). By analysing a sample of 199 European regions between 1985 and 2006, they find an absence of an overall process of TFP convergence: overall TFP dispersion is virtually constant over time however a few TFP leaders are emerging and are progressively differentiating themselves from the rest, while the cluster of low TFP regions is increasing.

Fabling and Maré – in Chapter "Productivity and Local Workforce Composition" – examine the link between firms productivity and the population composition of the areas in which these firms operate. They consider the case of New Zealand and find evidence of productive spillovers from firms active in areas with a high-skilled workforce, and with high population density, although the strength and nature of these spillovers varies across different types of firms.

Chapter "Regional Growth in Central and Eastern Europe: Convergence, Divergence and Non-linear Dynamics" by Monastiriotis views the collapse of central planning in Central and Eastern Europe (CEE) as a quasi-natural-experiment, particularly suitable for the study of the applicability and relevance of competing theories of regional growth: the neoclassical convergence hypothesis (NC), the cumulative causation theory (CC) and the evolutionary approach of the regional Kuznets curve (KC). The empirical results provide evidence in support of all three processes (convergence, cumulative growth and a regional Kuznets curve) and, on balance, seem to favour a hybrid explanation. Regional growth is non-linear process that depends on the level of national development and produces aggregate divergence and polarisation at later stages of development despite an overall tendency towards convergence.

In chapter "Some Practical Elements Associated with the Design of an Integrated and Territorial Place-Based Approach to EU Cohesion Policy" – that concludes the first section of the book – McCann and Ortega-Argilés discuss how bottom-up place-based development policies can be used in order to address the disparities in economic development and innovation identified in the previous chapters as the result of the interaction between 'global' (e.g. FDI flows) and highly localized (e.g. knowledge spillovers) effects. In particular, this chapter focuses on practical ways in which an integrated place-based approach to territorial and social cohesion can be exploited to tailor policy-design to the specific needs and potential of the EU regions.

After the examination of long term economic dynamics and their determinants, the second part of the book deals more specifically with the relevance of cultural and institutional factors for regional economic performance. Bellini, Ottaviano, Pinelli and Prarolo (chapter "Cultural Diversity and Economic Performance: Evidence from European Regions") investigate the relationship between diversity and productivity in Europe using an original dataset covering the NUTS 3 regions of 12 countries of the EU15. The key hypothesis explored in the paper is that cultural diversity may affect both production and consumption through positive or negative externalities. In this context, a joint estimation of price and income equations is necessary to identify a dominant effect: empirical analysis shows that diversity is positively and causally correlated with productivity, and, as such, consistent with the results obtained for US cities by previous literature.

The role of cultural factors as drivers for regional economic dynamism is further explored in chapter "The 'Bright' Side of Social Capital: How 'Bridging' Makes Italian Provinces more Innovative" by Crescenzi, Gagliardi and Percoco. The authors study the differential role of 'bonding' and 'bridging' social capital in the genesis of innovation in the Italian Provinces. Quantitative analysis shows that social capital is an important predictor of innovative performance after controlling for 'traditional' knowledge inputs (R&D investments and human capital endowment) and other characteristics of the local economy. However, only 'bridging' social capital – based on weak ties – can be identified as the key driver of innovation.

Localized cultural factors can also influence regional development and growth by shaping local attitudes towards entrepreneurship. In chapter "Explicitly Implicit: How Institutional Differences Influence Entrepreneurship", Bauernschuster, Falck, Gold and Heblich compare individuals born and raised in the former socialist East Germany to their West German counterparts and find that the socialist regime had a strong impact on attitudes that are negatively associated with entrepreneurship. The analysis of East Germans who moved to West Germany after the fall of the Berlin Wall reveals that the socialist legacy not only runs through the channel of a less developed economic environment but indeed through individuals' mindsets.

Entrepreneurship and its determinants are also analysed in chapter "Academic Entrepreneurship and the Geography of University Knowledge Flows in the UK" by Abreu and Grinevich who look at the impact of personal, institutional and spatial factors on academic entrepreneurship in the UK. This analysis shows that while academics face a number of obstacles (including lack of experience, cultural differences and time constraints), academic entrepreneurship is still a significant phenomenon in most disciplines. The analysis shows that informal activities are common, but do not always translate into measurable outcomes such as patenting, licensing and spin-offs. Personal characteristics such as experience in business or the public sector are important determinants of academic entrepreneurship, as are the social norms and culture that prevail at the institutional level. The authors also find that the spatial dimension of knowledge-exchange activities varies by type of institution; academics in more vocational institutions are more likely to collaborate with local private, public and third sector institutions.

In the third part of the book the role of geographical factors and agglomeration economies for the location of economic activities is explored in depth. Cluster analysis has attracted considerable interest over the last few decades. The translation of clusters into complex relational networks and their embeddedness in (regional) systems of innovation has been generally associated with the delivery of greater innovation and economic dynamism. However there has been little systematic research into the relationship between clusters and other (social and institutional) factors promoting innovation and economic growth. Chapter "Evaluating the Role of Clusters for Innovation and Growth in Europe" by Rodríguez-Pose and Comptour addresses this issue by examining the relationship between clusters, innovation and economic growth in the EU regions. The authors show that the presence of clusters matters for regional growth but only in combination with a good 'social filter', and this association wanes over time. In addition, more traditional R&D variables have a weak initial connection to economic development, but this connection increases with time and remains contingent on the existence of adequate socioeconomic conditions.

Chapter "Input–Output Linkages, Proximity to Final Demand and the Location of Manufacturing Industries" by Mion zooms into the micro-economics of agglomeration processes by developing an empirical framework to estimate the role of agglomeration externalities, (and in particular those stemming from input–output linkages) in the location process of US manufacturing plants. The empirical results suggest that intermediate inputs have a positive impact on local specialization while consumers' demand has a negative effect. However, the majority of both effects comes from very local interactions, with spatial spillovers being quite weak over a large geographical radius: while very close interactions are extremely important, outside the boundaries of local markets the importance of geographical distance substantially dwindles.

Chapter "Agglomeration and Labour Markets: The Impact of Transport Investments on Labour Market Outcomes" (D'Costa, Gibbons, Overman and Pelkonen) addresses the role of geography as an explanation for economic dynamics by focusing on cities – rather than 'industrial' clusters as in previous chapters – and exploring how changes in accessibility impact upon the composition of the local workforce through sorting, and/or because people change their characteristics in response to changes in economic mass. The econometric analysis (based on the cases of Leeds and Manchester) shows that individual wage growth is faster in places with better access to economic mass, but this effect is driven by the fact that these cities tend to have more educated workers. After controlling for this there is essentially no relationship between city size and wage growth. The results suggest that the aggregate effects of closer integration/better accessibility may be greater than the individual effects. This aggregated effect relies on structural changes moving the composition of better integrated labour markets towards higher skilled jobs.

In chapter "Firm Capabilities and Cooperation for Innovation: Evidence from the UK Regions", Iammarino, Piva, Vivarelli, and von Tunzelmann examine the relationship between firms by exploring the impact of inter-organisational cooperation on their technological competence and capabilities and how this relationship is influenced by firms' geographical location. Their analysis is based on a large sample of UK firms and suggest that linkages with other enterprises within the group, suppliers, clients, and public research and higher education institutions are highly significant predictors of firms' technological status. These findings show remarkable regional specificities in terms of the association between collaborative patterns and technological capabilities at the firm level, suggesting that contextual and geographical conditions do play a very important role in shaping micro-economic outcomes.

The final part of the book is devoted to the analysis of the dynamic interactions between economic agents across space by looking at trade flows and factor (labour and capital) mobility. In chapter "Assessing Regional Economic Performance: Regional Competition in Spain Under a Spatial Vector Autoregressive Approach", Marquez, Ramajo and Hewings consider the impact of trade-related spatial externalities on regional economic dynamics. In order to measure these geographical spillovers, a trade-based weighted measure of transmission of externalities is considered in the context of a dynamic spatio-temporal analysis of regional growth. A spatial vector autoregressive (SpVAR) model for the Spanish regions is specified including both spatial and temporal lags of the state variables. The empirical results suggest the existence of inter-regional externalities related to trade for the Spanish regional economic system. In addition, the estimated SpVAR is used to calculate impulse responses that provide insights on the effects of macroeconomic events in different locations.

Further insights on the relationship between geography, economic dynamics and trade are developed by looking – in chapter "The Long Run Interplay Between Trade Policy and the Location of Economic Activity in Brazil Revisited" by Ferraz, Haddad and Terra – at the question of whether trade liberalization may exert some influence on the heterogeneous economic geography of Brazil. Through the specification of an interregional CGE model for the Brazilian economy, the chapter argues that the adoption of (horizontal) liberal trade policies, beyond traditional gains from trade, can also contribute towards ameliorating regional inequality in the country. In this sense, trade liberalization in Brazil is advocated as an additional (horizontal) public policy apparently effective in fighting regional inequality trough traditional market forces.

In chapter "Environmental Standards, Delocalization and Employment: The Case of the EU Cement Industry", by Bosco and Altomonte, further attention is given to the link between the geography of trade and public policies by analysing the effect of environmental regulations on industrial conditions for the entire EU. This chapter provides an estimate for an empirical formulation in two steps: first, by estimating the direct impact from increased imports on the EU cement industry employment; secondly, by estimating how clinker imports are affected by increases in environmental costs. The result suggest that the overall impact on employment is negative.

In chapter "Migrant Heterogeneity and Urban Development: A Conceptual Analysis", Bakens and Nijkamp examine the contribution of migrant heterogeneity to the attractiveness of cities from both the production and the consumption side. Based on an extensive literature review, the authors hypothesize that the interaction of people from different cultural groups in cities will increase labour productivity in line with the concepts of Jacobs externalities. As concerns the consumption side of the model – a far less researched issue – they hypothesize that urban cultural diversity increases heterogeneity in the private goods provided, which, in turn, will increase the utility of living in that area. Finally, Bakens and Nijkamp argue that future research should focus on the interaction between people from different cultures in the workplace in order to determine urban productivity externalities, and on migrant-induced product heterogeneity in a city in order to determine migrant-induced urban amenities.

Chapter "What Drives Chinese Multinationals to Italy?" by Pietrobelli, Rabellotti and Sanfilippo addresses the question of capital mobility by exploring the geography of Chinese FDI in Italy. The in-depth empirical analysis developed in this chapter confirms the importance of both geographical (clustering) and cultural factors as drivers for the location decisions of foreign investments. Chinese investments in Italy are increasingly targeting the acquisition of design and brands in key Italian sectors of specialization, and technological capabilities in sectors such as metalworking. Chinese multinational enterprises also are investing in Italy in order to acquire access to local competitive advantages: their location choice is clearly linked to the intention to tap local competences available at the cluster level.

Interregional relations are also crucial when evaluating the effect of economic policy, as highlighted by Ichihara, Guilhoto and Imori in chapter "Regional Development and Government Income Transfer Programs: Combining Input–Output Systems and Geoprocessing as Tools for Planning in São Paulo State, Brazil". To measure the regional impacts and possible development and planning policies of the "Bolsa Familia" Program, an income transfer program founded by the Brazilian Federal Government, the authors combine geoprocessing with input–output theory. The analytical potential of the tool is shown through the estimation and analysis of the São Paulo State Municipalities supply and demand relations. These relations were estimated in the process of the construction of the state interregional input–output system. Bearing in mind that the creation of accurate development strategies depends on regional peculiarities, the results show that this program must be understood not only as a form of income transference, but also as a catalytic agent for decreasing regional inequality inside the state.

The book's preparation brought together scholars from several countries and different 'sister' disciplines (Urban and Regional Economics, Economic Geography, Innovation Studies etc.) with different approaches to the same crucial issue: how geography, culture and institutions influence regional development (conceptualised and measured in the broadest sense). The book includes a number of relevant insights into these complex relations and these – in line with the four sections around which the book is organised – may be summarized in four main directions for future research.

First, formal economic theories for the analysis of regional economic performance can be successfully cross-fertilised with other disciplines in the social sciences in order to account for the role of geographical and institutional factors, which remains crucial for the understanding of empirical phenomena.

Second, culture and *de facto* institutions promote economic performance through different channels. Societies characterized by a large endowment of bridging social capital are more open and knowledge spreads faster, leading to higher growth rates and higher propensity to innovate. Similarly, culture may influence individual decision to become an entrepreneur and this, in aggregate, implies that part of the spatial variation in entrepreneurship rate is explained by cultural differences, other things being equal.

Third, the concentration of people and firms in space produces comparative advantage and higher productivity. These productivity advantages increase when levels of spatial interactions are higher or are promoted by decreasing transport costs (although the effect of such innovation can be non-linear).

Fourth, spatial interactions may take several forms, among them being trade and migration. Both phenomena have long been investigated in regional science literature, although their net effect on regional development is so complex that there is no consensus on the sign and magnitude of their impact. In fact, the effect of trade liberalisation on regional disparities has been found to be unclear as well as the one of globalization and firm migration. Similarly, migration flows have long been considered to be an equilibrating mechanism for spatial imbalances. However, increasing attention is now being paid to their human and social capital content. The development of these four key areas of research forms the basis of an ambitious multidisciplinary research agenda.

To conclude, the book is organized around the idea that regional development is a complex multifaceted phenomena whose in-depth understanding calls for the joint consideration of several strands of literature by taking explicit account of both "hard" and "soft" factors of development, especially in terms of geography, culture and institutions.

References

Acs ZJ, de Groot HLF, Nijkamp P (2002) The emergence of the knowledge economy. A regional perspective. Springer, Heidelberg

Audretsch DB, Feldman MP (1996) R&D spillovers and the geography of innovation and production. Am Econ Rev 86:630–640

Becattini G (1987) L'unità di indagine. In: Becattini G (ed) Mercato e forze locali: il distretto industriale. Bologna, Il Mulino, pp 35–48

Blanchard OJ (1991) Comments on [Barro and Sala-i-Martin, 'Convergence across states and regions']. Brooking Pap Econ Activity 1:159–174

Camagni R (1995) Global network and local Milieux: towards a theory of economic space. In: Conti S, Malecki E, Oinas P (eds) The industrial enterprise and its environment: spatial perspective. Avebury, Aldershot, pp 195–216

Cantwell J, Iammarino S (2003) Multinational corporations and regional systems in Europe. Routledge, London

Cheshire PC, Carbonaro G (1995) Convergence-divergence in regional growth rates: an empty black box? In: Armstrong HW, Vickerman RW (eds) Convergence and divergence among European regions. Pion, London

Chisholm M (1995) Britain on the edge of Europe. Routledge, London

De la Fuente A (2000) Convergence across countries and regions: theory and empirics. EIB papers 5(2):25–45

Dosi G, Freeeman C, Nelson R, Silverberg G, Soete L (eds) (1988) Technical change and economic theory. Pinter, London

Freeman C (1994) Critical survey: the economics of technical change. Camb J Econ 18:463–512

Freeman C, e Soete L (1997) The economics of innovation. Continuum, London

Fujita M, Krugman P, Venables A (1999) The spatial economy. MIT Press, Cambridge, MA

Grossman GM, Helpman E (1991) Innovation and growth in the global economy. MIT Press, Cambridge, MA

Hewings JGD, Sonis M, Guo J, Israilevich PR, Schindler GR (1998) The hollowing-out process in the Chicago economy, 1975–2011. Geogr Anal 30(3):217–233

Hirschman AO (1958) The strategy of economic development. Yale University Press, New Haven
Iammarino S (2005) An evolutionary integrated view of regional systems of innovation: concepts, measures and historical perspectives. Eur Plan Stud 13(4):497–519
Jaffe AB (1986) Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market share. Am Econ Rev 76:984–1001
Kaldor N (1961) Capital accumulation and economic growth. In: Lutz FA, Hague DC (eds) The theory of capital. MacMillan, London
Krugman P (1991) Increasing returns and economic geography. J Polit Econ 99:484–499
Lucas R (1988) On the mechanics of economic development. J Monet Econ 22:1
Lundvall BÅ (1992) National systems of innovation: towards a theory of innovation and interactive learning. Pinter, London
Myrdal G (1957) Economic theory and under-developed regions. Duckworth, London
Nelson RR, Winter SG (1982) An evolutionary theory of economic change. Belknap Press/ Harvard University Press, Cambridge
Putnam R (1993) Making democracy work: civic traditions in modern Italy. Princeton University Press, Princeton
Romer PM (1986) Increasing returns and long-run growth. J Polit Econ 94:5
Romer PM (1990) Endogenous technological change. J Polit Econ 98:5
Rossi-Hansberg E, Sarte P-D, Owens R (2009) Firm fragmentation and urban patterns. Int Econ Rev 50(1):143–186
Solow RM (1956) A contribution to the theory of economic growth. Q J Econ 70:1
Sonn JW, Storper M (2005) The increasing importance of geographical proximity in technological innovation: an analysis of U.S. patent citations, 1975–1997. Mimeo
Swan TW (1956) Economic growth and capital accumulation. Econ Rec 32:334–361
Tabellini G (2008) The scope of cooperation: values and incentives. Q J Econ 3:905–950

Part I
Space, Growth and Development

Globalization and Endogenous Regional Growth

Roberta Capello and Ugo Fratesi

Abstract Globalisation is not a state of the world but an evolutionary process, which entails the increasing planetary integration of markets for goods and services, markets of location sites for economic activities, markets of production factors as technologies and information. Regions are involved in the globalization process to a different extent depending on their structure and specialization.

The first aim of this paper is to highlight the importance of a regional dimension in the analysis of globalization trends, and to explore the debate between exogenous and endogenous factors driving economic development, with their relative importance appearing to be different according to the development stage of the economies.

The paper also investigates the factors of growth, showing that, after national effects and innovative capabilities, one of the most important aspects is represented by FDI penetration, whose impact is shown to differ according to the source, sector and technological level.

1 Introduction

Globalisation is not a state of the world but an evolutionary process, which entails the increasing planetary integration of markets for goods and services, markets of location sites for economic activities, markets of production factors as technologies and information.

For sure, globalisation is not a new phenomenon and in many periods of last century it reached very high and even comparable levels than today; moreover, it did not show up in a single, catastrophic jump, as the sudden adoption and fortune

---

JEL code: R10, R11

R. Capello (✉) • U. Fratesi
Department of Building, Environment, Science and Technology, Politecnico di Milano, Milan, Italy
e-mail: Roberta.Capello@polimi.it; Ugo.Fratesi@polimi.it

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_2,

of the term in the political debate could suggest. What is new is the long-term, contemporary acceleration of many parallel integration processes, reinforcing and integrating one another in multiple ways. Since almost 30 years, international trade has been steadily growing at a rate which is the double of world GDP. Foreign direct investments, on their turn, have grown at rates which are the double with respect to international trade, and four times those of world GDP. Most of these investments are directed towards developed countries (they were 80 % in the years 1986–1990, around 60 % in years 1993–97 and still slightly more than 50 % in 2009 despite of the differential effects of the crisis at World level, UNCTAD 2010) and look particularly attracted by accelerations in economic integration processes: in fact EU countries, at the top of the process of creation of the Single Market in 1991–92, received up to 50 % of world FDI (UNCTAD 1997; Camagni 2002) and still attract 2/3 of FDI of developed countries (UNCTAD 2004, 2008). Moreover, mobility (and volatility) of financial capital has grown spectacularly: in 1995 financial exchanges reached 1,000 billion dollars a day, more than the foreign exchange reserves of all national governments together. The short term profit objective of these movements imposes serious constraints on the governance of the international financial system. Finally, the nature of international trade has evolved from pure exchange of (final) goods among national production systems, to exchange of intermediate goods through the internationalisation of functions within production networks organised on a world-wide scale, to the most recent unbundling of functions themselves in specific tasks, leading to a trade-in-task economy. With these qualitative changes, local production systems find themselves increasingly tied together and interdependent, mainly through the global strategies of multinational corporations.

Much theoretical and empirical work has been developed on globalization, trying to capture different effects of the quali-quantitative changes imposed by the integration of markets through either multilateral or "regional" liberalization policies (Panagariya 2000)[1]; new international trade patterns which see more and more developing and emerging countries as exporters of manufacturing goods, thus forcing industrialized countries to change their specialization towards high quality goods and, mainly services (Bergoeing et al. 2004; Kucera and Milberg 2003), new composition of intermediate vs. final goods traded at international level, also as a result of multinational firms' new strategies (Yi 2003; Hummels et al. 1998, 2001;

[1] We refer here to Regional Trade Agreements (RTAs), which have characterised the present wave of globalization not only because their number has exponentially increased since the World War II, but mainly because they have changed both in nature and motivations (Fiorentino et al. 2007). RTAs may have both positive and negative effects on international trade relationships. They can play an important role in promoting the liberalisation and expansion of trade and fostering growth and development and so acting as stepping stones on the way to a multilateral agreement. But regional agreements also risk making it harder for countries outside the region to trade with those inside and may therefore discourage further opening up of markets, ultimately limiting growth prospects for all. On this still open debate on RTAs as stumbling or building blocks toward multilateralism see Winters 1996; Panagariya 1999, 2000; Baldwin 1995; Baldwin and Venables 1995.

Hanson et al. 2005), new location patterns of foreign direct investments and consequent new growth opportunities for developing economies (Hansen and Rand 2006; Lall and Narula 2004; Moran et al. 2005), migration trends and international trade flows (Soubbotina 2004; Lucas 2008), represent some of the main issues treated in the recent literature.

From the perspective of the above mentioned studies, though, globalisation can be regarded as neutral for what concerns its spatial effects: opportunities and threats may look equivalent and specular. A number of good reasons exists, however, for claiming that a *regional perspective* is instead fundamental in order to understand the real economic effects of globalization, and that conceptual and empirical analyses at regional level are fundamental.

In particular, in front of the intensifying both quantitative and qualitative trends in globalization an old debate in regional growth becomes again a hot topic of discussion: regional patterns of growth can be the result of either internal forces and endogenous capacity of a region to grow, or of exogenous forces, that reach a local economy from outside and give rise, in a cumulative self-reinforcing mechanism, to a local process of growth. In particular, in this paper the aim is to analyze the role of endogenous vs. exogenous factors allowing regions to grow. These factors are expected to vary between Western and Eastern regions.

The paper will hence be organized as follows: in the next section the importance of a regional dimension in the analysis of globalization trends is highlighted, and the debate between exogenous and endogenous factors driving economic development highlighted. Our impression is that the role of exogenous and endogenous factors highly depends on the development stage of the economies, and that it is therefore different between Eastern and Western countries. Section 3 presents the recent trends in regional disparities highlighting that the development stage is still rather different between eastern and Western countries and that growth disparities among regions in Europe persist.

Sections 4 and 5 will be devoted to highlight success factors behind the different regional performances, putting most emphasis on FDI as an exogenous factor of growth, compared to the traditional material factors explaining endogenous growth. Section 6 will conclude the paper.

2 Globalization and Regional Competitive Assets

As a consequence of the increase in globalization processes regional economies face fiercer competition, that leads to a worsening of regional disparities, especially driven by intra-national disparities, exacerbated by the concentration of economic resources in most advanced and dynamic places, where the most successful cities lie, and by resource inefficiency and lack of competitive advantage in peripheral regions. All this is even more remarkable, if one thinks that regional economic systems are more vulnerable to external shocks than nations; regions are by definition very open economies, highly dependent from external trade conditions and international terms of trade, from external final goods for internal consumption and from external intermediate goods and natural resources for local productions.

Moreover, a situation of fiercer competition is even more dramatic if one thinks that, differently from nations, regions compete on the basis of absolute rather than comparative advantages. The two "classical" equilibrating processes of a comparative advantage rule *à la* Ricardo do not work properly or do not exist at the regional level: the first process relies on downward flexibility of prices and wages, which is widely hampered by the existence of national wage contracts in both private and public structures and by the homogeneity of import prices (we remind that regions are very open economies); the second "modern" process relies on the devaluation of the currency, and it is automatically excluded in an inter-regional context (Camagni 2002). The ricardian conclusion that each country will always be granted some specialisation and role in the interregional division of labour is not valid for regions. A region can well be pushed "out of business" if the efficiency and competitiveness of all its sectors is lower than that of other regions, and its fate is, in this case, mass unemployment and, in case of insufficient public income transfers, emigration and possible desertification. In front of this possible scenario, taking care of the regional effects of stronger global competition bears a strong economic rationale.

The capabilities of a region to grow require deep understanding. As it is widely accepted nowadays by the most advanced literature on the subject, long-term local development is largely a supply-side phenomenon, based on general rules and institutional frames and above all nourished by the internal entrepreneurial capabilities of regions and places and by the local capability of efficiently exploiting existing resources, local policies require a deep knowledge of local resources and potentialities. This means that the possibility for any region to contribute to the general EU growth is dependent on the fact that it creatively exploits its territorial capital, enriches it in the right ways setting appropriate priorities to local and regional policies, and "taps" and mobilizes previously "untapped" assets of its territorial capital.

Territorial capital may be seen as the set of localised assets – natural, human, artificial, organizational, relational and cognitive – that constitute the competitive potential of a given territory (Camagni 2009). It was launched explicitly in the early 2000's by the OECD (OECD 2001) and re-launched by the EU Commission in its *Guidelines to Structural Funds* in 2005[2]: agglomeration economies, equilibrated and polycentric urban structures, accessibility, skilled labour force, R&D and high level education facilities, business networks and social capital, natural resources and cultural heritage, territorial diversity and territorial identities are indicated as the assets and preconditions for regional growth that need to be properly identified, wisely protected and strengthened, smartly utilised, continuously re-interpreted and re-oriented.

The strategic factors that enable a region to achieve and maintain a position in the international division of labour over the long run are more and more non-material

[2] "Each Region has a specific 'territorial capital' that is distinct and generates a higher return for specific kinds of investments than for others. Territorial development policies should first and foremost help areas to develop their territorial capital" (CEC 2005).

factors, linked to knowledge, culture, taste and creativity (see also Crescenzi and Percoco in the introductory chapter to this volume). The laws of accumulation of these elements are especially dependent on local aspects: in fact all these elements develop through slow learning processes, fed with information, interaction, long term investments in research and education. Like all learning processes, they are inherently localised and cumulative, as they embed in human capital, interpersonal networks, specialised and highly skilled local labour markets and local innovative *milieux*; therefore they are highly selective in spatial terms (Camagni 1991a; 1999). Moreover, while traditional material production factors are subject to a hyper-mobility, marketed and utilised everywhere (playing apparently no role in a competing environment), the skills and "relational capital" required for their most efficient or innovative use are by no means available everywhere, and are these elements that make the difference: trust (Glaeser et al. 2000; McCloskey and Klamer 1995), social capital (Glaeser et al. 2002; Knack and Keefer 1997; La Porta et al. 1997; Beugelsdijk and van Schaik 2005), sense of belonging to a society (Bowles et al. 2001; Lazear 1999; Alesina and La Ferrara 2000) are nowadays the main sources of increasing returns for traditional economic production factors (Capello et al. 2011a; Caragliu 2009). These elements are highly heterogeneously distributed at regional level, something that explains the high and persistent level of TFP heterogeneity across regions (see Diliberto and Usai in this volume).

All these reflections lead to the consideration that the capabilities of regions to compete in a global world mainly lie in endogenous territorial assets (Affuso et al. 2011; Fratesi 2012). However, some realities exist in which exogenous forces, in the form of foreigner productive capital, play an important role. This is especially true in areas where the stage of development is still lower than the EU average and where endogenous territorial capital assets are, in qualitative and quantitative terms, lower than the EU average. A firm becomes multinational in order to exploit three kinds of advantages, summarized in the acronym of the well-known OLI paradigm: ownership, internalization and locational advantages (Dunning 2001). For what concerns the locational advantages, foreigner firms are attracted by either large market potentials or by labour cost advantages, as also most of the empirical literature has demonstrated (Resmini 2007, 2008).

Whether the presence of exogenous or endogenous factors play a role in regional growth is empirically investigated in this chapter in the context of globalization. In particular, our hypothesis is that endogenous factors play a more important role in advanced stages of development, while exogenous elements are more striking in lower development stages. This will be empirically analysed by running the empirical analysis in Western and Eastern Europe respectively. Moreover, our a-priori is that regional territorial capital assets are fundamental for explaining the capacity of a local area to grow more than its nation; however, among the causes of regional success and failure are factors which are directly linked to certain pervasive and generalized characteristics of the national economy. We refer in particular to institutional factors such as the performance of the high functions of the nation-state – legislative, judicial and governmental; to organizational factors such as the efficiency of services of general interest like education, transportation, communication, health and security services; to economic factors such as general fiscal pressure, the effectiveness of public expenditure, the pervasiveness of environmental regulations, and the efficiency of contract enforcement procedures. Once competition is at world level, the international disparities in legislative, judicial and governmental factors, as well as in the efficiency of public services widen dramatically; as a consequence, the role of these elements in regional growth explanations grows (Capello et al. 2011b).

In the rest of the paper, we will hence investigate the effects of endogenous and exogenous factors on regional growth, considering territorial capital assets as the endogenous factors and considering two different exogenous ones: on the one hand national growth, which reflects exogenous national factors and, on the other hand, foreign direct investments, which account from global push forces.

3 Different Stages of Development and Regional Disparities in Western and Eastern Countries

We claim in the chapter that Eastern and Western countries are still in a different development stage. An analysis of the trends in regional disparities is presented here to confirm that, despite the growth rates measured in recent years in Eastern countries, Eastern regions still lag behind.

To represent regional disparities the Theil index is used, which has the precious characteristics of being decomposable into parts, i.e. of allowing to disentangle how much disparities depend on one factor or another.

Figure 1 represents the general Theil indexes of regional disparities, i.e. without taking into account globalization forces in order to work as benchmark. If we look at the total European regional disparities (Fig. 1a) we can observe that the total Theil index of regional disparities has decreased significantly from 1995 to 2005 (our period of analysis). This is due, as found in other works, to the decrease of the between countries disparities, whereas within countries there is a small but consistent increase of disparities, signalling that lagging countries have generally outperformed the strongest ones, but lagging regions have generally been unable to catch up with their national frontrunners.

The aggregate effects, however, hide the fact that an important effect in the convergence process has been due to the stronger performance of the New12 member countries of the EU, which are still significantly less rich than their western counterparts but have been growing much faster. This can be observed in Fig. 1b, where it can be observed that a large part of EU total disparities (about two thirds) is due to the difference between Old15 and New12 member countries and that, while this part has decreased fast, the disparities within the two parts of Europe have increased, though with a lower pace.

By using an additional and innovative decomposition it is possible to examine the role of three levels at the same time. For this reason in Fig. 1c it can be observed that, once the very large and decreasing effect of Old15-vs-New12 countries has been

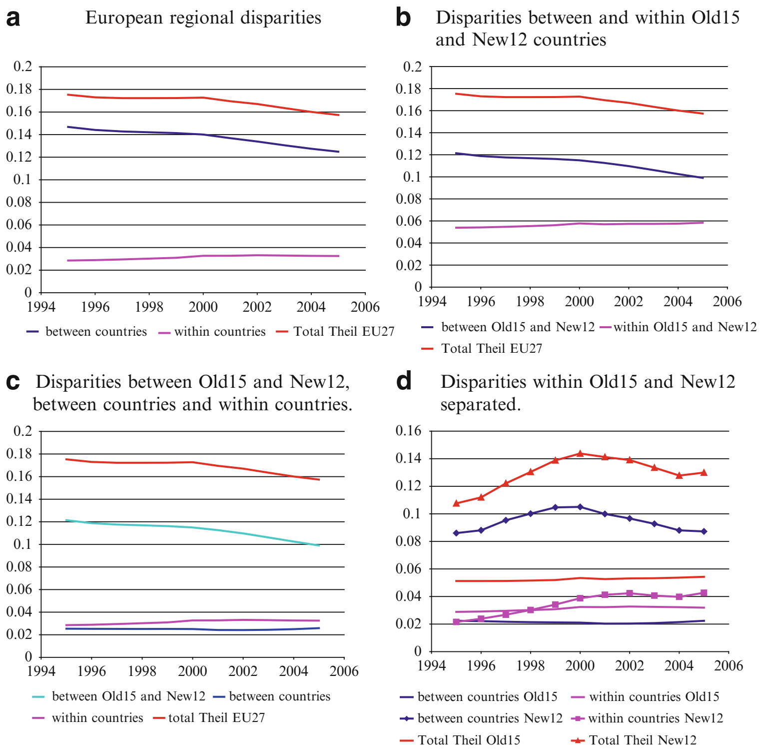


Fig. 1 General Theil indices of regional disparities

extracted, the remaining of regional disparities can be attributed in similar parts to between country and within country disparities, with the latter slightly larger than the former, signalling that the results of Fig. 1a are biased by the difference between Old15 and New12, and once it is wiped out, within country disparities are even more relevant than between country ones. Notice that the two effects are slightly increasing, differently from the disparities between New and Old member states, which decreases steadily and consistently throughout the period of analysis.

In order to see if the two groups of countries hide different patterns, Fig. 1d represents in the same picture (for comparative purposes) the Theil indexes calculated between and within countries for Old15 and New12 countries separately. It is immediately evident that the total level of disparities within the New Member States is considerably higher; moreover, in these countries total disparities exhibit in many years a tendency to increase, whereas they are substantially stable in Old15 countries. This is due to the fact that, in New12 member states, between country disparities first increase and then decrease, whereas within countries exhibit a clear growing pattern, due to the fact that the core areas of these countries have normally outperformed the rest of their respective countries, probably because they were better fit for global challenges. This appears to be consistent with the old

Williamson (1965) curve, which sees regional disparities grow as income per capita grow in the earlier stages of development and with the findings of Monastiriotis in this volume.

Interestingly enough, within countries disparities in the New12 member states have now exceeded those in Old15 countries, which have only marginally increased. All types of disparities (total, between countries and within countries) have remained quite stable in Old 15. The last aspect which is interesting to observe is that in Old15 countries the disparities between countries are lower than those within, signalling that dualisms between rich and poor regions are more important than differences among countries, whereas in the New Member states the disparities between countries remain significantly higher than those within countries despite the doubling of the latter.

4 Endogenous and Exogenous Regional Success Factors

The previous sections have evidenced that the patterns of growth of European regions are differentiated. The descriptive analysis has evidenced that the characteristics associated with the performance of regions are very different between Eastern and Western regions.

Figure 1 also evidenced that country effects take place, so that the general performance of regions highly depends on the country to which they belong. National economic trends, especially some trends linked to globalization processes, like the movements of financial capitals, interest rates and exchange rates, are exogenous growth factors which exert their effects at national level. Other aspects of the globalization processes, on the contrary, despite being exogenous, show their effects at regional level, for instance most aspects linked to the reorganization of the production processes, on which the attraction of local economies of high or low value added tasks and phases plays a crucial role. It becomes therefore interesting to analyze what of the regional structure affects regional performance once country-wide, mainly monetary, variables are kept separate.

In this section of the paper, therefore, the use of regression analysis will be made with the purpose of investigating in a causal way the factors which explain the growth rate of European regions, focusing on FDI, the main globalization channel for which we have data. One way to do such an analysis would be to use as dependent variable the GDP differential growth of the regions with respect to their respective countries, but this assumes that country effects are taking place for all of them in the same way; a better option is to use regional growth as dependent variable and introduce the national growth rate among the regressors, so that the data are allowed to estimate the elasticity of regional growth to national growth without imposing the restriction that it is 1, a restriction which is implicitly assumed when regressing the differential growth rate.

To detect the structural features more associated with positive regional economic performance of regions in an age of globalization we hence use multivariate regression analysis in which a number of factors be used together in an explicative model.

The choice of the success factors explaining regional performance was based on the consideration that the differential growth rate is what remains to be explained once the national effects have been considered by including the national growth rates among the regressors. A recent theory summarises the elements explaining endogenous regional growth in what is termed 'territorial capital', which consists of material and intangible, private and public, soft and hard elements (Camagni 2009). For this reason, a rather differentiated set of local assets were chosen: some were traditional material factors like transport infrastructure, geographical position, and the functions in which a region is specialised; others were intangible, like agglomeration economies, with the usual limitations that characterise a database that must cover the entire European territory.[3]

In particular, the following endogenous territorial capital variables were selected (see Table 1 for the description of variables):

- The growth effects induced by the regional geographical position, whether it is close to fast-growing regions or close to regions unable to grow fast (*spill*). These effects can be positive or negative depending on the role of neighbouring regions, which may induce growth through demand effects or steal it away through competition. The indicator used was a spatial growth spillover indicator for a generic region *r*, capturing economic potential (Clark et al. 1969) as the sum of the annual absolute difference between income growth rates of all other regions *j* divided by the distance between each region *r* and region *j*, defined as[4]:

$$SP_{rt} = \sum_{j=1}^{n} \frac{\Delta Y_{jt}}{d_{rj}}; r \neq j \tag{1}$$

where:

ΔYjt = income growth rate of region j at time t
j = all regions except region r
drj = physical distance between region r and j
n = all regions of the sample;

[3] All independent variables were lagged in order to reduce problems of endogeneity and reverse causation.

[4] An indicator weighting each regional growth rates for the share of each regional economy (GDP) on the European total GDP was calculated in addition to the non-weighted one. A high statistical correlation emerged between the two, as shown by a Pearson correlation coefficient of 0.93. Moreover, the difference between the two standardised indices showed a low spatial autocorrelation, with a Moran's I index of 0.30. On removing a few outliers (mainly Nordic and Spanish regions), the Moran's I index was 0.18. On the basis of this correlation, it was decided to use the non-weighted spillover indicator, given its closer similarity to the classic spatially-lagged models of spatial econometrics. This indicator is an economic potential measure which is generally calculated as the accessibility to total income at any location allowing for distance, following Clark et al., 1969. Here the concept of economic potential is measured in terms of accessibility to the income growth rates.

Table 1 Variables description and data sources

Data and indicators	Definition	Source of raw data
Regional GDP	Regional GDP in real terms at NUTS2 level in the period 1995–2005, computed from the nominal one, using national GDP deflators.	Eurostat
Regional average annual differential GDP growth rate	Annual average regional GDP growth rate less national GDP growth rate in the period 1999–2002	Eurostat
FDI	Number of new foreign firms per million inhabitants. Reference period 1999–2001	FDIRegio database
Regional employment by function (ISCO)	Regional employment by function at ISCO 2 digit classification at Nuts 2 level	European Labour Force survey
Innovation/regional share of human resources in S&T	Share of people working in S&T on population in the year 2000	Eurostat
Regional infrastructure endowment	Km of high speed railways, main rails, express roads, motorways and inland waterways in year 2000	KTEN data within the Espon database
Per capita structural funds	Total structural funds expenditure/population in the period 1994–1999. Also divided into 5 types of expenditure	Espon database
Spatial growth spillovers	Calculated for the period 1999–2002	Eurostat
Agglomerated regions	With a city of > 300,000 inhabitants and a population density > 300 inhabitants/km sq. or a population density 150–300 inhabitants/km sq.	Espon database
Urban regions	With a city of between 150,000 and 300,000 inhabitants and a population density 150–300 inhabitants/km sq. (or a smaller population density – 100–150 inh./km with a bigger centre (>300,000) or a population density between 100 and 150 inh./km sq).	Espon database
Rural regions	With a population density < 100/km sq. and a centre > 125,000 inh. or a population density < 100/km sq. with a centre < 125,000	Espon database

– a soft and private element of territorial capital, namely the degree of innovation of regions (*inno*), expected to affect positively the regional growth rates, as a large body of literature suggests.[5] Innovation was proxied by the share of human resources in science and technology;
– a hard element of the territorial capital: the transport infrastructural endowment of regions, which ought be positive but may also be negative if this variable

[5] For a review of the role of innovation in regional growth, see Howells 2005; Johansson and Karlsson 2009; de Groot et al. 2009; Audretsch and Aldridge 2009.

measures the density of roads and congestion effects prevail.[6] This aspect was measured by the endowment of roads per square kilometre;

– regional specialisation in high-value functions. In a period of globalization, it is to be expected that the higher the functions that a region performs, the higher its growth rate.[7] These functions were approximated by the share of high-value service functions (i.e. share of corporate managers) (*funct*) reported by the labour force survey;
– a mixed (hard/soft) element of territorial capital: agglomeration economies, which were captured with a dummy measuring the settlement structure of regions (*Daggec*). In particular, agglomeration economies were proxied by dummies measuring the presence in regions of dense and large cities. Specifically, use was made of two different dummies built on different thresholds of densities and sizes of cities;
– last, but not least, an important variable explaining regional differential growth is the presence of public funds (*pol*) which, because they are aimed at either demand-side support or supply-side development, should yield positive growth effects. We used structural funds expenditure per capita as a proxy for this factor

In addition to the endogenous growth factors, we take into account two main exogenous factors, i.e. the national trends and the regional specific globalization trends, by using the following variables:

– the national growth rate (*natgrowth*), which measures all the national factors with an equal impact for all regions of the same country. In order to avoid endogeneity, this national growth rate was calculated using only the GDP of the other regions of the country to which the region belonged.[8] We expected national factors to positively influence regional growth;
– FDI penetration in a region as a measure of regional attractiveness and global flows (*fdi*). We consider total FDI as well as separate between FDI originating from within Europe and FDI originating out of Europe[9]:

The model estimated is therefore the following:

$$\begin{aligned} regrowth_r = \alpha_0 + \beta_1 natgrowth_r + \beta_2 inno_r + \beta_3 spill_r + \beta_4 in_r + \beta_5\, pol_r + \\ + \beta_6\, funct_r + \beta_7\, fdi_r + \alpha_1 D_{aggec_r} + \varepsilon_r \end{aligned} \tag{2}$$

[6] For a review of the role of transport infrastructure in regional growth, see Bröcker and Rietveld 2009.

[7] On the role of functions in regional growth, see Capello et al. 2011b.

[8] This had the drawback of eliminating from the regressions countries which have only one NUTS 2 region, namely Luxembourg, Cyprus, Malta, Estonia, Latvia, Lithuania. Nevertheless, this decreased the sample by only 6 observations.

[9] This is possible thanks to the FDI-Regio database, kindly provided us by Laura Resmini of Bicocca University of Milan.

Table 2 Growth factors for regions in a period of globalization

	Model0 stand. coeff.	p-value	sig.	Model1 stand. coeff.	p-value	sig.	Model1 stand. coeff.	(Western regions) p-value	Sig.
National growth rates of all other regions in the country 2002–05	0.774	0.00	***	0.784	0	***	0.603	0.00	***
Share of science and technology employment 2000	0.147	0.019	**	0.142	0.026	**	0.055	0.413	
Growth spllovers 1999–2000	−0.031	0.00	***	−0.028	0.001	***	−0.050	0.00	***
Total infrastructure on sqm 2000	−0.140	0.00	***	−0.140	0	***	−0.227	0.00	***
Structural funds per capita 1994–1999	0.047	0.046	**	0.053	0.015	**	0.071	0.008	***
High level functions (share of private managers) 1999–2001	0.063	0.164		0.055	0.219		0.226	0.005	***
Urban dummy	0.104	0.00	***	0.114	0	***	0.104	0.005	***
Total FDI penetration rate 2001–03				0.071	0.049	**	0.047	0.339	
Constant		0.191			0.127			0.724	
Obs	246			246			195		
R2	0.5938			0.5984			0.4927		
F	43.3			42.29			34.84		
Moran's I	1.678	0.093	*	1.288	0.198		3.282	0.001	***
Spatial error									
Lagrange multiplier	0.146	0.702		0.005	0.941		1.808	0.179	
Robust lagrange multiplier	0.797	0.372		0.243	0.622		0.222	0.637	
Spatial lag									
Lagrange multiplier	0.446	0.504		0.363	0.547		2.053	0.152	
Robust lagrange multiplier	1.097	0.295		0.6	0.438		0.468	0.494	

*** $p < 1$ %; ** $p < 5$ %; * $p < 10$ %. Independent variable: regional growth rates 2002–05. Coefficients are standardized

The results of the regression model are reported in Table 2.[10] In the first reported model (Model 0, for reasons which will be evident below), which has all variables except from the settlement structure, all coefficients have the expected sign, apart

[10] We tested for spatial autocorrelation our regression models, but, due to the fact that among the regressors spatial spillovers and country effects are present, the spatial tests turned out as non significant so that the reported standardized coefficients are those of OLS with robust standard errors.

from infrastructure which has a negative one, and all coefficients are highly significant, with the exception of high-value functions which are positive and (nearly) significant. In particular, the coefficient of the national growth rate is positive and, being close to 0.8, the highest, implying that exogenous national factors are very important and being part of a country which grows 1 % faster imply 0.8 % faster growth rate for a region ceteris paribus, i.e. independently from the regional endogenous characteristics.

The human resources in science and technology are also positive and significant, meaning that the hypothesis that they are needed for regions to thrive in an age of globalization is confirmed.

Growth spillovers are on the contrary negative and significant, meaning that being close to strong and fast growing regions has more negative effects due the presence of strong competitors nearby than positive effects due to induced demand.

Somehow puzzling, infrastructure endowment has a negative and significant coefficient, probably due to the fact that road endowment is not able to capture the smoothness of traffic on these roads, but rather tends to capture the excessive density of some areas, which are hence subject to congestion diseconomies.

Public policy support has a small but positive and significant coefficient, implying that within their countries, the most assisted regions take benefit of this assistance ceteris paribus.

High level functions are not significant at 0.1 threshold. However, their coefficient is positive, quite stable (as we will see in the rest of the Section) and has a p-value which is only slightly higher than 0.1. For this reason it is possible to keep this variable in the regressions, also because of the theoretical importance of this variable for the globalization processes.

The last variable which is used in this general regression model is the amount of total FDI on population received by regions, whose coefficient is positive and significant, meaning that being able to attract FDI has a growth effect in a period of globalization.

To the first model, we added the dummy capturing the settlement structure of the regions (*Daggec)*, which turns out to be significant without altering the other coefficients, nor their significance. In particular, it turns out in Model 1 that the "urban" regions (i.e. the intermediate ones in terms of density and presence of large cities[11]) have outperformed the rest of the regions ceteris paribus. Probably, the most "agglomerated" regions suffer for decreasing agglomeration economies, if not congestion diseconomies, which are still not present in the intermediate category. On the contrary, the more sparsely populated "rural" regions, also due to the absence of large cities within, are unable to reach the critical mass needed to generate agglomeration economies and, consequently, growth.

[11] Urban regions are defined as those regions with a city of between 150,000 and 300,000 inhabitants and a population density 150–300 inhabitants/km sq. (or a smaller population density – 100–150 inh./km with a bigger centre (>300,000) or a population density between 100 and 150 inh./km sq).

As a technical but important note, being regional growth the dependent variable, the models have been tested for spatial effects using different matrices, including standardized distance matrix and a standardized distance matrix with a threshold. All test reject the presence of spatial autocorrelation in the regressions, and the need to use a spatial lag or spatial error model. This is likely to be due to two concomitant effects: on the one hand the regressions have an explicit growth spillover coefficient inside, i.e. some sort of spatial lag; on the other hand, the fact that the regressions include the national growth rate is another way in which growth in neighbouring regions is taken into account.

To test whether our hypothesis that more advanced stages of development imply a more important role of endogenous factors, we re-run model 1 in the sub-sample of Western regions.[12] The results are reported in Table 2.

It is immediately evident that the exogenous factors are clearly less important for the richest Western regions with respect to the poorest Eastern ones, as evidenced by the fact that the national growth coefficient is still highly significant but considerably lower, and the fact that the coefficient for FDI is lower and much less significant (and this is in line with the findings of Pietrobelli et al., in this volume).

It is also evident that the coefficient of science and technology employment is no longer significant, but we interpret this result as evidence that beyond R&D personnel, which is quite diffused in Western Europe, what matter are the socio-economic conditions for regional innovation (see also the contribution by Rodríguez-Pose and Comptour in this volume).

In order to go more in depth into the role of exogeous factors in regional growth, the next section will investigate the growth effects of different types of FDI, first among all the European regions and then in the sub-sample of Western regions.

5 The Effects of FDI by Source, Sector and Technological Level

FDI are not homogeneously distributed over space, they belong to different sectors, and vary in terms of origin and technological development. Given their important role in explaining regional growth, a more detailed analysis is worth inspection.

In order to test all differential effects in FDI, starting from Model 1 of Sect. 4 a number of FDI typologies were regressed to see if they have a more or less important role in regional growth with respect to generic FDI. The results are presented in Table 3, where the first two models are in order to differentiate between intra-European and extra-European FDI.

First it has to be observed that the coefficients and the significance of the other regressors are stable, including the endogenous territorial capital factors and the exogenous national growth. Concerning FDI, those coming from intra-EU appear to have a very similar coefficient with respect to those coming from extra-EU and the

[12] It is on the contrary not possible to run a similar regression on Eastern regions only because of lack of degrees of freedom.

Table 3 The effects of FDI on regional growth by source, sector and technological level

	Model1 stand. coeff.	p-value	Sig.	Model2 stand. coeff.	p-value	Sig.	Model3 stand. coeff.	p-value	Sig.	Model4 stand. coeff.	p-value	Sig.	Model5 stand. coeff.	p-value	Sig.	Model6 stand. coeff.	p-value	Sig.	Model7 stand. coeff.	p-value	Sig.
National growth rates of all other regions in the country 2002–05	0.784	0	***	0.789	0	***	0.786	0	***	0.801	0	***	0.784	0	***	0.800	0	***	0.746	0	***
Share of science and technology employment 2000	0.142	0.026	**	0.144	0.027	**	0.146	0.019	**	0.148	0.018	**	0.144	0.022	**	0.150	0.017	**	0.155	0.013	**
Growth spllovers 1999–2000	−0.028	0.001	***	−0.023	0.011	**	−0.030	0.001	***	−0.028	0.001	***	−0.022	0.032	**	−0.022	0.024	**	−0.028	0.002	***
Total infrastructure on sqm 2000	−0.140	0	***	−0.156	0.001	***	−0.119	0.003	***	−0.127	0.001	***	−0.131	0.001	***	−0.113	0.003	***	−0.134	0.001	***
Structural funds per capita 1994–1999	0.053	0.015	**	0.054	0.015	**	0.054	0.012	**	0.054	0.014	**	0.056	0.009	***	0.056	0.012	**	0.058	0.006	***
High level functions (share of private managers) 1999–2001	0.055	0.219		0.063	0.167		0.056	0.215		0.055	0.248		0.081	0.071	*	0.030	0.632		0.061	0.167	
Urban dummy	0.071	0.049	**	0.072	0.05	*	0.066	0.065	*	0.066	0.065	*	0.065	0.067	*	0.064	0.076	*	0.071	0.051	*
Total FDI penetration rate 2001–03	0.114	0	***																		
Total intra-European FDI penetration rate 2001–03				0.103	0.038	**															
Total extra-European FDI penetration rate 2001–03							0.104	0	***												
High value-added service FDI penetration rate 2001–03										0.075	0.007	***									
Low value-added service FDI penetration rate 2001–03													0.101	0.001	***						
High value-added manufacturing FDI penetration rate 2001–03																0.074	0.149				

(continued)

Table 3 (continued)

	Model1 stand. coeff.	p-value	Sig.	Model2 stand. coeff.	p-value	Sig.	Model3 stand. coeff.	p-value	Sig.	Model4 stand. coeff.	p-value	Sig.	Model5 stand. coeff.	p-value	Sig.	Model6 stand. coeff.	p-value	Sig.	Model7 stand. coeff.	p-value	Sig.
Low value-added manufacturing FDI penetration rate 2001–03																			0.149	0.002	***
Constant		0.127			0.107			0.111			0.095	*		0.097	*		0.088	*		0.092	*
Obs	246			246			246			246			246			246			246		
R2	0.5984			0.5953			0.5977			0.5923			0.5974			0.5912			0.6061		
F	42.29			36.42			56.5			35.49			37.34			35.6			41.5		
Moran's I	1.288	0.198		1.679	0.093	*	1.137	0.256		1.601	0.109		0.979	0.327		1.667	0.095	*	2.158	0.031	**
Spatial error																					
Lagrange multiplier	0.005	0.941		0.124	0.724		0.001	0.973		0.107	0.743		0.023	0.878		0.113	0.737		0.432	0.511	
Robust lagrange multiplier	0.243	0.622		0.537	0.464		0.113	0.737		0.275	0.6		0.082	0.774		0.358	0.55		1.873	0.171	
Spatial lag																					
Lagrange multiplier	0.363	0.547		0.202	0.653		0.315	0.574		0.034	0.854		0.493	0.482		0.074	0.786		0.786	0.375	
Robust lagrange multiplier	0.6	0.438		0.615	0.433		0.427	0.513		0.202	0.653		0.552	0.457		0.319	0.572		2.228	0.136	

*** $p < 1\ \%$; ** $p < 5\ \%$; * $p < 10\ \%$. Independent variable: regional growth rates 2002–05. Coefficients are standardized significance of these coefficients is high for both but higher for extra-EU FDI (models 2 and 3 in Table 3). This implies that both types of FDI are significantly able to help regional growth, though with different intensity.

One could also analyze the effects of service FDI, which have increased their importance over the last few years. In Models 4 and 5 service FDI are divided into high value added and low value added, with the first ones being more of support to production and/or producing services which can be exported, whereas the low value added ones are those related to the personal services and retail.[13] It turns out that the second coefficient is slightly higher and also slightly more significant, but both types of services FDI are relevant in explaining regional growth.

The same distinction between high level and low level has been done for manufacturing FDI (models 6 and 7 in Table 3), which have been divided in high-tech and low-tech using the Pavitt classification.[14] The results show an interesting result: differently from the service sector, high-tech manufacturing FDI, though having a positive coefficient, do not have a statistically significant impact on regional growth, whereas the coefficient for low-tech FDI is positive and significant. Among the different explanations for this apparently counterintuitive result, one can recall that high tech manufacturing FDI in Europe are very limited in their number and that manufacturing FDI in general are a phenomenon typical of Eastern countries, where they induce growth and what matter is more their quantity than their specialization; manufacturing FDI are less common and less important in the West, providing further support to the possibility that the role of exogenous growth factors is more important for regions at a lower stage of development.

This calls for a more specific analysis on Western vs. Eastern regions, in order to disentangle their possible different behaviours as far as FDI are concerned. The number of regions in the East is insufficient to provide regressions for them alone, but by analysing the behaviour of Western regions it is possible to induce that any difference with respect to the total sample is due to Eastern regions.[15]

Table 4 hence reports the same regressions of Table 3 performed on the sub-sample of European countries which belong to the Old 15 members of the EU. The two tables are rather similar, spatial effects are still not present and the significance of the general regression coefficients is also normally the same.

In addition to the far lower coefficient for national growth which was already pointed out in Sect. 5, the only noticeable differences regard the significance of the dummy for urban areas, which is no longer significant, and the share of employment

[13] In particular, High-value service FDI are those of Ateco 1.1 sectors I (Transport, storage and communication), J (Financial intermediation) and K (Real estate, renting and business activities). The other service sectors are included in low-value.

[14] In particular, high-tech manufacturing FDI are those in sectors classified high-tech or medium-high tech by Pavitt (1984), with the other sectors (Pavitt medium-low tech and low-tech) composing low-tech FDI. Notice that the results are consistent when using only the high-tech and the low-tech of Pavitt.

[15] We also performed a Chow test which, due to the high significance of national growth in regional growth, did not identify a differentiation of growth model between the two.

Table 4 The effects of FDI on regional growth by source, sector and technological level in western regions

	Model1 stand. coeff.	p-value	Sig.	Model2 stand. coeff.	p-value	Sig.	Model3 stand. coeff.	p-value	Sig.	Model4 stand. coeff.	p-value	Sig.	Model5 stand. coeff.	p-value	Sig.	Model6 stand. coeff.	p-value	Sig.	Model7 stand. coeff.	p-value	Sig.
National growth rates of all other regions in the country 2002–05	0.603	0	***	0.593	0	***	0.613	0	***	0.609	0	***	0.598	0	***	0.612	0	***	0.598	0	***
Share of science and technology employment 2000	0.055	0.413		0.047	0.481		0.064	0.34		0.054	0.42		0.053	0.431		0.066	0.333		0.069	0.307	
Growth spllovers 1999–2000	−0.050	0	***	−0.044	0.001	***	−0.051	0	***	−0.051	0	***	−0.044	0.003	***	−0.043	0.008	***	-0.046	0.001	***
Total infrastructure on sqm 2000	−0.227	0	***	−0.252	0	***	−0.201	0.001	***	−0.219	0	***	−0.245	0	***	−0.196	0.001	***	-0.207	0.001	***
Structural funds per capita 1994–1999	0.071	0.008	***	0.071	0.009	***	0.073	0.006	***	0.071	0.008	***	0.072	0.007	***	0.077	0.005	***	0.076	0.005	***
High level functions (share of private managers) 1999–2001	0.226	0.005	***	0.243	0.002	***	0.226	0.007	***	0.230	0.004	***	0.249	0.001	***	0.240	0.015	***	0.239	0.004	***
Urban dummy	0.047	0.339		0.047	0.338		0.042	0.389		0.046	0.34		0.045	0.36		0.036	0.464		0.038	0.435	
Total FDI penetration rate 2001–03	0.104	0.005	***																		
Total intra-European FDI penetration rate 2001–03				0.117	0.018	**															
Total extra-European FDI penetration rate 2001–03							0.074	0.006	***												
										0.100	0.001	***									

High value-added service FDI penetration rate 2001–03																					
Low value-added service FDI penetration rate 2001–03													0.101	0.089	*						
High value-added manufacturing FDI penetration rate 2001–03																0.027	0.679				
Low value-added manufacturing FDI penetration rate 2001–03																			0.052	0.407	
Constant		0.724			0.665			0.828			0.714			0.741			0.895			0.894	
Obs	195			195			195			195			195			195			195		
R2	0.4927			0.494			0.4894			0.4928			0.4922			0.4858			0.4872		
F	34.84			30.84			44.28			34.39			25.18			22.99			24.17		
Moran's I	3.281	0.001	***	3.384	0.001	***	3.14	0.002	***	3.05	0.002	***	3.499	0	***	3.085	0.002	***	3.505	0	***
Spatial error																					
Lagrange multiplier	1.808	0.179		1.923	0.165		1.601	0.206		1.446	0.229		2.087	0.149		1.381	0.24		2.117	0.146	
Robust lagrange multiplier	0.222	0.637		0.136	0.713		0.278	0.598		0.136	0.713		0.189	0.663		0.161	0.688		0.48	0.489	
Spatial lag																					
Lagrange multiplier	2.053	0.152		2.508	0.113		1.616	0.204		1.795	0.18		2.549	0.11		1.605	0.205		1.878	0.171	
Robust lagrange multiplier	0.468	0.494		0.721	0.396		0.292	0.589		0.484	0.487		0.652	0.42		0.385	0.535		0.241	0.623	

*** $p < 1$ %; ** $p < 5$ %; * $p < 10$ %. Independent variable: regional growth rates 2002–05. Coefficients are standardized in science and technology which is less significant with respect to high-value added functions.

Our main interest lies here in FDI coefficients: first, the standardized coefficient for intra-European FDI is now, for Western regions, larger with respect to the extra-European one, though the latter remains more significant (models 2 and 3 of Table 4). Even more interesting, the high-value service FDI increase their coefficient and their significance, whereas low-value-added service FDI have about the same coefficient but lose significance (models 5 and 6 of Table 4).

It hence appears that manufacturing FDI is less important within Western countries, as supported by the observation that high value added manufacturing FDI have a lower and much less significant coefficient and, even more strikingly, low value added ones were significant overall are now no longer significant for the Old15 regions, signalling that the growth rate effects of manufacturing foreign investments is mainly a feature of countries in transition or restructuring (models 6 and 7 of Table 4).

6 Conclusions

This paper has analyzed the growth performance of European regions in a period of fast globalization, focusing on the importance of endogenous and exogenous factors in different stages of development, indentified by the belonging to Old15 or New12 countries.

It first emerges that regional disparities have a clearly different pattern in the two groups, with the decrease of EU-wide disparities due to the decline of the East–west divide, but with between country disparities nearly stable in the West and clearly increasing in the East.

Our explicative analysis analyzed the role of endogenous and exogenous factors of regional growth, using many assets of territorial capital as the endogenous ones and as exogenous ones national factors and an internationally, globalization-related one, namely FDI.

A first result which emerges is that despite the strong integration process which is in place in Europe since many years, the national component of growth plays an important role in the explanation of regional growth. This result is strong in both the East and the West, but considerably stronger in the former, characterised by a lower stage of development.

A second important result is that the capacity of a region to grow depends on both territorial capital success factors, and the presence of foreigner direct investments. The latter appear to be more significant for Eastern regions, which again show a higher growth dependency on exogenous factors. Once the analysis is developed at sectoral level, it turns out that the presence of FDI in manufacturing low-tech sectors plays a significant role on growth in Eastern regions whereas manufacturing FDI does not play any role in explaining growth patterns in Western country regions.

Our intuition that exogenous vs. endogenous forces of development play a different role according to the stage of development of regions finds solid empirical support; also in a period of strong globalization, endogenous factors explain regional growth trajectories in areas characterised by a more advanced stage of development.

References

Affuso A, Capello R, Fratesi U (2011) Globalization and competitive strategies in European vulnerable regions. Reg Stud 45(5):657–675

Alesina A, La Ferrara E (2000) Participation in heterogeneous communities. Q J Econ 65(3):847–904

Audretsch D, Aldridge T (2009) Knowledge spillovers, entrepreneurship and regional development. In: Capello R, Nijkamp P (eds) Handbook of regional dynamics and growth: advances in regional economics. Edward Elgar, Cheltenham, pp 201–210

Baldwin R (1995) A domino theory of regionalism. In: Baldwin R, Haaparnata P, Kiander J (eds) Expanding membership of the European union. Cambridge University Press, Cambridge, pp 25–53

Baldwin R, Venables A (1995) Regional economic integration. In: Grossman G, Rogoff K (eds) Handbook of international economics, vol III. North Holland, Amsterdam, pp 1597–1644

Bergoeing R, Kehoe T, Strauss-Kahn V, Yi K (2004) Why is manufacturing trade rising even as manufacturing output is falling? Am Econ Rev 94:134–138

Beugelsdijk S, van Schaik T (2005) Social capital and growth in European regions: an empirical test. European J Pol Econ 21(2):301–324

Bowles S, Gintis H, Osborne M (2001) The determinants of earnings: a behavioral approach. J Econ Lit 39(4):1137–1176

Bröcker J, Rietveld P (2009) Infrastructure and regional development. In: Capello R, Nijkamp P (eds) Handbook of regional dynamics and growth: advances in regional economics. Edward Elgar, Chelthenham, pp 152–181

Camagni R (ed) (1991a) Innovation networks: spatial perspectives. Belhaven-Pinter, London

Camagni R (1991b) Regional deindustrialization and revitalization processes in Italy. In: Sazanami H, Rodwin L (eds) Industrial change and regional economic transformation. Harper Collins, London, pp 137–167

Camagni R (1999) La ville comme Milieu: De l'application de l'approche GREMI à l'évolution urbaine. Revue d'Economie Régionale et Urbaine 3:591–606

Camagni R (2002) On the concept of territorial competitiveness: sound or misleading? Urban Stud 13:2395–2412

Camagni R (2009) Territorial capital and regional development. In: Capello R, Nijkamp P (eds) Handbook in regional growth and development theories. Edward Elgar, Cheltenham, pp 118–132

Capello R, Caragliu A, Nijkamp P (2011a) Territorial capital and regional growth: increasing returns in knowledge use. Tijdschrift voor Economische en Sociale Geographie (TESG) 102 (4):385–405

Capello R, Fratesi U, Resmini L (2011b) Globalisation and regional growth in Europe: past trends and scenarios. Springer Verlag, Berlin

Caragliu A (2009) Innovation, territorial capital and regional growth. Ph.D. thesis, Politecnico di Milano

CEC – European commission (2005) Europe 2020, Brussels

Clark C, Wilson F, Bradley J (1969) Industrial location and economic potential in Western Europe. Reg Stud 3(2):197–212

De Groot H, Poot J, Smit M (2009) Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In: Capello R, Nijkamp P (eds) Handbook of regional dynamics and growth: advances in regional economics, Edward Elgar, Cheltenham, pp 256–281

Dunning J (2001) The eclectic (OLI) paradigm of international production: past, present and future. Int J Econ of Bus 8(2):173–190

Fiorentino R, Verdeja L, Toqueboeuf C (2007) The changing landscape of regional trade agreements: 2006 update. WTO, discussion paper no 12

Fratesi U (2012) A globalization-based taxonomy of European regions. Regional Science Policy and Practice 4(1):1–23

Friedmann J (1986) The world city hypothesis. Develop Change 17:69–83

Glaeser E, Laibson D, Sacerdote B (2002) An economic approach to social capital. Econ J 112:437–458, November

Glaeser E, Laibson D, Scheinkman J, Soutter C (2000) Measuring trust. Q J Econ 65(3):811–846

Hanson G, Mataloni R, Slaughter M (2005) Vertical production network in multinational firms. Rev Econ Stat 87(4):664–678

Hansen H, Rand J (2006) On the causal links between FDI and growth in developing countries. The World Economy 29(1):21–41

Howells J (2005) Innovation and regional economic development: a matter of perspective? Res Pol 34(8):1220–1234

Hummels D, Ishii J, Yi K (2001) The nature and growth of vertical specialization in world trade. J Int Econ 54(1):75–96

Hummels D, Rapoport D, Yi K (1998) Vertical specialization and the changing nature of world trade". Federal Reserve Bank of New York Economic Policy Review 4(2):79–99

Knack S, Keefer P (1997) Does social capital have an economic pay-off? A cross-country investigation. Q J Econ 112:1251–1288

Kucera D, Milberg W (2003) Deindustrialization and changes in manufacturing trade: factor content calculations for 1978–1995. Rev World Econ 139:601–624

Johansson B, Karlsson C (2009) Knowledge and regional development. In: Capello R, Nijkamp P (eds) Handbook of regional dynamics and growth: advances in regional economics. Edward Elgar, Cheltenham, pp 239–255

La Porta R, Lopez-de-Salanes F, Shleifer A, Vishny R (1997) Trust in large organisations. American economic review papers and proceedings, vol 87, pp 333–8

Lall S, Narula R (2004) Foreign direct investment and its role in economic development: do we need a new agenda? The European J Dev Res 16:447–464

Lazear E (1999) Culture and language. J Pol Econ 107(6):S95–S126, part 2

Lucas R (2008) International labour migration in a globalizing economy. Carnegie paper no 92

McCloskey D, Klamer A (1995) One third of GDP is persuasion, The American economic association papers and proceedings, pp 191–195

Moran T, Graham E, Blomström M (eds) (2005) Does foreign direct investment promote development? Institute for International Economics, Washington

OECD (2001) Territorial outlook. OECD, Paris

Panagariya A (1999) The regionalism debate: an overview. World Econ 22(4):477–511

Panagariya A (2000) Trade liberalization: the traditional theory and new developments. J Econ Lit 38(2):287–331

Pavitt K (1984) Sectoral patterns of technical change: towards a taxonomy and a theory. Res Pol 13:343–373

Resmini L (2007) Regional patterns of industry location in transition countries: does economic integration with the EU matter? Reg Stud 41(6):747–764

Resmini L (2008) The location of multinational enterprises in central and eastern European countries. In: Krieger-Boden C, Morgenroth E, Petrakos G (eds) The impact of European integration on regional structural change and cohesion. Routledge, London, pp 266–284

Soubbotina T (2004) Beyond economic growth, chapter 12. Globalization: international trade and migration. The World Bank, Washington, DC, pp 83–94

UNCTAD (1997) World investment report. UN, Washington

UNCTAD (2004) World investment report. U.N., New York/Geneva

UNCTAD (2008) World investment report: transnational corporations and the infrastructure challenge. U.N., New York/Geneva

UNCTAD (2010) World investment report: investing in a low carbon economy. U.N., New York/Geneva

Williamson JG (1965) Regional inequality and the process of national development: a description of the patterns. Econ Dev Cultural Change 13:3–45

Winters A (1996) Regionalism vs. multilateralism. The World Bank policy research working paper no 1697. The World Bank, Washington, DC

Yi K (2003) Can vertical specialization explain the growth of world trade. Polit Econ 111(1): 52–102

TFP Convergence Across European Regions: A Comparative Spatial Dynamics Analysis

Adriana Di Liberto and Stefano Usai

Abstract This paper proposes a fixed-effect panel methodology to estimate Total Factor Productivity and investigate the spatial dimension of regional EU TFP from both a static and a dynamic perspective. The sample includes 199 regions in EU15 (plus Norway and Switzerland) between 1985 and 2006. First of all, we find the absence of an overall process of convergence, since TFP dispersion is virtually constant along time. Furthermore, exploratory spatial data techniques show that there are interesting interregional dynamic patterns. We find that polarization patterns in Europe have significantly changed along time. Overall, results suggest that only few TFP leaders are emerging and they are distancing themselves from the rest, while the cluster of low TFP regions is widening.

Keywords TFP • Technology catching up • Panel data • Exploratory spatial data analysis

1 Introduction

Even if a large body of economic theory suggests that differences in estimated Total Factor Productivity (TFP) should imply flows of technology moving from advanced to less developed areas, data often reveal that these diffusion processes are neither effortless nor instantaneous. Consequently, a persistent gap in the rate of technology adoption is observed together with a weak (or absent) process of absolute convergence in income per capita (Pritchett 1997; Durlauf et al. 2005; Grier and Grier 2007). Even more puzzling is the evidence that slow processes of technology adoption are observed also across similar leading countries of the world economy

JEL code: C23, O47, O33, R11

A. Di Liberto (✉) • S. Usai
Università di Cagliari and Crenos, via S. Ignazio 78, 09123 Cagliari, Italy
e-mail: diliberto@unica.it

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_3,

(Comin and Hobijn 2004 and Comin et al. 2008) or across regions within the same country or within union of states (see Magrini 2004, for a review).

This evidence is usually explained by differences in human capital stocks, as firstly suggested by Nelson and Phelps (1966) and/or by barriers to technology adoption due to monopoly rights of various forms, as in Parente and Prescott (1999). More recently, the debate focused on the role of other possible important and strictly related[1] productivity determinants such as institutional quality heterogeneity (Hall and Jones 1999; Acemoglu et al. 2001, 2006; Comin and Hobijn 2009), social capital (Knack and Keefer 1997 among the many others) and cultural aspects (Easterly and Levine 1997; Alesina and La Ferrara 2005). As emphasised by various contributions in this book (Crescenzi et al. 2012; Bellini et al. 2012) these factors represent important determinants of technological progress and of innovation diffusion at the regional level too.

This paper does not focus on productivity and/or innovation determinants, an issue that is extensively discussed in other contributions of this book (Capello and Fratesi 2012 and Rodriguez-Pose 2012, among others), but rather on TFP di/convergence across European regions. TFP heterogeneity and dynamics are assessed with the use of some innovative tools.

First of all, we use a recently developed technique to estimate TFP levels for EU regions. The former has been firstly suggested by Islam (1995) and further extended by Di Liberto et al. (2008) and exploits the properties of fixed-effects panel estimators to obtain TFP measures.[2] This approach is particularly suited to convergence dynamics analysis since it enables us to disentangle the two potential components of observed convergence processes: technological catching up and capital deepening. Consequently, it also limits the otherwise likely risk of overstating the role of TFP dynamics within convergence processes.

Secondly, we investigate the underlying spatial dimension of regional TFP from both a static and a dynamic perspective following some recent contributions in Exploratory Spatial Data Analysis (ESDA) by Rey et al. (2010) and Rey and Ye (2010). In particular, we examine if, and how, geographical proximity can influence divergence and convergence trends as a result of either agglomeration or dispersion forces which operate at the local level.

We use data on per capita Value Added (VA) of 199 regions in 17 countries (EU15 plus Norway and Switzerland) over the period 1985–2006. It is worth underlining that this time span includes the decade characterized by the Information and Technology (IT) revolution, which is known to be the source of a significant asymmetric shock on productivity levels, with the more developed economies as

[1] On this see Tabellini (2010).

[2] A large array of methodologies is currently available to estimate TFP and none has emerged as a recognized standard. See Del Gatto et al. (2011) for a recent survey.

the major beneficiaries. We choose EU15 + 2, instead of EU27,[3] for a two main reasons. Firstly, time series availability is restricted to a shorter period for the regions of the new accession countries. Secondly, a preferential preliminary test of the catching up hypothesis requires a scenario which is ideal for technology transfer, that is, with a homogeneous institutional and economic setting.

Our results confirm that cross-region gaps in TFP levels are significant, that they are persistent, and that they are an important component of Value Added (VA) per capita dynamics. In particular, we do not observe a process of global convergence in TFP, as it is not detected in VA per capita. At the same time this does not imply the absence of cross-region dynamics in TFP. During the two decades under examination, we notice the presence of strong intra-distribution movements with significant changes in regional rankings and cluster composition. Such clusters prove to have an important geographic component since global and local spatial dependence in TFP levels is found. Finally, thanks to new visualisation techniques we highlight that polarisation patterns have changed profoundly along time.

The paper is organized as follows. In Sect. 2 we describe our methodology for the estimation of TFP levels at different points in time. Sections 3 and 4 examine the role of space in the evolution of regional TFP distributions by means of some traditional (Sect. 3) and some new and advanced tools (Sect. 4) of exploratory spatial data analysis (ESDA). Conclusions are in Sect. 5.

2 A Panel Data Approach to Estimate TFP Convergence

In the last 20 years, the study of the convergence process at the regional and international levels has been the centre of interest of regional science and growth literature. Most studies have focused on income instead of productivity differentials, mostly because the former are easier to measure. Nonetheless, according to Easterly and Levine (2001), the study of the productivity dynamics is essential whenever we aim at explaining economic inequality across countries and regions. TFP provides probably the best definition of the economic notion of productivity since it describes how efficiently each region transforms physical capital and labour into output.

There are just a few studies on the economic growth of European regions based on the analysis of TFP determinants (see Dettori et al. 2011; Marrocu et al. 2012; Scherngell et al. 2007; Le Sage and Fischer 2012), and to the best of our knowledge only one focuses on convergence, even though with a different methodology (Ezcurra et al. 2009).

Measuring productivity or TFP in empirical analysis is not an easy task as there is no consensus on the best way to measure it.[4] Islam (1995) was among the first to suggest to investigate cross-country (or region) TFP heterogeneity by using a fixed-effect panel

---

[3] EU27 TFP dynamics is analysed in Marrocu et al. (2012) whilst regions of the EU10 group are analysed in this book by Monastiriotis 2011.

[4] See Del Gatto et al. (2011).

estimator.[5] In particular, the author extended the standard Mankiw et al. (1992) structural approach by allowing TFP levels to vary across individual economies, together with saving and population growth rates. Following Di Liberto et al. (2011) we use this approach to firstly obtain TFP measures for EU regions in different periods of time and then test for the presence of technological catching up.

Our period of analysis, from 1985 to 2006, includes some years which have been strongly influenced by the introduction of IT technologies. In terms of TFP convergence, such years are important since the development of IT has meant "... a rapidly rising source of aggregate productivity growth throughout the 1990s".[6] More precisely, we use different datasets (Cambridge Econometrics and Eurostat) to estimate the following equation:

$$\tilde{y}_{it} = \beta \tilde{y}_{it-\tau} + \sum_{j=1}^{2} \gamma_j \tilde{x}_{j,it} + \mu_i + u_{it} \tag{1}$$

$$\tilde{y}_{it} = y_{it} + \bar{y}_t, \quad \tilde{x}_{it} = x_{it} + \bar{x}_t \tag{2}$$

Equation 1 represents a standard convergence equation (Islam 1995), where the dependent variable is the logarithm of per capita VA (measured in terms of population working age), u is the transitory term that varies across countries and time and $\bar{y}_t$ and $\bar{x}_t$ are the EU regional averages in period t. In other words, data are taken in difference from the sample mean in order to control for the presence of a time trend component η_t and of a likely common stochastic trend (the common component of technology) across countries. The remaining terms are defined as follows:

$$x_{1,it} = \ln(s_{it}) \tag{3}$$

$$x_{2,it} = \ln(n_{it} + g + \delta) \tag{4}$$

$$\gamma_1 = (1 - \beta)\frac{\alpha}{1 - \alpha} \tag{5}$$

$$\gamma_2 = -(1 - \beta)\frac{\alpha}{1 - \alpha} \tag{6}$$

$$\mu_i = -(1 - \beta)\ln A(0)_i \tag{7}$$

$$\eta_t = g(t_2 - \beta t_1) \tag{8}$$

[5] See also Caselli et al. (1996) and Islam (2003) among others.

[6] See Jorgenson (2005).

$A(0)_i$ represents the initial level of technology, and s, n, δ are the saving rate, the population growth rate and the depreciation rate, respectively; g is the exogenous rate of technological change,[7] assumed to be invariant across individual economies; α is the usual capital share of a standard Cobb-Douglas production function; finally, $\beta = e^{-\lambda\tau}$, where $\lambda = (1-\alpha)(n+g+\delta)$ represents the convergence parameter and $\tau = t_2 - t_1$ is the time span considered.

Estimates from Eq. 1 enable us to separate out the neoclassical convergence mechanism (through the lambda coefficient) from the technological catching up hypothesis (captured by $A(0)_i$ computed in different time periods). This is important since, as stressed by Bernard and Jones (1996), we often do not know "...how much of the convergence that we observe is due to convergence in technology versus convergence in capital-labour ratios". Most empirical analyses do not discriminate between these two mechanisms and are therefore likely to overstate the role of TFP dynamics within convergence process analysis.

We use a 3-year time span in order to control for business cycle fluctuations and serial correlation, which are likely to affect the data in the short run. As a result, we obtain a sample with T = 8, which is the longest possible one.[8] Moreover, we augment Eq. 1 to take into account differences in human capital across regions and we take all regressors at their t-3 level to control for likely endogeneity problems. Our regional TFP index are obtained through the following formula based on fixed effects μ_i:

$$A(0)_i = \exp\left(\frac{\mu_i}{1-\beta}\right) \tag{9}$$

Finally, the computation of regional TFP values over subsequent periods, which is presented in the next section, enables us to test whether the observed time pattern is consistent either with the catch-up hypothesis or with the alternative hypothesis that the current level of technology heterogeneity across regions is constant or even increasing over time.

Equation 1 represents a dynamic panel model and this implies that the choice of the best estimator to use is not a simple one since in this case even consistent estimators are characterized by small sample problems.[9] Consequently, Table 1 compares the results obtained by using four different fixed effects estimators applied to the whole period: Least Squares Dummy Variable (LSDV), LSDV with spatial correction suggested by Anselin et al. (2008) and with the correction advocated by Kiviet (1995) and the GMM suggested by Arellano and Bond (1991).[10]

---

[7] As is standard in this literature, $(g + \delta)$ is assumed equal to 0.05.

[8] Therefore, our sample includes the following years: 1985, 1988, 1991, 1994, 1997, 2000, 2003 and 2006.

[9] See Di Liberto et al. (2011) for a full discussion of all technical econometric issues.

[10] This estimation may be performed under very different assumptions about the endogeneity of the included regressors. In this study we adopt three different hypotheses (column V, VI and VII in table 1) on the additional regressors *x's*.

Table 1 Panel estimations

Sample: 199 regions EU, (1985–2006)

Dependent variable: per capita VA

Observations: 1393

	I OLS	II LSDV	III SEM*	IV KIVIET	V GMM-AB1	VI GMM-AB2	VII GMM-AB3
$\tilde{y}_{it-\tau}$							
	0.977***	0.599***	0.600***	0.742***	0.599***	0.492***	0.557***
	(0.004)	(0.021)	(0.000)	(0.026)	(0.055)	(0.050)	(0.047)
$\tilde{x}_{1,it-\tau}$	−0.005	0.041***	0.051***	0.038***	0.017	0.088***	0.062***
$\tilde{x}_{2,it-\tau}$	(0.005)	(0.010)	(0.000)	(0.011)	(0.013)	(0.023)	(0.017)
	0.011	0.027	−0.010	0.011	0.023	0.008	0.077***
	(0.013)	(0.017)	(0.516)	(0.018)	(0.019)	(0.030)	(0.019)
Spatial autocorr.			0.922***				
			(0.000)				
Sargan (p-value)					0.000	0.000	0.000
AB-2 test (p-value)					0.39	0.11	0.16

Notes: Standard errors in parentheses: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. Pooled model with spatial error autocorrelation and spatial fixed effects

First of all, results show that the Sargan test in each of the three models estimated by GMM implies that these specifications are not valid. As a consequence, in the remaining part of the paper we no longer use or report results based on this estimator. Further, when Eq. 1 is estimated with LSDV (column 2) we find that regional dummy coefficients $\hat{\mu}_i$, and thus TFP estimates, are almost invariably statistically significant. In particular, the F-test of the joint hypothesis that all the coefficients on our dummies are equal to zero is 23.15 (p-value = 0.00) and clearly rejects the hypothesis of no difference among regions. Moreover, we find an AR(1) coefficient of 0.60 and a corresponding high speed of convergence of 17 %. Among the regressors, both the coefficients on the lagged dependent variable and on population growth are significant and have the expected sign, while the coefficient on human capital is not significant.[11]

These results are confirmed when alternative estimators, the so called spatial error model (or SEM, see Anselin 1988)[12] and the KIVIET correction procedure

[11] The lack of empirical support for human capital in convergence regressions based on large international datasets is a well known problem. A number of possible explanations have been put forward. See Pritchett (1997), Temple (1999), and Krueger and Lindahl (2001).

[12] According to Rey and Montouri (1999) and López-Bazo et al. (2004) among others, the study of convergence across states and regions should take into account the possibility for spatial spill-overs across territorial units which may lead to spatial dependence. Such a possibility has been tested and the suggested model, the so called spatial error model (or SEM, see Anselin 1988), has been estimated by means of Maximum Likelihood and results are reported in column 3.

Table 2 Spearman's rank correlation coefficient

Estimators	LSDV	SEM	KIVIET	VAP85
LSDV	1.000	0.999	0.996	0.927
SEM		1.000	0.995	0.925
KIVIET			1.000	0.953
VAP85				1.000

(see Kiviet 1995), are used. Nevertheless, we prefer KIVIET estimates in light of Monte Carlo results which show that, for balanced panel and small (less or equal to ten) or moderate T (T = 30), this estimator has more attractive properties than the other two available options. [13]

With LSDV, SEM and KIVIET estimates in hand, we, therefore, compute three sets of TFP indexes respectively. Such sets are later standardised with the formula $T\tilde{F}P_i = \hat{A}(0)_i/\hat{A}(0)_{DK}$, with $\hat{A}(0)_{DK}$ being the estimated TFP value for Denmark, which is consistently the TFP leader. A closer inspection of our estimates reveals that best and worst performers are almost identical across the estimators. Moreover, Table 2 reports the Spearman rank order correlation which enables us to assess if the regional rankings in terms of TFP levels differ across the three procedures: LSDV, SEM and KIVIET. Figures show that regional rankings obtained with these three methods are remarkably similar, with the index always above 0.99. High values (from 0.92 to 0.95) are also obtained when we compare our estimates with the ranking of regional per capita value added in the initial year, 1985 (VAP85).

To sum up, the pattern and the magnitude of TFP heterogeneity, whatever the chosen estimator, confirm that cross-region TFP inequality is wide and that it is strongly associated with differences in per capita VA. In other words, controlling for possible convergence due to capital deepening processes, we still find that a potential for technological catch-up of lagging regions does exist. In the following, we implement the same methodology to compute TFP at two points in time in order to assess how much of the actual convergence can be explained by the occurrence of that potential. Spatial analysis is subsequently applied to investigate the geographical dimension of TFP dynamics.

3 Regional EU TFP Convergence: Preliminary Results

To detect how much TFP convergence is present in our sample, we estimate TFP using the same methodology described in the previous section over two sub-periods: 1985–1997 and 1997–2006. As before, we estimate Eq. 1 and save the two different series of $\hat{\mu}_i$ and then compute the two indices $T\tilde{F}P_{i,1} = \hat{A}_{i,1}/\hat{A}_{DK,1}$ (for the initial period, 1985–97) and $T\tilde{F}P_{i,2} = \hat{A}_{i,2}/\hat{A}_{DK,2}$ (for the subsequent period, 1997–2006). Estimation results are summarised in Table 3, where we focus on the

[13] See Kiviet (1995); Judson and Owen (1999); Everaert and Pozzi (2007).

Table 3 Panel estimations

Sample: 199 regions EU, 3 years time-span						
Dependent variable: per capita VA						
Observations: 796						
	1985–1997			1997–2006		
	OLS	LSDV	KIVIET	OLS	LSDV	KIVIET
$\tilde{y}_{it-\tau}$	0.973***	0.231***	0.378***	0.985***	0.649***	0.881***
	(0.006)	(0.032)	(0.042)	(0.004)	(0.029)	(0.038)
$\tilde{x}_{1,it-\tau}$	−0.013	0.018	0.020	0.001	0.035***	0.024***
	(0.007)	(0.019)	(0.019)	(0.005)	(0.011)	(0.011)
$\tilde{x}_{2,it-\tau}$	0.012	0.067***	0.072***	−0.032**	0.086***	−0.095***
	(0.019)	(0.023)	(0.024)	(0.013)	(0.020)	(0.022)

Notes: Standard errors in parentheses: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

KIVIET estimates even though the OLS and LSDV results are also shown for comparative reasons.

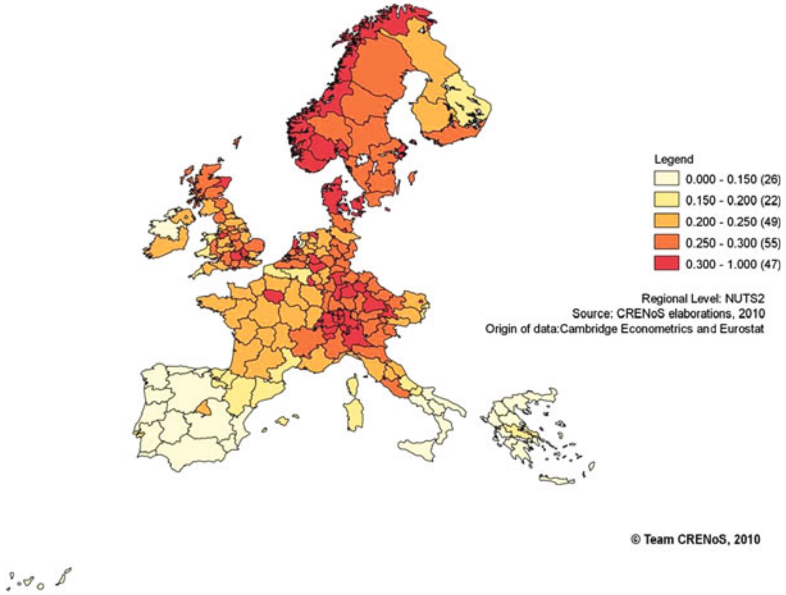
With these TFP estimates in hand,[14] we examine the main features of the two distributions for the two periods and focus on their geographical characteristics in both a static and dynamic perspective. We accept the possibility that TFP dynamics can be related to geographical localisation and to phenomena which are dependent on spatial features, such as distance among agents. Following Ertur and Koch (2006) and Rey et al. (2010), we investigate directly those features which may prove crucial in the catching up process and the diffusion of technology. In particular, we focus on geographical distance which can influence some channels of communications, such as trade, externalities and knowledge circulation, between the origin and the destination regions. Thus, we implement a spatial criteria to study the distribution of regional disparities in total factor productivity in order to see if the local environment of each region relative to its neighbours has a role in determining TFP distribution and its dynamics along time.

The analysis is mainly descriptive and based on global and local spatial measures of autocorrelation and on some visualisation techniques of the latest developments of ESDA.

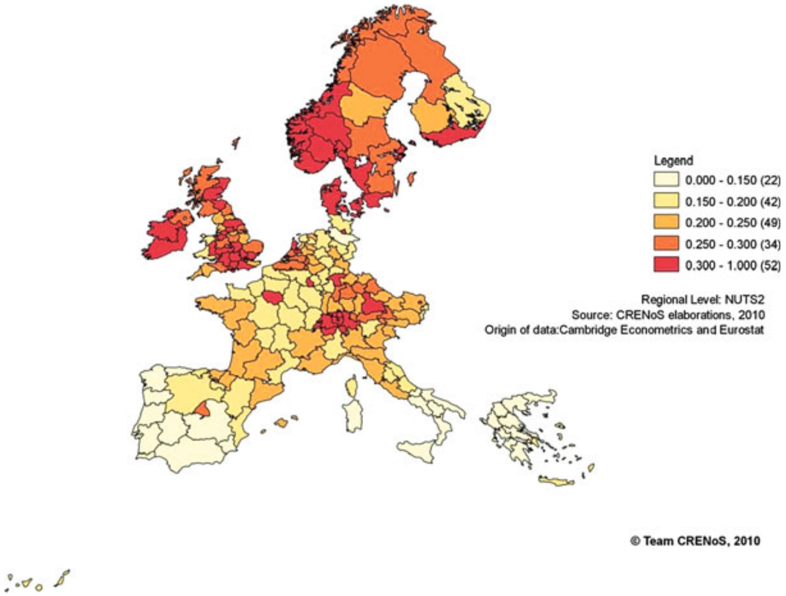
The starting point of the exploratory analysis is the inspection of the map of the phenomenon under examination. Map 1a, b show TFP levels in our two sub-periods of analysis: 1985–97 and 1997–2006. As expected, regional TFP levels in the first period are higher in the centre of Europe, United Kingdom and in some Northern Scandinavian countries (especially Norway and Sweden). Backward regions are concentrated in the South of Europe. This confirms a well known stylized fact, that the northern EU regions are at the top of the technology ladder and southern ones at

[14] The whole set of estimated regional TFP values and the variation of the rankings in the two sub-periods can be found in a previous version of this study. See Di Liberto and Usai (2010).

a 1985-1997



b 1997-2006



Map 1 TFP levels. (**a**) 1985–1997, (**b**) 1997–2006

the bottom. Among the former we find Denmark followed by some capital city areas, that is Inner London, Oslo and Brussels. At the bottom of the ranking the TFP laggard regions are all in southern Europe, they are in Italy, Spain, Portugal and

Table 4 CV and theil index statistics

Periods	Index CV	Theil
1985–1997	0.446	0.088
1997–2006	0.546	0.102

Greece. The second period shows significant differences in terms of the spatial distribution of regional TFPs: some progress is recorded in United Kingdom, Finland and Ireland. At the same time, several regions in central Europe (notably French and German regions together with Italian regions in the centre-north) lose ground. Some capital regions keep their position (Ile de France and Madrid), some others do not (Lazio). Therefore, comparing the two maps we find evidence of a possible presence of polarisation of high and low levels of TFP. This might be a result of the IT revolution, which has been put in action by Northern regions and neglected in Central-South of Europe.

To further investigate the dynamics of polarisation, we use two popular and intuitive measures of inequality in regional economics, that is, the Theil index and the coefficient of variation (CV, henceforth). They are therefore proposed in Table 4. In case of convergence we expect their values to decrease from the first to the second period of analysis.

Conversely, we find that the main characteristic of our TFP distributions is the absence of an overall process of TFP convergence. Actually, both the Theil index and the CV increase along time, implying a slight process of divergence in TFP levels across EU15 + 2 regions in the two decades starting from 1985. This is also suggested by Fig. 1 which illustrates the kernel distribution of TFP in the two periods. In this figure we can see the absence of significant changes in the distribution between the initial TFP levels (dashed line) and subsequent TFP levels (straight line). Secondly, it suggests that a club of highly productive regions (those ones on the extreme right of the distribution) is moving away from the mass of the other less efficient regions.

This result suggests that the absence of a clear cut overall dynamics may, in fact, hide complex and interesting intra-distribution patterns. In particular, when we focus on detailed regional data we find that the intra-distribution dynamics across EU regions has been remarkable. Data show[15] that EU regions have experienced significant changes of rank. Among the losers we mainly find German, Italian and Dutch regions: Rheinhessen-Pfalz and Trentino (−100 positions), Hannover (−98), Arnsberg (−93), Groningen (−63). Results are less region specific among the winners, with the exception of Ireland. The best performers are regions in the UK and Ireland: Border (160), Southern and Eastern (+140), Herefordshire (+89), Northern Ireland (+85). Finally, notice the remarkable association between the trends of TFP and VA per capita: regions which have significantly improved their TFP ranking are also the regions which have achieved high growth in VA per

[15] See also Di Liberto and Usai (2010) where for each region in Table A1 in the appendix we include both the first and second sub-period ranking position in terms of relative TFP levels and in the last two columns we include the change of rank in relative TFP levels and that observed in per capita VA levels between the initial and the final observation.

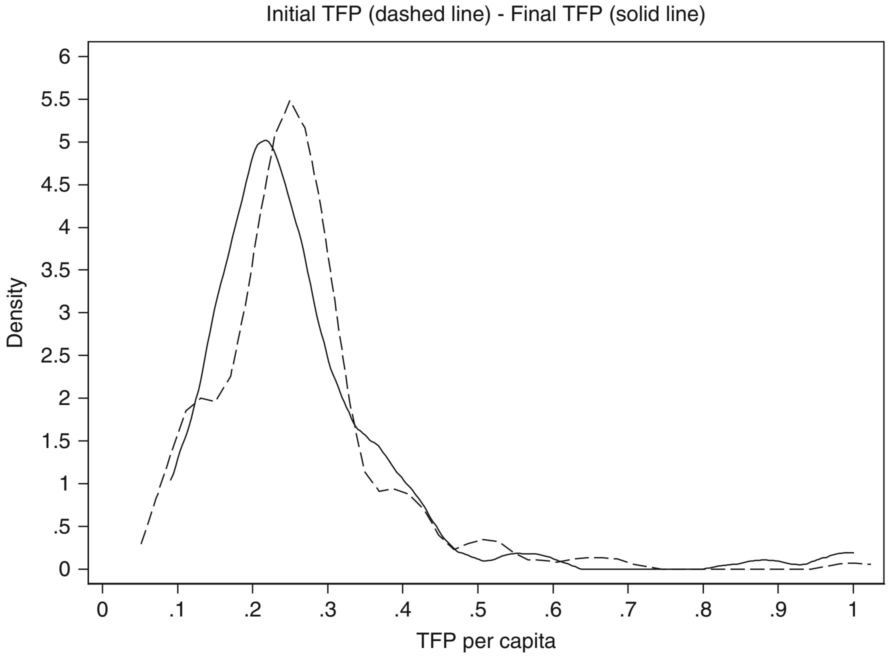


Fig. 1 Kernel TFP

capita. While obtaining fast growth in TFP is not simple, it appears to be a key factor in achieving fast VA per capita growth.

4 TFP, Geography and Spatial-Dynamic Analysis

The analysis above raises additional research questions about the nature of these intra-distribution dynamic patterns. In particular, we wonder where these mobile regions are located. Are they close to each other? Are neighbouring regions following the same trends while forming geographical clusters?

To answer these questions we carry out the application of ESDA with the study of spatial autocorrelation, a useful way to analyse territorial patterns in a certain period and along time. Local spatial movements can be traced by means of the Moran Scatterplot (Anselin 1996), which illustrates different types of spatial association (each corresponding to a quadrant) between a region and its neighbours.[16] The Moran scatterplot in the two periods, reported in Fig. 2a, b

[16] To model spatial dependence, a connectivity grid, that is a spatial weight matrix, has to be specified. A spatial weight matrix W specifies exogenously the connection among regions and it can refer to either contiguity or to distance. In this paper we refer to contiguity, which implies that the *wij* element of the W matrix is set to unity when regions are contiguous and zero otherwise.

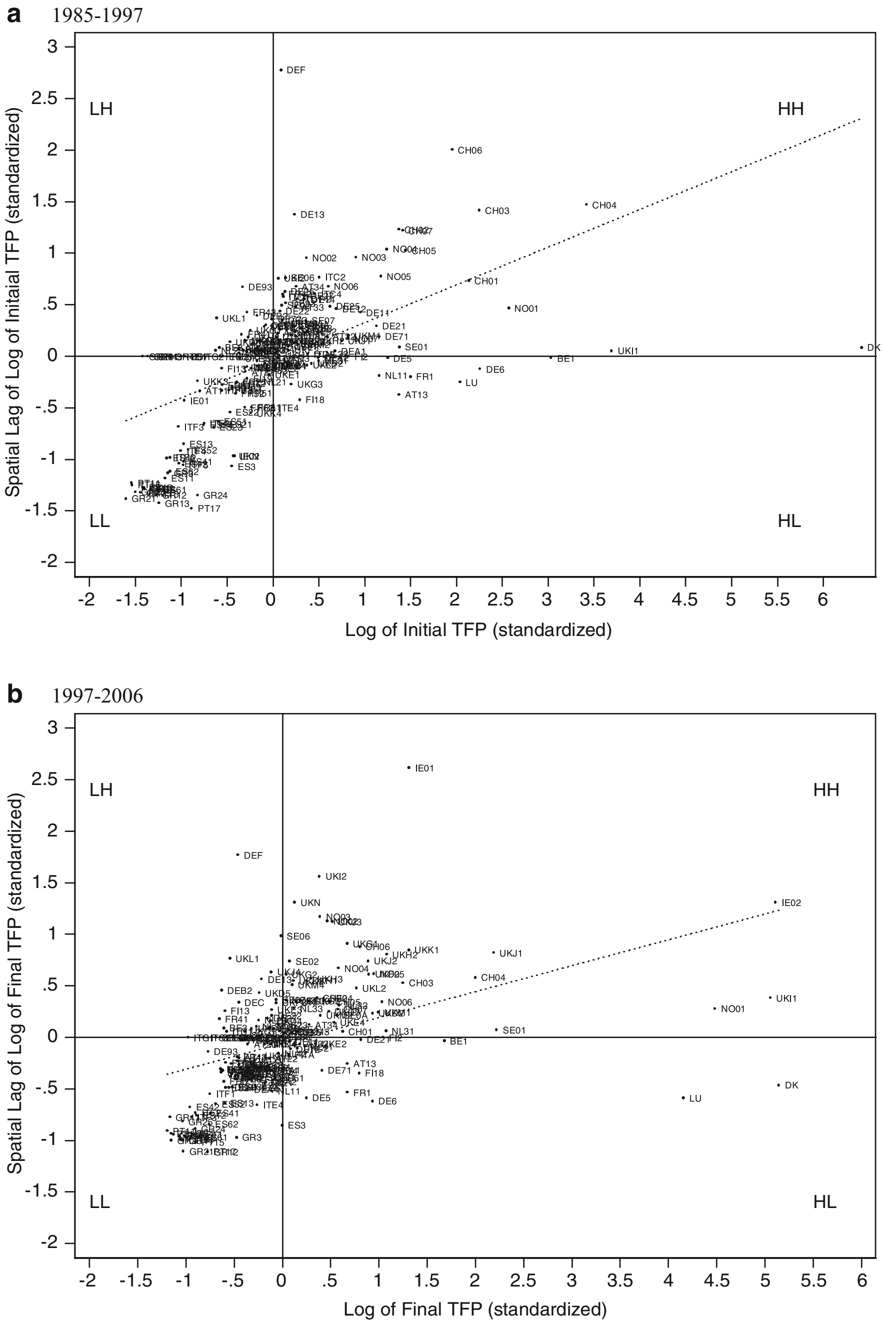


Fig. 2 Moran scatterplot for per capita TFP. (**a**) 1985–1997, (**b**) 1997–2006

respectively, is a useful tool since it immediately visualises spatial clustering of similar values (quadrants HH and LL) as much as cases of atypical values (quadrants LH and HL). More specifically:

- The North-Eastern quadrant of the plot contains regions which have above average TFP surrounded by regions which also have high TFP (HH, high cluster regions);
- In the North-western quadrant we find below average TFP regions whose neighbours are above the TFP average (LH, backward regions);
- The South-western quadrant consist of regions with low TFP surrounded by other low TFP regions (LL, low cluster regions);
- In the South-eastern quadrant one finds regions with high levels of TFP but with low TFP regions as their neighbours (HL, leader regions).

  The Moran scatterplot is also useful since it allow to assess the global spatial dependence (by means of the Moran I), which is represented by the slope of the linear regression (the dashed line) of the spatially lagged TFP on the original TFP.

Figure 2a, b provide interesting information on both cross-sectional and time varying patterns of regional TFP clustering processes. Firstly, they show that in both periods spatial autocorrelation is present and significant. Nonetheless, as far as the global spatial autocorrelation is concerned, the Moran I decreases from 0.366 to 0.251, which implies that clusters of similar regions (either with high or with low values) are becoming weaker in time. The quota of regions which are characterised by the presence of positive spatial association is however constant along time, at around 70 % in both periods.[17] We also identify a few outliers, that is, regions which appear far away from the majority of regions located either around the origin or along the dashed line in the graph. Outliers with respect to the x axis (regions which are relatively poor but in rich surroundings) are Schleswig-Holsen in Germany (DEF) in the first period and Border in Ireland (IE01) in the second period; with respect to the y axis we find one region, that is Denmark (DK) in the first period and five regions in the second period, that is Southern and Eastern (IE02) in Ireland, Oslo (NO01), Inner London (UKI1) and Luxembourg (LU). These subsets of regions, which are relatively richer than neighbouring regions, are often capital regions and/or highly urbanised areas.

Another interesting aspect which can be extracted from the comparison of the two figures is the presence of an upward movement of the low-cluster regions (in the LL quadrant) that seem to move closer to the origin in the second period. This weak sign of TFP convergence for the backward regions is associated with the simultaneous enlargement of the two clusters in the LL quadrant. As a matter of fact this alarming evidence is confirmed in Table 5, that includes the share of regions in

[17] It is interesting to note that quite a similar result is obtained by Ertur and Koch (2006) in their analysis of income per capita across EU15 and EU27 regions: in 2000 they find 75 % of positive spatial dependence.

Table 5 Spatial association in the moran scatterplot

	Quadrant HH (%)	Quadrant LL (%)	Quadrant HL (%)	Quadrant LH (%)	No quadrant (%)
1985–1997	0.37	0.33	0.08	0.17	0.05
1997–2006	0.28	0.43	0.07	0.18	0.05

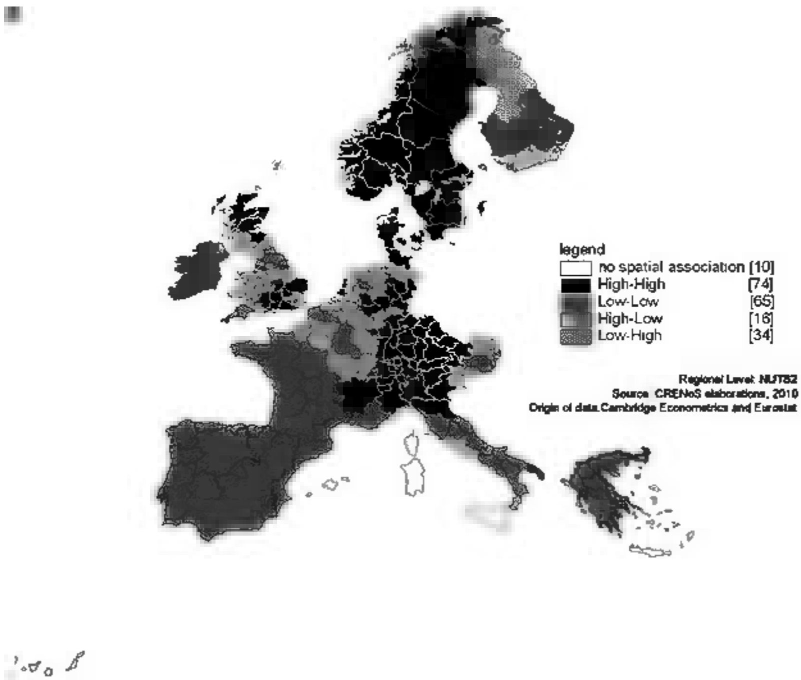
each quadrant in both periods of the analysis. We find that, the strength of the positive association for HH regions fades along time, since the quota of such regions goes from 37 % to 26 %, conversely, the fraction of the low-cluster regions (LL regions) increases significantly, from 33 % to 45 %. This implies that a non virtuous geographical clustering of TFP levels has appeared in the EU.[18] Regional shares of dissimilar regions (LH and HL), that is, regions with TFP levels higher or lower than their neighbours, are much more stable: LH regions stay, around 20 % and HL around 9 %.

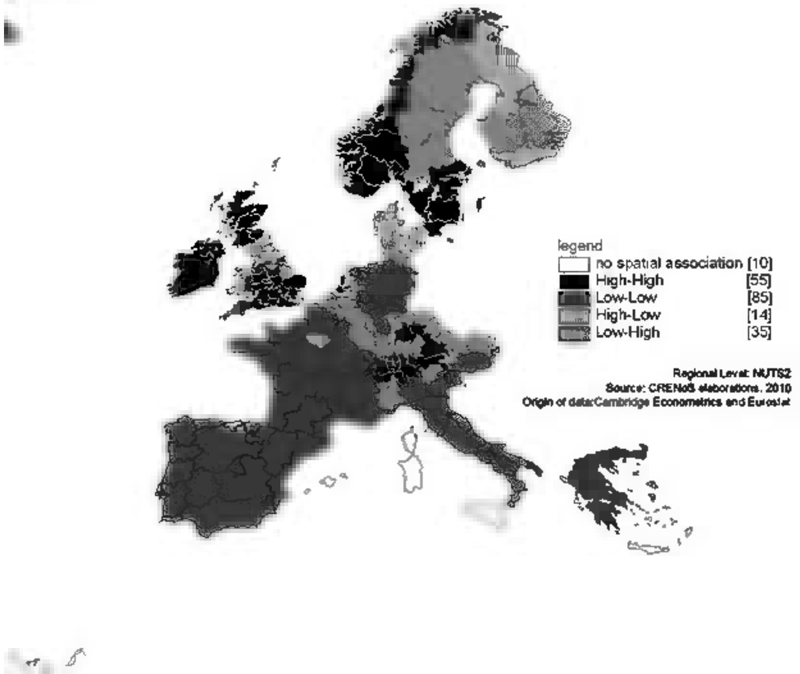
Another helpful statistical tool is the so called Moran map, where regions are positioned in a map[19] according to the type of association, that is LL, HH, LH or HL. Map 2a, b distinguish regions corresponding to the four quadrants of the scatterplot above with different colours, in order to identify clusters of similar regions or peculiar isolated cases. Map 2a shows two big red clusters of highly productive regions, which are close to each other, arising in the first sub-period. One is located in Scandinavia, more precisely in Norway and Sweden; another one is in Central Europe and goes across several countries from Italy to Denmark. There are also two small clusters in the South of the United Kingdom around London, and another one among the Scottish regions in the North. Conversely, aggregations of poorly productive regions are typically located in the South of Europe (the Mezzogiorno of Italy, the Iberian peninsula and Greece) but also in some Northern regions: Ireland and Finland in particular.

Map 2b describes a dramatically different scene 10 years later. First of all, the central European cluster reduces to just some regions in Switzerland, Austria and Germany. There is also a small cluster between Belgium and the Netherlands, but all the regions in between are now characterised by the presence of territories with low productivity. In the North of Europe, the Scandinavian cluster loses some Swedish regions while gaining some high productivity regions in Finland, even though these are still contiguous with low TFP regions. Only the cluster of highly productive regions in the southern UK widens while almost joining with the one in the North and, most importantly, the one in Ireland. At the same time, a major change is observed among the low-TFP cluster regions: the mass of low productivity regions now stretches along the whole of the South of Europe and includes France and many German regions.

[18] Similar result have been found in Di Liberto et al. (2011) at the country level.

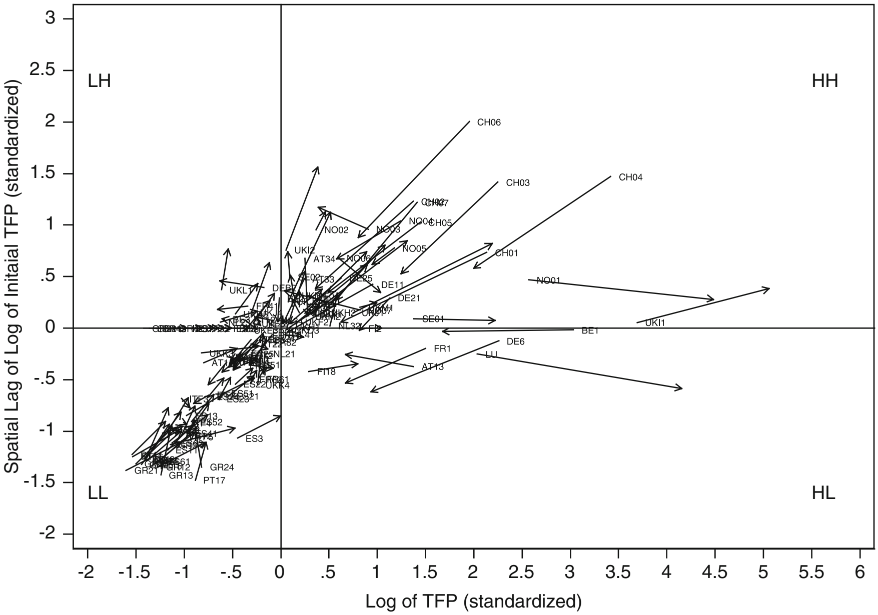
[19] In Table 2A in Di Liberto and Usai (2010) more detailed information is given about the Local Indicator of Spatial Association which provides the significance of the relationship for each region and its neighbours (Anselin 1995).





Map 2 Moran maps. (**a**) 1985–1997, (**b**) 1997–2006

• **a** Regions which do not change quadrant



• **b** Regions which change quadrant

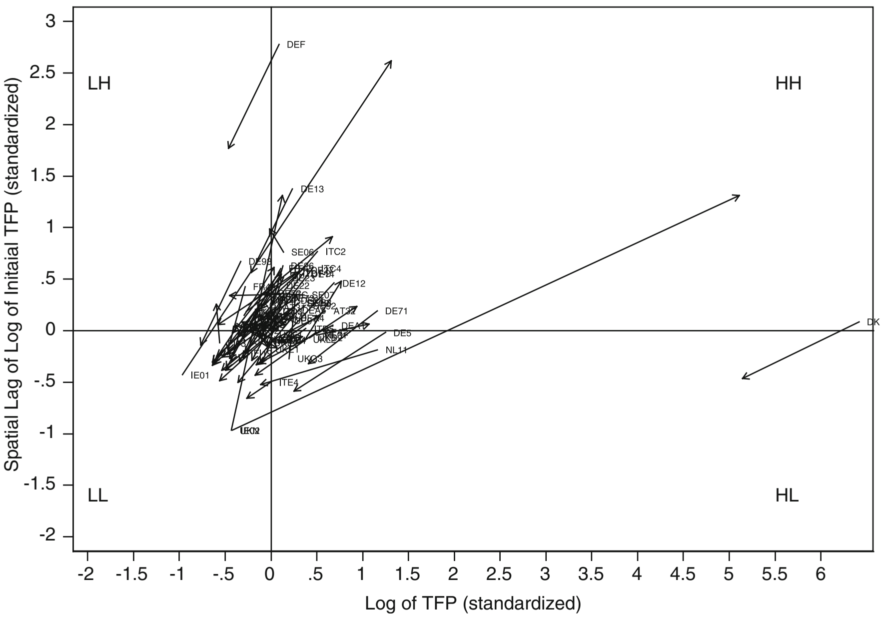


Fig. 3 Directional moran scatterplot. (**a**) Regions which do not change quadrant, (**b**)Regions which change quadrant

Table 6 Transition matrix in the moran scatterplot

	(1997–2006)			
(1985–1997)	Quadrant HH (%)	Quadrant LL (%)	Quadrant HL (%)	Quadrant LH (%)
Quadrant HH	54.05	16.22	6.76	22.97
Quadrant LL	6.15	86.15	1.54	6.15
Quadrant HL	37.50	18.75	37.50	6.25
Quadrant LH	14.71	41.18	5.88	38.24

The image provided by the comparison of the two maps above is thus worrisome: a dual Europe is taking shape without the extensive fringe zone which was guaranteed in the past by central European regions. The risk is that the inertia of spatial association will keep on working in a positive way in the North and mainly negatively in the South.

To further interpret the evolution captured by the analysis of the Moran scatterplot and the Moran maps Rey et al. (2010) suggest using an innovative graphic tool called Directional Moran Scatterplot, reported in Fig. 3 above. This tool is meant to unmask and identify regional individual movements.

Figure 3 displays the Directional Moran Scatterplot applied to our data. This allows us to pinpoint each region's transitions along time as a vector, where the arrowhead marks the movement towards the location in the final period. Since a clear visualization of 199 regional transitions is almost impossible, we distinguish between (1) movements within the same quadrant (reported in Fig. 3a), and (2) movements across quadrants (in Fig. 3b).

From the first figure we can spot several upward movements in the LL quadrant from Greek and Portuguese regions: a sign of feable convergence of the least productive regions. In the HH quadrant the most interesting moves are those of Swiss regions which reposition closer to the origin. From the second figure we can easily discern the dramatic change of the two Irish regions moving from the LL quadrant to the HH one, as well as the opposite path shown by German regions which move from the HH to the LL quadrant.

More information about the regional TFP dynamics can be found in Table 6 that introduces a standard transition matrix for regional TFP levels as usually done in income distribution studies. Unlike income distribution transition matrix, this table depicts the transitions across the scatterplot quadrants over time (see Rey and Ye 2010). Therefore, the main diagonal includes the quota of regions that do not change their "state" across HH, LL, HL or LH between the first and the second sub-periods. We notice that the most stable quadrant is the LL (85 % regions keep their position), the least stable is the HL. It is interesting to note that most regions go from HL to HH, so there is a cluster effect referring to HH regions. However, this result is in contrast with the fact that almost half the regions which were in the HH quadrant in the first period have moved, going either to HL (19 %) or to LL (22 %). This implies that some positive spatial dependence working among rich regions is now still positive but working among poor regions. The strength of this cluster is therefore quite low. Another interesting aspect illustrated in Table 6 is the fact that 35 % of those regions which were in the LH quadrant are now in the LL, as if the spatial dependence among poor regions was getting stronger along time.

5 Conclusions

This paper explores a potential process of technology convergence across a sample of 199 European regions between 1985 and 2006. To this aim, we first suggest the use of a panel approach to estimate TFP levels in different periods of time and then apply a recently developed spatial analysis technique to thoroughly investigate the geographical innovation patterns across the EU.

The main advantage of using the fixed-effect panel methodology to estimate TFP levels is that it enables us to distinguish between two possible components of convergence: one due to capital deepening processes and a distinct one due to technology transfers. In fact, despite the large array of methodologies proposed to measure TFP heterogeneity across regions, only a few of these try to capture the presence of technology convergence as a separate component from the standard (capital-deepening) source of convergence. Robustness of results is assessed using different estimation procedures, such as simple LSDV, spatially-corrected LSDV, Kiviet-corrected LSDV, and GMM à la Arellano and Bond (1991) but, given the characteristics of our sample, we use Kiviet-corrected estimates to calculate our final TFP levels.

The analysis on the geographical patterns of TFP reveals the presence of a high and persistent level of TFP heterogeneity across EU regions and the absence of a global process of TFP convergence. Within this aggregate persistence, important changes are nevertheless detected. Such changes show a relevant geographical component since spatial dependence has proved to be a constant feature of TFP distribution along time. In particular, we observe that there is a polarisation of richer regions in the North of Europe while southern regions keep losing ground. In other words, the cluster of successful regions has become smaller and is turning more distant in economic and in geographical terms from the cluster of less productive regions. That is, this last cluster seems to have grown in numbers and in territorial extension.

In sum, our ample descriptive analysis offers a broad picture, though not always reassuring, about TFP dynamics and technology diffusion processes across EU regions. First of all, we find that while the global distribution of regional TFP levels seems to remain quite stable over time, the intra-distribution (or single region movements) dynamics shows significant changes in ranking across regions. Additional inspection suggests the presence of regional productivity polarisation between high and low TFP levels which can be attributed to the recent asymmetric shock due to the IT revolution.

Overall, few new regions have recently joined the TFP leaders, while results suggests that the leading club has expelled several previous members located mainly in the centre of Europe. Such regions are now members of the low TFP club which is enlarging its reach far beyond the initial group from Southern Europe. This calls for new analysis, not only to investigate the reasons of the apparent absence of technology diffusion processes but also the cause of the prevalence of the spatial dependence among low TFP regions rather than among high TFP regions. If such a prevalence is to settle, the future ahead is one of divergence rather than convergence.

Acknowledgments The research leading to these results has received funding from the ESPON project KIT, Knowledge, Innovation, Territory. We thank Claudio Deiana for his excellent research assistance.

References

Acemoglu D, Johnson S, Robinson J (2001) The colonial origins of comparative development: an empirical investigation. Am Econ Rev 91:1369–1401
Acemoglu D, Aghion P, Zilibotti F (2006) Distance to frontier, selection and economic growth. J European Econ Assoc 4(1):37–74
Alesina A, La Ferrara E (2005) Ethnic diversity and economic performance. J Econ Lit, Am Econ Assoc 43(3):762–800
Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dordrecht
Anselin L (1995) Local indicators of spatial association-LISA. Geogr Anal 27:93–115
Anselin L (1996) The moran scatterplot as an ESDA tool to assess local instability in spatial association. In: Fischer M, Scholten H, Unwin D (eds) Spatial analytical perspectives on GIS. Taylor and Francis, London, pp 111–125
Anselin L, Le Gallo J, Jayet H (2008) The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice (3rd edn). In: Matyas L L, Sevestre P (eds), Springer-Verlag, Berlin Heidelberg, pp 627–662
Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 58:277–297
Bellini E, Ottaviano G, Pinelli D (2012) Cultural diversity and economics performance: evidence from European regions, in this volume
Bernard A, Jones C (1996) Technology and convergence. Econ J 106:1037–1044
Capello R, Fratesi U (2012) Globalization and endogenous regional growth, in this volume
Caselli F, Esquivel G, Lefort F (1996) Reopening the convergence debate: a new look at cross country growth empirics. J Econ Growth 1:363–389
Comin D, Hobijn B (2004) Cross-country technology adoption: making the theories face the facts. J Monetary Econ 51:39–83
Comin D, Hobijn B (2009) Lobbies and technology diffusion. Rev Econ Stat 91(2):229–244
Comin D, Hobijn B, Rovito E (2008) Technology usage lags. J Econ Growth 13(4):237–256
Crescenzi R, Gagliardi L Percoco M (2012) Social capital and innovation in Italy, in this volume
Del Gatto M, Di Liberto A, Petraglia C (2011) Measuring productivity. J Econ Surv 25(5): 952–1008
Dettori B, Marrocu E, Paci R (2011) Total factor productivity, intangible assets and spatial dependence in the European regions. Regional Studies (in press)
Di Liberto A, Usai S (2010) TFP convergence across European regions: a comparative spatial dynamics analysis. Crenos working papers 2010/30
Di Liberto A, Pigliaru F, Mura R (2008) How to measure the unobservable: a panel technique for the analysis of TFP convergence. Oxford Econ Pap 60(2):343–368
Di Liberto A, Pigliaru F, Chelucci P (2011) International TFP dynamics and human capital stocks: a panel data analysis, 1960–2003. Rev Income Wealth 57(1):156–182
Durlauf S, Johnson P, Temple J (2005) Growth econometrics. In: Aghion P, Durlauf S (ed) Handbook of economic growth, vol 1, North-Holland, Amsterdam. pp 555–677
Easterly W, Levine R (1997) Africa's growth tragedy: policies and ethnic divisions. Q J Econ 112 (4):1203–1250

Easterly W, Levine R (2001) It's not factor accumulation: stylized facts and growth models. World Bank Econ Rev 15:177–219
Ertur C, Koch W (2006) Regional disparities in the European Union and the enlargement process: an exploratory spatial data analysis, 1995–2000. Annal Reg Sci 40:723–765
Everaert G, Pozzi L (2007) Bootstrap-based bias correction for dynamic panels. J Econ Dyn Control 31(4):1160–1184
Ezcurra R, Iraizoz B, Pascual P (2009) Total factor productivity, efficiency, and technological change in the European regions: a nonparametric approach. Environ Plann A 41(5):1152–1170
Grier K, Grier R (2007) Only income diverges: a neoclassical anomaly. J Dev Econ 84:25–45
Hall R, Jones C (1999) Why do some countries produce so much more output per worker than others? Q J Econ 114:83–116
Islam N (1995) Growth empirics: a panel data approach. Q J Econ 110:1127–1170
Islam N (2003) Productivity dynamics in a large sample of countries: a panel study. Rev Income and Wealth 49:247–272
Jorgenson D (2005) Accounting for growth in the information age. In: Aghion P, Durlauf S (ed) Handbook of economic growth, vol 1, North-Holland, Amsterdam. pp 743–815
Judson R, Owen A (1999) Estimating dynamic panel data models: a guide for macroeconomists. Econ Lett 65:9–15
Kiviet J (1995) On bias inconsistency and efficiency of various estimators in dynamic panel data models. J Econ 68:53–78
Knack S, Keefer P (1997) Does social capital have an economic payoff? A cross-country investigation. Q J Econ 112(4):1251–1288
Krueger A, Lindahl M (2001) Education for growth: why and for whom? J Econ Lit 39:1101–1136
Le Sage JP, Fischer M (2012) Estimates of the impact of static and dynamic knowledge spillovers on regional factor productivity. Int Reg Sci Review 35(1):103–127
López-Bazo E, Vayá E, Artís M (2004) Regional externalities and growth: evidence from European regions. J Reg Sci 44(1):43–73
Magrini S (2004) Regional (di)convergence. In: Henderson V, Thisse JF (eds) Handbook of regional and urban economics: cities and geography. Elsevier, the Netherlands
Mankiw N, Romer D, Weil D (1992) A contribution to the empirics of economic growth. Q J Econ 107:407–437
Marrocu E, Paci R, Usai S (2012) Productivity growth in the old and new Europe: the role of agglomeration externalities. Int J Reg Studies (in press)
Monastiriotis V (2012) Regional growth in central and eastern Europe: convergence, divergence and non-linear dynamics, in this volume
Nelson R, Phelps E (1966) Investments in humans technological diffusion and economic growth. Am Econ Rev 56:69–75
Parente S, Prescott E (1999) Monopoly rights: a barrier to riches. Am Econ Rev 89:1216–1233
Pritchett L (1997) Divergence: big time. J Econ Perspect 11:3–18
Rey SJ, Montouri BD (1999) US regional income convergence: a spatial econometric perspective. Reg Stud Taylor Francis J 33(2):143–156
Rey SJ, Ye X (2010) Comparative spatial dynamics of regional systems. In progress in spatial analysis: theory and computation, and thematic applications, pp 441–464
Rey SJ, Murray AT, Anselin L (2010) Visualising regional income distribution dynamics. GeoDa spatial center working papers, p 14
Rodriguez-Pose A, Comptour F (2012) Do clusters generate greater innovation and growth? An analysis of European regions, in this volume
Scherngell T, Fischer MM, Reismann M (2007) Total factor productivity effects of interregional knowledge spillovers in manufacturing industries across Europe. Romanian J Reg Sci 1 (1):1–16
Tabellini G (2010) Culture and institutions: economics development in the regions of Europe. J European Econ Assoc 8:677–716
Temple J (1999) A positive effect of human capital on growth. Econo Lett Elsevier 65(1):131–134

Productivity and Local Workforce Composition

David Christopher Maré and Richard Fabling

Abstract This chapter examines the link between firm productivity and the population composition of the areas in which firms operate. We combine annual firm-level microdata with area-level workforce characteristics obtained from population censuses. Overall, the results confirm the existence of agglomeration effects that operate through local labour markets. We find evidence of productive spillovers from operating in areas with high-skilled workers, and with high population density. The strength and nature of spillovers varies across different types of firms. Our findings demonstrate the importance of controlling for multiple dimensions of local workforce composition, and of analysing effects for subpopulations of firms.

Keywords Productivity • Agglomeration • Workforce composition

1 Introduction

This chapter examines the links between economic performance and the spatial context in which firms operate. In particular, it focuses on the impact on productivity of local workforce density, composition and diversity. A dependence of firm productivity on local workforce characteristics is predicted by a range of economic theories that emphasise productive spillovers, particularly in dense urban areas.

JEL codes: R1, R3, D24

D.C. Maré
Motu Economic and Public Policy Research and University of Waikato, Cuba Street, Wellington 6142, New Zealand
e-mail: dave.mare@motu.org.nz

R. Fabling
Motu Economic and Public Policy Research, Cuba Street, Wellington 6142, New Zealand
e-mail: richard.fabling@motu.org.nz

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_4,

The geography of factor inputs has long been identified as a key source of urban agglomeration economies. Smith (1904, I.3.2) highlights the gains from the greater labour specialisation that is made possible in “great towns”. Marshall (1920, Bk IV. X) famously emphasised the operation of skill accumulation and innovation in urban labour markets, and the improved access to specialised skills in thick labour markets. More recent analyses of the microfoundations of agglomeration continue to place a strong emphasis on urban labour market mechanisms. These include sharing the gains from specialisation and pooling labour market risks, improving the quality of labour market matching, and supporting the generation, diffusion and accumulation of information and knowledge (Duranton and Puga 2004).

Other chapters in this book focus on productive advantages that may arise in regions with culturally diverse populations. Bakens and Nijkamp (chapter “Migrant Heterogeneity and Urban Development: A Conceptual Analysis”) outline the impacts of migrant heterogeneity on urban productivity and amenities, discussing strategies for identifying separately the two effects. Bellini and Pinelli (chapter “Cultural Diversity and Economic Performance: Evidence from European Regions”) examine the relationship between regional productivity and the presence and diversity of foreigners, concluding that cultural diversity raises economic performance, and possibly also consumption amenities.

The current chapter presents estimates of the productivity effects of local workforce characteristics, including the presence of migrants, obtained from direct estimation of firm-level production functions. We combine annual firm-level microdata on production, covering a large proportion of the New Zealand economy, with area-level workforce characteristics obtained from population censuses. The use of firm microdata allows us to examine the heterogeneity of productivity impacts across different industries, providing an indication of which firms benefit most from different dimensions of local workforce composition.

The following section outlines our analytical approach and estimation strategy. This is followed by a description of the data that we use before we present our main results. The chapter concludes with a discussion of our main findings.

2 Analytical Approach

Firm performance is characterised in terms of a standard production function, which is potentially augmented by the productivity impacts of the spatial context in which the firm operates. It is common to characterise the influence of spatial characteristics as a Hicks-neutral factor, as in the following production function:

$$Y_i = A_j f_i(K_i, L_i, M_i) \tag{1}$$

where gross output Y_i produced by firm i depends not only on its inputs of capital (K_i), labour (L_i) and materials (M_i), but also on characteristics (A_j) of the local area j. We focus on four features of the local area population – the density of the local population, the proportion that is highly qualified, the proportion that is newly arrived in the area, and the proportion that is foreign-born. These features capture key dimensions of the local workforce that have been identified as significant in previous studies.

Those studies have found a clear positive relationship between the productivity of firms and the density of economic activity in the locations where they operate (Ciccone and Hall 1996). Density is a rather coarse proxy for a broad range of potential advantages associated with agglomeration. Identifying and disentangling the different potential causes of these productivity advantages remains a challenge (Rosenthal and Strange 2004). There is a well-established body of literature that documents the important role played by labour market interactions and knowledge spillovers that are associated with variation in local workforce characteristics.

Moretti (2004a) reviews empirical approaches to estimating local human capital spillovers, distinguishing studies that identify spillovers through their simultaneous impacts on wages and rents (such as Bellini and Pinelli's chapter in this volume), and those that rely on the estimation of firm productivity. The current paper takes the latter approach. Moretti's own empirical study (Moretti 2004b) is a leading example of the approach of estimating firm production functions. He finds positive evidence of human capital spillovers between local industries. Moreover, he finds that spillovers are stronger between industries that are close in terms of input–output linkages, technological similarity, and patent citation links, providing support for knowledge transfer explanations.

Moretti captures the local skill level as the proportion of the local population that has a college degree – analogous to our measure of the percent with a degree qualification. Moretti's findings echo those of Rauch's (1993) influential study, which found that workers in areas with a more highly qualified workforce earn higher wages, controlling for their own human capital, arguably as a result of knowledge spillovers. Rodríguez-Pose and Comptour (chapter "Evaluating the Role of Clusters for Innovation and Growth in Europe" of this volume) include average local educational attainment and educational participation in their 'social filter' measures that captures spatial influences on regions' innovativeness.

Economic performance can be affected not only by the average quality of the local workforce, but also by the local diversity of skills and cultures. Interactions between people with different knowledge and perspectives are more likely to generate new innovative and productive ideas, through the sort of externalities that were highlighted by Jacobs (1969). To capture the diversity within the local workforce, we include two additional measures in A_j. First, we include the proportion of residents who are new to the area, who may have brought new ideas and perspectives. Second, we include the proportion of the local workforce that is foreign-born.

More broadly, the composition and density of the local workforce can improve a firm's productivity performance through any of the three mechanisms identified by Duranton and Puga (2004) – sharing, matching and learning. Recent studies have found support for each of these mechanisms. Overman and Puga (2010) show the advantages associated with sharing of labour market risks in dense, skilled urban labour markets. Amiti and Pissarides (2005) show the potential agglomeration gains from better matching of heterogeneous workers. Studies of the localisation of patent citations (Jaffe et al. 1993) and the links between patenting and the presence of migrants locally (Hunt and Gauthier-Loiselle 2010) add further weight to explanations involving knowledge flows. More direct evidence of local knowledge interactions comes from Zucker and Darby's (2009) study of the location patterns of 'star scientists'.

We base our estimations on an empirical version of Eq. 1. Specifically, we estimate the relationship between productivity and local workforce characteristics using a gross output Cobb-Douglas production function augmented with area-level workforce composition measures,

$$GO_{it} = \phi_{it}^{A} + \beta_j^K K_{it} + \beta_j^L L_{it} + \beta_j^M M_{it} + \left(\lambda_i + \alpha_{jt} + \varepsilon_{it}\right), \tag{2}$$

where i denotes a firm, t refers to time period and j indicates parameters that vary by industry. Output (GO_{it}), capital services (K_{it}), labour input (L_{it}), and intermediate consumption (M_{it}) are all measured in logarithms. The error term potentially has components corresponding to firms, industries, and time periods. ϕ_{it}^{A} is the Hicks-neutral contribution to productivity in period t of characteristics of the area (A_i) in which firm i operates. This contribution is entered as a linear combination of local workforce measures,

$$\begin{aligned}\phi_{it}^{A} =& \gamma^{Dens}[\%\text{Population density}]_{it}^{A} + \gamma^{HS}[\%\text{Degree qualified}]_{it}^{A} \\ &+ \gamma^{New}[\%\text{New to area}]_{it}^{A} + \gamma^{Mig}[\%\text{Foreign-born}]_{it}^{A} + e_{it}\end{aligned} \tag{3}$$

We use annual production data, combined with area information that is available only every 5 years. Consequently, we estimate Eq. 2 in two stages. In the first stage, we estimate productivity using an annual firm-level panel, but omitting area characteristics. We estimate a separate regression for each industry, allowing for clustered errors at the firm level.

In the second stage, we regress the residuals from the first-stage regression (multifactor productivity) on the right-hand-side terms of Eq. 3. The second stage regression is estimated using five-yearly firm-level data, with separate intercepts for industry and for year. We allow for area-clustered errors, since the area-level characteristics are common to all firms with the same geographic distribution (Moulton 1990).[1]

Workforce composition is potentially endogenous, as entrants and high-skilled workers may be attracted to areas with high-productivity firms (analogously to the sorting mechanisms discussed by D'Costa et al. in chapter "Agglomeration and Labour Markets: The Impact of Transport Investments on Labour Market Outcomes"). We use an instrumental variables approach to adjust for this endogeneity. Specifically, we use 5-year lags of the composition variables as instruments in the second stage regression.

---

[1] In practice, we observe firms operating in more than one location and measure geographic variables as the firm's average (employment-weighted) exposure to area characteristics. Clustering of errors is corrected for based on clusters identified from common combinations of area characteristics. Our standard errors do not allow for the variability associated with the use of generated regressors obtained from the first stage, and will therefore be somewhat understated. We generated one-step estimates for our main specifications and found that coefficients and standard errors were very similar to those obtained using our two-step procedure. On this basis, we judge that our results would be largely unchanged if we were to use one-step estimation or generate bootstrap standard errors for our two-stage estimates.

We also control for selected firm-level workforce characteristics that may be correlated with the area-level composition measures. Firms in areas where there is a high proportion of the workforce with a degree qualification will themselves employ more highly qualified personnel. Productivity in Eq. 2 is estimated using a headcount measure of labour input, which is likely to understate the effective labour input used by firms in high-skilled areas. Similarly, a high proportion of people new to an area may be reflected in higher worker turnover rates for local firms, which may have an independent influence on productivity. Consequently, we augment Eq. 3 by adding firm-specific labour quality and turnover measures.

3 Data

We combine firm-level microdata on production with area-level workforce characteristics. The workforce characteristics are drawn from the Census of Population and Dwellings, summarised at Area Unit level (roughly equivalent to a city suburb). Productivity is estimated using rich firm microdata contained in Statistics New Zealand's prototype Longitudinal Business Database (LBD).[2]

3.1 Production Data

The LBD dataset is based around the Longitudinal Business Frame (LBF), which provides longitudinal information on all businesses in the Statistics New Zealand Business Frame since 1999, combined with information from the tax administration system. The LBF population includes all employing businesses. We make use of the permanent enterprise identifiers developed by Fabling (2011), which uses plant transfers to improve the tracking of firms over time.

The primary unit of observation in the LBD is an enterprise (firm) year. We make use of business demographic information from the LBF, linked with financial performance measures for the 1999/2000 to 2007/08 years. Plant location and employment information from the Linked Employer-Employee Dataset (LEED) is used to link to local area information from the Population Census.

To calculate multifactor productivity (MFP), we follow Fabling and Maré (2011). Gross output is measured as the value of sales of goods and services, less the value of purchases of goods for resale, with an adjustment for changes in the value of stocks of finished goods and goods for resale. Gross output and factor inputs are measured in current prices.[3] Capital services has four components: depreciation; rental and leasing

[2] See Fabling (2009) for further information on the LBD.

[3] Changes over time in current price inputs and outputs will reflect both quantity and price changes. We double deflate to isolate quantity adjustment over time at the (one- or two-digit) industry level using Statistics New Zealand's PPI input and output indices. Measures of productivity premia for firms within the same industry will reflect both quantity and relative price differences. Spatial price indices are not available.

costs; rates; and the user cost of capital. The inclusion of rental and leasing costs and rates ensures consistent treatment of owned and rented or leased capital. The user cost of capital is calculated as the value of total assets, multiplied by an interest rate equal to the average 90-day bill rate plus a constant risk-adjustment factor of 4 percentage points. Intermediate consumption is measured as the value of other inputs used in the production process, with an adjustment for changes in stocks of raw materials.

The primary source used to obtain gross output, intermediate consumption, and capital services is the Annual Enterprise Survey (AES). This information is available for around 10 % of enterprises, which are disproportionately larger firms, accounting for around half of total employment in New Zealand. Where AES information is not available, we derive comparable measures from annual tax returns (IR10s). Enterprise total employment comes from LEED and comprises the count of employees in all of the enterprise's plants, annualised from employee counts as at the 15th of each month, plus working proprietor input, as reported in tax returns.

3.2 Local Workforce Composition

Information on local workforce composition is obtained from the 2001 and 2006 New Zealand Censuses of Population and Dwellings. Within urban areas, we use information for individual Area Units. Outside urban areas, population composition is measured as the average for non-urban Area Units in each territorial authority. This averaging is necessary to ensure that populations are large enough to support the required disaggregation.[4]

From the census data, we classify each member of the population aged 18–65 according to qualification, nativity, and recency of arrival. The workforce is classified into two qualification levels (tertiary qualified and other), two nativity groups (born in New Zealand, born elsewhere), and recency of arrival in the current Area Unit (within previous 5 years, or earlier).[5] For each qualification group, we have six sub-groups: two groups of people who were in the same location 5 years earlier (New Zealand-born and earlier migrants), two of people who were elsewhere in New Zealand 5

[4] On average Area Units contain around 2,000 people. Area Units with populations of less than 100 are dropped from our analysis. There is a small number of Area Units for which disaggregated population information could not be separately released within the protections of the Statistics New Zealand confidentiality policy. Population composition for these areas was measured as the average across all such areas pooled. For the merged non-urban areas, the population within each Area Unit was estimated based on the Area Unit's share of the merged area's population, using data on the distribution of the 20–64 year old population, available from *Table Builder* on the Statistics New Zealand website.

[5] The Census collects information on each person's location (Area Unit) 5 years prior to the Census. Where responses identified prior location less precisely than Area Unit, it was assumed that respondents had not moved, unless their response indicated a Territorial Authority, Regional Council, island, or country different from their census-night location.

years earlier (New Zealand-born and earlier migrants), and two of people who were overseas 5 years earlier (returning New Zealand-born and recent migrants).

Geographically-smoothed workforce composition measures are calculated as a proportion of the population living within 10 km of each Area Unit centroid.[6] For businesses operating in more than one location, the composition of their local workforce is calculated as a weighted average of the compositions of each of the areas in which they employ, using the distribution of the firms' employment across the different locations.

4 Results

Table 1 summarises the productivity and workforce composition variables that are the main focus of the analysis. The first two rows show summary statistics for each of the two census years, with comparable figures for the pooled data in the third row. Productivity (MFP) is zero mean within each year, by construction. Workforce characteristics reflect the average composition faced by New Zealand firms. Because firms cluster in high employment-density areas, these 'exposure' means differ from population averages. On average, firms are located in areas where 21.3 % of the population aged 18–65 is foreign-born, with a slightly higher migrant penetration in 2006 than in 2001. Around half of the population (48.2 %) is new to the area, and 13.7 % are degree qualified. Population density increased between 2001 and 2006, due mainly to the greater clustering of firms in densely populated areas.

The final row of Table 1 presents comparable statistics for the subsample of firms that have no employees. In some of the analysis that follows, we control for the composition and turnover of each firm's workforce. These measures are available only for employees, so we are unable to include working-proprietor-only (WPO) firms in that analysis. WPO firms account for around one half of all firms but these are smaller, have lower mean productivity than the total population of firms, and have a standard deviation of MFP that is 0.15 higher.

High productivity firms are disproportionately located in areas with a high proportion of skilled workers, new entrants, and immigrants. The bivariate relationships are summarised in Fig. 1, for 58 Labour Market Areas (LMAs).[7] Figure 1 shows the LMA means of firm-level productivity (MFP) and local workforce composition within a 10 km radius of firms operating in the LMA.[8] The strongest relationship is between productivity and the fraction of the workforce

[6] Measures are smoothed using an Epanechnikov kernel with bandwidth of 10 km. Weights are calculated as ¾*(1-(distance/10)^2) where distance < 10, and zero otherwise.

[7] LMAs are defined using travel to work information following Papps and Newell (2002).

[8] The LMA means are calculated by regressing (a) firm MFP and (b) local workforce exposure, on a full set of LMA share dummies, where the shares represent the proportion of firm employment in each LMA. The coefficients on these share dummies are the measures that are graphed in Fig. 1.

Table 1 Data summary

	N	Productivity (MFP)	Percent migrants within 10 km	Percent new to area within 10 km	Percent degree-qualified within 10 km	ln(Population density within 10 km)
2001	173,022	0.00	19.1 %	44.9 %	11.1 %	4.21
		(0.68)	(10.9 %)	(8.7 %)	(6.6 %)	(2.37)
2006	186,747	0.00	23.4 %	51.3 %	16.1 %	4.47
		(0.67)	(12.8 %)	(6.6 %)	(7.9 %)	(2.37)
Total	359,769	0.00	21.3 %	48.2 %	13.7 %	4.35
		(0.67)	(12.1 %)	(8.3 %)	(7.7 %)	(2.37)
Working proprietor	190,071	−0.05	21.4 %	48.0 %	13.6 %	4.27
Only (WPO)		(0.82)	(12.0 %)	(8.3 %)	(7.8 %)	(2.36)

Notes: standard errors in *brackets*. N is random-rounded (base 3) in compliance with Statistics New Zealand confidentiality rules

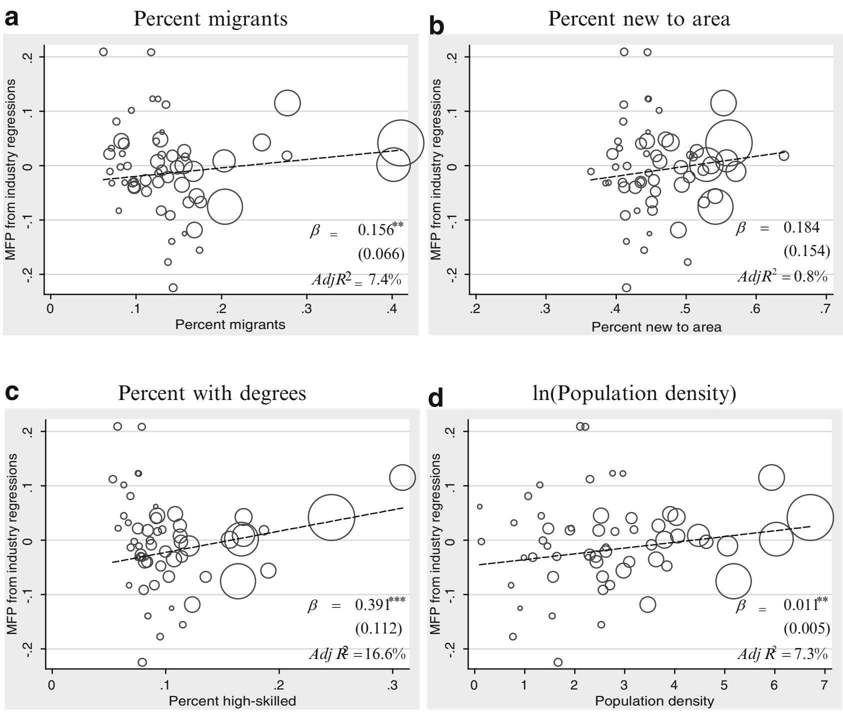


Fig. 1 Relationship between productivity (MFP) and workforce characteristics in 2006. (Notes: Each *symbol* represents a Labour Market Area (LMA). The *size of the symbol* reflects employment in the LMA. *Dashed lines* are weighted regression lines. Workforce composition is measured as an average within a 10 km radius of each Area Unit. See text for fuller explanation. Significance indicators: 1 % (***); 5 % (**))

Table 2 Basic specifications

	mfp: OLS	mfp: OLS	mfp: OLS	mfp: OLS	mfp: OLS	mfp: IV	Δmfp: IV
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Percent migrants	0.292**				0.026	0.024	−0.852
	[0.018]				[0.032]	[0.033]	[0.699]
Percent new to area		0.312**			−0.282**	−0.492**	2.422**
		[0.040]			[0.059]	[0.075]	[0.726]
Percent degree qualified			0.616**		0.586**	0.695**	1.377
			[0.037]		[0.065]	[0.069]	[0.813]
ln(Population density)				0.017**	0.011**	0.014**	0.750*
				[0.001]	[0.002]	[0.002]	[0.323]
Industry intercepts	Y	Y	Y	Y	Y	Y	N
Year intercept	Y	Y	Y	Y	Y	Y	N
Constant	−0.067**	−0.158**	−0.097**	−0.074**	−0.003	0.112**	−0.248**
	[0.005]	[0.020]	[0.007]	[0.006]	[0.023]	[0.029]	[0.088]
Observations	359,769	359,769	359,769	359,769	359,769	359,769	63,069
AdjR2	0.20 %	0.09 %	0.34 %	0.23 %	0.42 %	0.40 %	−0.78 %
UnderId F-stat (p)						266.3 (0)	21.66 (0)
WeakInst F-stat						1653	5.935

Notes: standard errors, clustered on Area Unit, in *brackets* (**; * significant at 1 %; 5 % level respectively). Counts are random-rounded (base 3). For specifications (6) and (7), the instrument set is (5-year) lagged workforce characteristics (including population density). Kleibergen-Paap F-statistics for tests of weak identification and under identification reported. Specification (7) is estimated in first differences (both dependent and independent variables) for firms located and staying in a single Area Unit

with a degree qualification. A 1 percentage point higher degree-share is associated with productivity that is 0.48 % higher ($e^{0.391}$-1). The comparable figure for a higher migrant share is 0.17 %, and for the share of the population new to the area the comparable figure is 0.20 % but not statistically significant. Population density has a clear positive relationship with productivity, with a 10 % higher density associated with productivity that is 0.1 % higher.

4.1 *Regression Analysis*

It is clear from Fig. 1 that LMA size is positively correlated not only with productivity but also with each of the workforce composition measures. In Table 2, we use multivariate regression methods to evaluate the independent contribution of each of these to productivity variation. In the first four columns, we enter each of the workforce composition measures separately into a productivity regression that includes industry and year intercepts. As in Fig. 1, each of the relationships is positive.[9] When the measures are entered together in the same regression (shown in

[9] The coefficients differ from those in Fig. 1 because the regressions in Table 2 use firm-level variation, including within-LMA variation, which is ignored in Fig. 1.

column 5), the influence of density and the proportion of the workforce with a degree qualification remain positive and significant, with coefficients of similar magnitude to those in columns 3 and 4. In contrast, the relationship between productivity and the presence of migrants is small and no longer significant, and the influence of people new to the area is negative. The workforce composition measures are clearly not independent, confirming the need for multivariate analysis to identify the contributions of each.

Columns 6 and 7 present estimates that control for the possible endogeneity of local workforce characteristics. In both columns, actual workforce composition measures are instrumented using their own lags.[10] Column 6 is a level regression, as in previous columns, while column 7 estimates the relationship between MFP and workforce composition in first differences. Consequently, the latter regression is estimated only on the subsample of firms present in both time periods. We further restrict this regression to firms that operate in a single Area Unit and that remain in that Area Unit over time. Thus, the first difference regression, as well as controlling for time-invariant firm characteristics, also removes potentially confounding fixed Area Unit characteristics.

Both sets of IV estimates confirm the general findings of a positive relationship between productivity and both density and degree share. While column 7 represents the more stringent test of the relationships we are interested in, our preferred specification in subsequent tables is the levels IV (column 6). We make this choice since both approaches suggest that workforce characteristics matter, but the first differences approach seriously reduces the sample size, raising questions of the broader applicability of the findings and restricting our ability to estimate effects for smaller subpopulations of firms. Additionally, the increase in the size of coefficients and standard errors associated with instrumenting in the first difference IV is suggestive of a weak instrument problem (despite the estimates passing the Kleibergen-Paap test with an F-statistic of 5.9).

The positive relationship between local skills and productivity may in part reflect the higher average quality of labour that firms employ, rather than an external effect of local skills. Similarly, the negative relationship between productivity and the proportion of the population new to the area may reflect the negative effect of higher average labour turnover at the firm level. In order to control for these firm-level factors, we present, in Table 3, estimates that include measures of firm-level skill and turnover.

Unfortunately, the LBD does not contain comprehensive firm-level information about worker skills. We use a proxy for worker quality derived from a two-way fixed effect model estimated using LEED data. The estimated worker effect is an index of each worker's portable wage premium. For each firm in a given year, we calculate the weighted average of worker fixed effects, using as weights a measure

[10] The specification passes an under identification F-test on the first-stage equation, as shown at the bottom of the table. The Kleibergen-Paap F-statistic for weak identification is also shown, and has a high value of 1653 for column 6, confirming the joint relevance of the instruments. In both cases, the equation is exactly identified, so it is not possible to test for instrument validity.

Table 3 Adding selected firm-level controls – IV estimates

Dependent variable: mfp	Working proprietors only (1)	Employing firms (2)	Employing firms (3)	Employing firms (4)	Employing firms (5)
Percent migrants	0.058	0.036	0.003	0.038	0.006
	[0.044]	[0.025]	[0.022]	[0.024]	[0.023]
Percent new to area	−0.754**	−0.056	−0.022	−0.051	−0.019
	[0.102]	[0.046]	[0.043]	[0.046]	[0.043]
Percent degree qualified	1.028**	0.204**	0.112*	0.204**	0.114*
	[0.087]	[0.063]	[0.051]	[0.063]	[0.051]
ln(Population density)	0.017**	0.011**	0.011**	0.010**	0.011**
	[0.003]	[0.001]	[0.001]	[0.001]	[0.001]
Average worker fixed effects			0.241**		0.236**
			[0.008]		[0.008]
Gross turnover				−0.031**	−0.024**
				[0.003]	[0.002]
Net turnover				0.000	−0.001
				[0.003]	[0.003]
Industry intercepts	Y	Y	Y	Y	Y
Year intercept	Y	Y	Y	Y	Y
Constant	0.081*	0.135**	0.170**	0.162**	0.190**
	[0.040]	[0.018]	[0.017]	[0.018]	[0.018]
Observations	190,071	160,719	160,719	160,719	160,719
AdjR2	0.77 %	2.10 %	3.47 %	2.27 %	3.57 %
UnderId F-stat (p)	291.4 (0)	200.7 (0)	200.6 (0)	200.7 (0)	200.7 (0)
WeakInst F-stat	1262	1862	1873	1860	1872

Notes: standard errors, clustered on Area Unit, in *brackets* (**; * denote significance at the 1 %; 5 % level respectively). Counts are random-rounded (base 3). Only workforce characteristic variables (including population density) are instrumented, using their (5-year) lags. Kleibergen-Paap F-statistics for tests of weak identification and under identification reported

of the workers' employment intensity during the year.[11] Worker effects are estimated only for employees, so WPO firms are excluded from the analysis.

Worker turnover at the firm is also calculated using LEED data, and is based on average quarterly turnover of employees.[12] We include two variables to capture variation in turnover rates. The first is gross turnover, calculated as the sum of accessions and separations during the year. The second is net turnover, which is the difference between accessions and separations. By including both measures, we can interpret the gross turnover as a measure of turnover in excess of what was required to achieve the observed employment growth or decline. Each is expressed as a proportion of average quarterly employment, so that the underlying accessions and separations measures range from −2 to 2.

[11] For further details of the two-way fixed effects estimation method and the employment intensity measure, see Maré and Hyslop (2006).

[12] Excluding quarters related to the first transition into employment and the last transition out of employment.

The first column of Table 3 shows the same IV specification as in column 6 of Table 2, for the subsample of WPO firms, which account for most of the firms excluded from the analysis of firm level labour quality and turnover. The second column of Panel (b) shows the same specification but for firms for which we have labour quality and turnover measures. The coefficients on local workforce measures are significantly smaller for the subset of firms with employees, suggesting that WPO firms may be more affected by local workforce characteristics. They are also estimated with greater precision, reflecting the greater volatility in the productivity measure for self-employed firms. The findings of a positive effect of local skills and population density are maintained.

In the third column of the table, we include the proxy for worker skills within the firm. As expected, the coefficient on local skills is reduced (by 55 %). A 1 percentage point higher share of degree-qualified residents is nevertheless still associated with 0.12 % higher productivity ($e^{0.112}$-1). The relationship between local population density and productivity remains significant and the insignificant coefficients on the percent new to the area and the migrant share do not change materially.

Including controls for labour turnover within the firm has a negligible impact on the other coefficients. Column 4 of Table 3 presents the estimates. Gross turnover is associated with lower productivity, though the effect is modest in size. On average, gross turnover is 53 % of average employment during the year. The coefficient of -0.031 implies that a 10 percentage point increase in this figure is associated with productivity that is 0.3 % lower. Net turnover has a very small and statistically insignificant positive relationship with productivity. The fifth column includes both labour quality and turnover measures, and is our preferred specification. The only local workforce characteristics that are significantly related to productivity are population density (elasticity of 0.01) and the proportion of people with a degree qualification ($\beta = 0.114$).

These results reflect the influence of workforce composition on productivity, averaged across all firms. It is unlikely, however, that all firms are affected equally by the composition of their local workforce. We consider seven subsets of firms, chosen to highlight different accounts of what sort of firms benefit most from local labour and density spillovers. Descriptive statistics for these subsets of firms are presented in Table 4, with IV regression estimates of the relationship between productivity and local workforce composition for each subset presented in Table 5.

Users of high-skilled labour are more likely to benefit from a highly qualified local workforce, through mechanisms such as labour market pooling and matching. The first two subsets of firms shown in Table 4 are firms in industries that employ a high proportion of high-skilled workers, and in industries where research and development expenditure is relatively high.[13] These groups are located in relatively

[13] High-skilled industries are identified from the Business Operations Survey (BOS) as those in which more than 10 % of the workforce are in skilled occupations (*managers and professionals* or *technicians and associate professionals*). The two-digit ANZSIC'96 industries are: B12, C28, D36, D37, F46, G52, I63, I66, J71, K73, K74, K75, L77, L78, N84, O86, P91. High R&D industries are also identified from the BOS as those where more than 0.5 % of industry expenditure is on R&D. The two-digit industries are: A02, B11, B13, C25, C28, C29, L78, N84.

Table 4 Sample statistics for subgroups of firms

	N	mfp	Percent migrants within 10 km	Percent new to area within 10 km	Percent degree-qualified within 10 km	ln(Population density within 10 km)	Average worker fixed effects	Gross turnover	Net turnover
High-skilled industries	46,275	0.03	25.5 %	51.4 %	16.6 %	5.44	−0.03	39.2 %	4.8 %
		(0.48)	(12.4 %)	(6.9 %)	(8.2 %)	(1.80)	(0.26)	(48.6 %)	(36.2 %)
High R&D industries	28,812	0.04	25.0 %	51.0 %	16.3 %	5.25	−0.02	45.5 %	4.8 %
		(0.48)	(12.6 %)	(7.2 %)	(8.3 %)	(2.02)	(0.26)	(55.9 %)	(39.7 %)
Dense areas	40,131	0.06	36.9 %	55.0 %	21.4 %	6.99	−0.04	42.2 %	4.7 %
		(0.45)	(9.1 %)	(2.3 %)	(5.7 %)	(0.30)	(0.25)	(49.2 %)	(37.2 %)
Small firms ($L \leq 5$)	101,754	0.07	20.4 %	47.7 %	13.2 %	4.19	−0.12	63.6 %	5.9 %
		(0.47)	(12.1 %)	(8.5 %)	(7.6 %)	(2.45)	(0.24)	(72.9 %)	(51.5 %)
Large firms ($L > 5$)	58,965	0.04	23.2 %	49.9 %	14.9 %	4.97	−0.05	35.2 %	2.0 %
		(0.37)	(12.4 %)	(7.6 %)	(7.7 %)	(2.08)	(0.17)	(29.1 %)	(13.1 %)
New firms	10,374	0.04	22.8 %	49.7 %	14.6 %	4.76	−0.10	93.5 %	41.5 %
		(0.59)	(12.3 %)	(7.8 %)	(7.7 %)	(2.23)	(0.23)	(73.6 %)	(79.5 %)
Local service industries	46,521	0.02	22.3 %	49.7 %	14.6 %	4.97	−0.13	46.0 %	5.0 %
		(0.40)	(12.0 %)	(7.8 %)	(7.8 %)	(1.98)	(0.19)	(48.0 %)	(36.7 %)
Total	160,719	0.06	21.4 %	48.5 %	13.8 %	4.47	−0.10	53.2 %	4.5 %
		(0.43)	(12.3 %)	(8.3 %)	(7.7 %)	(2.35)	(0.22)	(62.2 %)	(41.8 %)

Notes: standard errors in *brackets*. Counts are random-rounded (base 3) in compliance with Statistics New Zealand confidentiality rules

Table 5 Subgroups of firms – IV estimates

Dependent variable: MFP	High-skilled industries (1)	High R&D industries (2)	Dense areas (3)	Small firms: $L \leq 5$ (4)	Large firms: $L > 5$ (5)	New firms (6)	Local service industries (7)
Percent migrants	0.001	−0.102*	−0.022	−0.003	0.022	−0.074	0.099**
	[0.039]	[0.050]	[0.046]	[0.027]	[0.030]	[0.076]	[0.031]
Percent new to area	−0.059	−0.286**	−0.364	−0.085	0.114*	−0.414**	0.201**
	[0.067]	[0.094]	[0.665]	[0.052]	[0.052]	[0.140]	[0.058]
Percent degree qualified	0.205**	0.432**	0.214	0.232**	−0.072	0.174	0.002
	[0.077]	[0.115]	[0.201]	[0.053]	[0.070]	[0.129]	[0.052]
ln(Population density)	0.007**	0.006*	0.040**	0.014**	0.005**	0.011*	0.005**
	[0.002]	[0.003]	[0.015]	[0.002]	[0.002]	[0.004]	[0.002]
Average worker fixed effects	0.382**	0.440**	0.286**	0.230**	0.285**	0.222**	0.200**
	[0.012]	[0.015]	[0.011]	[0.008]	[0.016]	[0.031]	[0.011]
Gross turnover	−0.075**	−0.052**	−0.048**	−0.023**	−0.097**	−0.019*	−0.056**
	[0.006]	[0.008]	[0.007]	[0.002]	[0.010]	[0.009]	[0.005]
Net turnover	0.015*	0.012	−0.017*	−0.003	0.035*	−0.013	−0.013
	[0.008]	[0.009]	[0.008]	[0.003]	[0.016]	[0.008]	[0.007]
Industry intercepts	Y	Y	Y	Y	Y	Y	Y
Year intercept	Y	Y	Y	Y	Y	Y	Y
Constant	0.111**	0.101**	0.109	0.214**	0.144**	0.402**	−0.065**
	[0.037]	[0.036]	[0.339]	[0.021]	[0.022]	[0.060]	[0.023]
Observations	46,275	28,812	40,131	101,754	58,965	10,374	46,521
AdjR2	5.62 %	7.32 %	3.33 %	3.77 %	4.18 %	4.53 %	2.63 %
UnderId F-stat (p)	117.6 (0)	191.5 (0)	12.76 (0)	222.2 (0)	155.5 (0)	158.7 (0)	109.3 (0)
WeakInst F-stat	1169	1420	3.199	1722	1869	1455	1555

Notes: standard errors, clustered on Area Unit, in *brackets* (**; * denote significance at the 1 %; 5 % level respectively). Counts are random-rounded (base 3). Only workforce characteristic variables (including population density) are instrumented, using their (5-year) lags. Kleibergen-Paap F-statistics for tests of weak identification and under identification reported

high density areas with higher-than-average proportions of migrants, degree-holders, and newcomers. They also have slightly higher-than-average labour quality, as captured by average worker fixed effects, and lower worker turnover rates.

The first two columns of Table 5 show regression estimates for these two groups. The estimates show a strong positive association of productivity with the percent of the local population with degree qualifications (coefficients of 0.205 and 0.432 respectively, compared with 0.114 overall). As expected, the coefficients on average worker quality within firms are also strongly positive for these two groups of firms, 0.382 and 0.440 respectively, compared with 0.236 overall, confirming the direct effect on measured productivity of having higher quality labour input within such firms.

Many theories of local labour market spillovers emphasise the operation of these effects in dense urban markets where interactions are greatest. In the third column, we show estimates for the quarter of firms operating in the areas with the highest population density. Within this group, the density of population is positively linked to productivity ($\beta = 0.040$) – more strongly than it is for firms generally. This suggests that there may be positive sorting on the basis of returns to density. The firms that have the most to gain from density are the ones that are disproportionately located in higher density areas. However, there are no significant spillovers from the composition of the local workforce for these firms. Firms in dense areas face even higher proportions of migrants, newcomers, and degree-holders than do firms in high skill or high-research and development industries, yet there is no significant relationship between productivity and these composition measures in dense areas.

Existing studies point to the importance of dense urban environments especially for small and newly established firms (Duranton and Puga 2001). Columns 4 and 5 of Table 5 show estimates for two size-classes of firms – those with employment of five or fewer, and those with employment greater than five. The advantages of operating in a dense area do appear to be more modest for larger firms, with the coefficient on population density being only half as big as for smaller firms. Smaller firms benefit more from being in a highly skilled local labour market, with a coefficient of 0.232 on the percent with degree qualifications. New firms (column 6) also benefit relatively strongly from being in densely populated areas, and appear to have lower productivity in areas with many newcomers ($\beta = -0.414$).

The composition of the local workforce may affect the pattern of demand for local goods and services as well as the operation of the labour market. The final column of Table 5 contains estimates for firms in industries that provide a high proportion of their output locally.[14] These firms are more productive in areas where new entrants ($\beta = 0.201$) and migrants ($\beta = 0.099$) are a relatively high proportion of the local workforce. The effect of being in a high-skilled area is small and statistically insignificant. For local services firms, the composition of the local workforce appears to raise productivity primarily through output markets rather than through factor markets.

5 Discussion

Overall, our findings confirm the existence of a significant relationship between economic performance and the spatial context in which firms operate. We find a positive bivariate relationship between productivity and each of the workforce

[14] Industries are identified from Statistics New Zealand's most recent published input–output tables (the 126 industry, 1996 classification) as those with approximately half or more of their output used directly by the household sector (defined as households plus the ownership of owner-occupied dwellings industry). We then drop Financial and Insurance Services (ANZSIC K) from the resulting industry group on the basis that they provide services largely to households outside the local area.

composition measures that we consider. Multivariate analysis highlights workforce qualifications as the single most important of the measures. In contrast, the proportion of the population that is new to the area, and the proportion that are foreign born are not positively related to firm productivity. This finding is maintained once we control for the possible endogeneity of workforce composition. We confirm the robustness of our findings to the inclusion of additional controls for firm-level labour quality and labour turnover.

We also provide separate estimates for various subgroups of firms in order to detect heterogeneous impacts of local workforce characteristics. The estimated overall relationship between economic performance and local workforce composition conceals marked differences among firms in the size and nature of local labour spillovers. Overall, local population density is positively related to productivity. The relationship is, however strongest for firms operating in the densest areas. In fact, in dense areas, the composition of the local workforce is not significantly related to productivity once we have controlled for density. The benefits of density are also relatively strong for small firms and for new firms, consistent with firm life cycle models of agglomeration (Duranton and Puga 2001).

Local workforce skills contribute most strongly to productivity for small firms, and for firms in industries with high levels of research and development or high usage of skilled workers. This is consistent with the advantages of thick labour markets for skilled and specialised workers. However, it may also indicate positive sorting based on the returns to local skill spillovers. Finally, the presence of newly arrived residents aids productivity most strongly for firms providing local services. This suggests a different source of local spillovers for these firms, linked more to product market effects rather than knowledge spillovers.

Our findings have two main implications for studies of the relationship between economic performance, geography, and institutional and cultural factors. The first relates to the correlation among alternative indicators of local workforce composition. The second relates to the heterogeneity of impacts across different firms.

The correlation among local workforce measures implies that caution is needed when estimating the impact of any one measure on economic performance. For instance, while it is reasonable to use the percent of migrants in an area as a proxy for cultural diversity, our estimates in Table 2 suggest that the positive relationship between migrants and productivity is largely accounted for by other factors. In particular, population density and skill composition are more strongly related to economic performance than is cultural diversity as measured by the percent of migrants. Analyses that focus on migrant presence alone may wrongly attribute the effects of average skill or skill diversity to cultural diversity.

The heterogeneity of impacts implies that care is needed when interpreting aggregate studies that relate regional workforce characteristics to average regional economic performance. Some of the positive relationship may reflect sorting of firms that benefit most from local skills or diversity. The estimated impact will therefore overstate the causal impact that would result from changes in the local workforce. The estimation of heterogeneous impacts can also shed light on the mechanisms by which local workforce spillovers operate, as with the implied product market effects of entrants on local service firms.

Further studies using firm-level microdata are needed to uncover which firms benefit from which aspects of local workforce composition, and to understand more fully the different mechanisms that are the source of local spillovers for particular types of firms.

Acknowledgments We thank the editors, Riccardo Crescenzi and Marco Percoco, for encouraging our contribution, and for comments on an earlier draft. This work has been partially funded as part of the New Zealand Department of Labour's Economic Impacts of Immigration research programme. Thanks to participants at the CRENoS/University of Southampton workshop on "Determinants and Effects of Interregional Mobility", and the 2010 ERSA Congress, and to Steve Stillman for comments on earlier versions of this paper. Access to the data used in this study was provided by Statistics New Zealand in accordance with security and confidentiality provisions of the Statistics Act 1975 and the Tax Administration Act 1994. The results in this paper have been confidentialised to protect individual businesses from identification. See Maré and Fabling (2011) for the full disclaimer.

References

Amiti M, Pissarides CA (2005) Trade and industrial location with heterogeneous labor. J Int Econ 67(2):392–412

Ciccone A, Hall RE (1996) Productivity and the density of economic activity. Am Econ Rev 86 (1):54–70

Duranton G, Puga D (2004) Micro-foundations of urban agglomeration economies. In: Henderson VJ, Thisse JF (eds) Handbook of regional and urban economics. North-Holland, Amsterdam

Duranton G, Puga D (2001) Nursery cities: urban diversity, process innovation, and the life cycle of products. Am Econ Rev 91(5):1454–1477

Fabling R (2009) A rough guide to New Zealand's Longitudinal Business Database. Global COE Hi-stat discussion paper GD09-103. Hitotsubashi University, Tokyo

Fabling R (2011) Keeping it together: tracking firms in New Zealand's Longitudinal Business Database. Motu working paper, pp 11–01

Fabling R, Maré DC (2011) Production function estimation using New Zealand's Longitudinal Business Database. Motu working paper (forthcoming)

Hunt J, Gauthier-Loiselle M (2010) How much does immigration boost innovation? Am Econ J Macroecon 2(2):31–56

Jacobs J (1969) The economy of cities. Vintage Books, New York

Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. Quart J Econ 108:577–598

Maré DC, Fabling R (2011) Productivity and local workforce composition. Motu working paper 11–06

Maré DC, Hyslop DR (2006) Worker-firm heterogeneity and matching: an analysis using worker and firm fixed effects estimated from LEED. Leed research report, Statistics New Zealand, Wellington. Available online at www.stats.govt.nz, Accessed 1 June 2011

Marshall A (1920) Principles of economics, 8th edn. Macmillan, London. Available online at http://www.econlib.org/library/Marshall/marP1.html. Last accessed 4 Nov 2005

Moretti E (2004a) Human capital externalities in cities. In: Henderson JV, Jacques-Francois T (eds) Handbook of urban and regional economics, vol 4, Cities and geography., pp 2243–2291

Moretti E (2004b) Workers; education, spillovers, and productivity: evidence from plant-level production functions. Am Econ Rev 94(3):656–690

Moulton BR (1990) An illustration of a pitfall in estimating the effects of aggregate variables on micro unit. Rev Econ Stat 72(2):334–338

Overman HG, Puga D (2010) Labour pooling as a source of agglomeration: an empirical investigation. In: Glaeser EL (ed) Agglomeration economics. Chicago University Press, Chicago (forthcoming)

Papps KL, Newell JO (2002) Identifying functional labour market areas in New Zealand: a reconnaissance study using travel-to-work data. IZA discussion paper 443. Institute for the Study of Labor, Bonn. Available online at ftp://repec.iza.org/RePEc/Discussionpaper/dp443.pdf. Last accessed 30 June 2005

Rauch JE (1993) Productivity gains from geographic concentration of human capital: evidence from the cities. J Urban Econ 34(3):380–400

Smith A (1904) An inquiry into the nature and causes of the wealth of nations, vol 5. Methuen and Co, London

Rosenthal S, Strange W (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson VJ, Thisse JF (eds) Handbook of regional and urban economics, vol 4. North-Holland, Amsterdam

Zucker LG, Darby MR (2009) Star scientists, innovation and regional and national immigration. In: Audretsch DB, Litan RE, Strom RJ (eds) Entrepreneurship and openness: theory and evidence. Edward Elgar, Cheltenham, pp 181–212

Regional Growth in Central and Eastern Europe: Convergence, Divergence and Non-linear Dynamics

Vassilis Monastiriotis

Abstract Regional disparities in the countries of Central and Eastern Europe have risen substantially in the process of transition and internationalisation of their economies. Still, prima facie evidence of beta-convergence is often found in the CEE data. In this chapter we develop a hybrid model of regional growth that incorporates key aspects of three central theories of regional growth, namely neoclassical convergence, cumulative causation and the regional Kuznets curve. We find partial evidence in support of all three approaches, but ultimately our results suggest that the regional growth process in CEE has not been obeying the rule of convergence, either monotonically (neoclassical) or through a bell-curve (regional Kuznets curve). The specific economic, institutional and cultural factors that account for this, ought to be the subject of continuing research in the filed.

1 Introduction

Regional disparities in the countries of Central and Eastern Europe (CEE) have risen sharply over the last two decades. With them, strong patterns of polarisation emerged, as the process of national convergence, stimulated by increasing openness and economic-political integration, has not been accompanied by a similar trend for cross-regional equilibration. Besides their policy relevance, these developments are particularly important for academic inquiry, as they challenge simple concepts of convergence and instantaneous equilibration, bringing to the fore some fundamental theoretical questions. Is the process of development inherently uneven? Is, inversely, convergence an automatic process driven by the properties of the production technology (diminishing returns)? Or is growth an endogenously-driven

JEL code: R11, R15, O11, O18, O47

V. Monastiriotis (✉)
European Institute, London School of Economics, Houghton Street, WC2A 2AE, London, UK
e-mail: v.monastiriotis@lse.ac.uk

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_5,

cumulative process, whereby leading economies, boosted by their past performance, are able to maintain and enhance their advantages over less developed ones? And is the process of convergence and divergence conditioned on the level of national development?

In recent decades, the study of these questions has been dominated by the so-called 'convergence hypothesis'. Based on the Solow one-sector growth model under the assumptions of a common technology, diminishing returns and no systematic external shocks (Barro and Sala-i-Martin 1991), the convergence hypothesis asserts that economies starting from higher development levels experience slower growth rates. As a consequence, less developed economies eventually (although, in the theoretical form of the model, asymptotically) catch-up, leading to a long-run stable equilibrium (steady-state) of convergence. Although more recent contributions have sought to move beyond the simplicity of the convergence hypothesis, either methodologically, by examining distributional dynamics (Magrini 1999; Rey and Janikas 2005), or substantively, looking in particular at the role of knowledge and institutions, including aspects of society and governance (Storper 1997; Martin and Sunley 1998), cultural diversity (Bellini et al. 2012), innovation/technology (Diliberto and Usai 2012; Rodriguez-Pose and Comptour 2012), and globalisation/openness (Nijkamp and Siedschlag 2008; Capello and Fratesi 2012), the macroeconomic analysis of regional growth is still driven by the simplistic notion of convergence. Two broad intellectual traditions in the analysis of regional economic performance have suffered as a result: those relating to the theory of cumulative causation (Myrdal 1957; Kaldor 1970) and ones that, deriving from the development economics tradition, emphasise the interwoven relationship between the processes of regional growth and national development (Williamson 1965).

In CEE this represents a significant deficit – as the collapse of central planning, and the processes of deindustrialisation, restructuring, marketisation and openness that it triggered, led to an unfolding of national and regional dynamics at an accelerated pace,[1] which offered itself for the study of competing theories of regional growth and of the related processes of equilibration and cumulative causation. As other contributions in this book show (see the chapters by Capello and Fratesi, Rodriguez-Pose and Comptour, and Bellini et al.), regional growth dynamics are significantly influenced not only by exogenous processes of openness and integration, which lead to sometimes radical reconfigurations of economic space, but also by locally embedded characteristics such as technological capacities, production networks, knowledge and cultural diversity. In this sense, the question of whether the relative economic performance of regions obeys an exogenously given rule (convergence to steady-state) or is instead conditioned on relational factors and national processes (of growth and development) is

[1] For example, countries like Hungary and the Czech Republic transformed their economies in a matter of just over 10 years, jumping from a level of development around 20 % of the EU15 average in 1995 to around 40 % in 2007 and shifting their export specialisations so that by 2005 over 80 % of their trade with the EU15 was of intra-industry character.

particularly important for evaluating the role of policy in strengthening and diversifying regional advantages to address cross-regional imbalances.

This chapter utilises this unique experience of fast transition, openness and restructuring in the CEE to examine how patterns of regional growth and disparity, as they developed over the last two decades in the region, relate to three analytically distinctive processes identified in the theoretical literature: the neoclassical convergence hypothesis, the cumulative causation theory and the evolutionary approach of the regional Kuznets curve. The focus here is not with the causal explanation of these patterns, in the sense of seeking to identify the specific variables that account for them (i.e., the cultural, technological and institutional factors referred to above), but rather with unveiling the underlying growth process that best describes these patterns. The next section offers a brief review of the evolution of regional disparities in CEE. Section 3 considers the three theoretical approaches and discusses how these can be instrumentalised in a nested model of regional growth. Section 4 presents the empirical investigation, while the last section concludes with some implications for theory and policy.

2 Transition, Accession and Regional Growth in the CEECs

Descriptive studies examining the extent and evolution of regional disparities in the CEE countries have found consistently that these have grown significantly over the last two decades. The rise in inequalities has been evident from the early stages of transition (Petrakos 1996; Römisch 2003), but it continued throughout the period and in some cases intensified (Ezcurra et al. 2007; Kallioras and Petrakos 2010). There is broad consensus in the literature, largely attributing these developments to the significant geographical and sectoral reallocation that has taken place in CEE over the last two decades. One the one hand, there is a notable shift of industrial activity towards metropolitan regions and regions bordering the EU (Petrakos and Economou 2002; Iara and Traistaru 2003), stimulated partly by the self-selective inflow of foreign investments in these areas (Altomonte and Resmini 2002; Tondl and Vuskic 2003). Trade integration also played a role in this, by favouring regions with significant specialisations and agglomeration economies, relative concentration of skilled labour and vibrant product demand (Traistaru et al. 2003; Resmini 2007). On the other hand, the literature identifies a process of structural change across sectoral lines, both in terms of internal structures (sectoral compositions) and external competitiveness (trade specialisations) (Resmini 2003; Niebuhr and Schlitte 2009; Kallioras and Petrakos 2010). Analyses along these lines confirm the inherent link between spatial and structural restructuring, finding that regions which have successfully restructured and thus benefited most from integration are those located closer to the EU borders and to metropolitan areas or large agglomerations.

Despite this general trend, econometric studies following the convergence approach often find evidence of convergence, at least in cross-country – cross-regional analyses (indicating regional convergence across the CEE space but not necessarily within each CEE country). Herz and Vogel (2003) use data for 31 regions across the CEECs and find evidence of divergence in the early transition period and of conditional convergence more recently. Using Eurostat data and examining cross-national and cross-regional convergence across the EU New Member States, Niebuhr and Schlitte (2009; at the NUTS2 level for the period 1995–2000) and Paas et al. (2007; at the NUTS3 level for the period 1995–2002), find evidence of fast cross-national convergence across countries and regional divergence or stability within countries – with an overall slow convergence of regional incomes across the CEE and EU countries. Using the same database in a simple NC framework, Petrakos et al. (2005a) also find evidence of convergence. Similar are the results obtained by Del Bo et al. (2010), who use NUTS2-level Cambridge Econometrics data in a spatial econometrics framework and find evidence of both conditional and (marginally) unconditional convergence across the CEE regions. Evidence of convergence is also obtained in country-specific studies (e.g., Banerjee and Jarmuzek 2010, for Slovakia).

In an analysis that departs somewhat from the neoclassical approach, Petrakos et al. (2005b) find simultaneous evidence of short-run divergence and long-run convergence, with the level of disparities following a pro-cyclical path and a long-run convergent trend. Kallioras (2010) shows that convergence trends are conditioned on the size of the regional economies, with evidence of divergence when population size is taken into account and evidence of convergence otherwise.[2] Given that population is typically higher in more advanced and more dynamic regions, these findings can be interpreted as evidence signalling intra-country polarisation and, possibly, club convergence: smaller (and poorer) regions tend to converge to their own steady-state, but larger regions tend to follow different, more dynamic, paths. Direct evidence for this, in the form of club convergence, with strong regional convergence within and persistent divergence across clubs, has been offered recently by Artelaris et al. (2010; for within-country clubs) and earlier by Fischer and Stirböck (2006; for cross-country clubs).

These patterns of polarisation and divergence are also reflected in the NUTS3-level dataset used here, covering the period 1990–2008.[3] In this dataset, regional disparities in GDP per capita rose by 80 % over the period, although this increase was not uniform across countries (ranging from 15 % in Bulgaria to 196 % in the Czech Republic). The increase was lower in terms of labour productivity (around 30 %) but, with the exception of Slovenia, it was evident in all CEE countries. Polarisation also increased significantly throughout the period and by 2007 GDP per capita in the-highest-income region in the average country was about 2.5 highest than in the corresponding median region. The lack of convergence is also manifested in the high persistence of regional rankings, with rank correlation coefficients across the sample at 87 % and 71 % for regional incomes and productivity, respectively. In the next section we discuss the key competing theoretical approaches that can account for these developments in regional growth performance in CEE.

[2] See also Petrakos and Artelaris (2009) for similar evidence for the pre-2000 EU member states.

[3] Source: Cambridge Econometrics European Regional Database (http://www.camecon.org).

3 Conceptualising Relative Regional Growth

As mentioned earlier, our focus is with three competing theoretical approaches: neoclassical convergence, cumulative causation and the regional Kuznets curve. We examine each of these sequentially.

3.1 Neoclassical Convergence

As is well known, the empirical formulation of the convergence hypothesis makes regional growth a function of initial regional incomes, as follows:

$$\Delta(y-l)_{i,t} = b_0 + b_1(y-l)_{i,t=0} \tag{1}$$

where y is the log of output, l is log employment, i and t index regions and time, respectively, and $b_1 < 0$, reflecting catch-up convergence. Two extensions of this model are possible. First, adding other controls, to capture region-specific structural characteristics (such as technology, preferences, propensities to save, etc.), takes us to the notion of "conditional convergence", where regions converge towards a region-specific steady-state. Second, by splitting the regions across a relevant dimension (e.g., large-small, metropolitan-peripheral, rich-poor, etc.), one can examine the so-called "club convergence" hypothesis, where regions converge towards two (or more) club-specific steady states (with club membership defined on the basis of similarity in initial conditions or time-persistent characteristics), resulting in polarisation in the distribution of regional incomes.[4] Thus, the general formulation of the neoclassical convergence story can be written as

$$\Delta(y-l)_{i,t} = (b_0 + b_0^C) + b_1(y-l)_{i,t-k} + b_2 C_i (y-l)_{i,t-k} + b_3 I_{i,t-k}(y-l)_{i,t-k} \tag{2}$$

where $k \in \{1, \mathrm{T}\}$, C is a binary variable indicating membership into a club and I is a variable summarising region-specific characteristics.[5] By setting $b_3 = 0$ we move from conditional to unconditional convergence and by setting $b_0{}^C$, $b_2 \neq 0$ we move from universal to club convergence.

As is well discussed in the literature, this model assumes a monotonic (although asymptotic) process towards equilibration, largely driven by the law of diminishing

[4] See the *Controversy* section at the July 1996 issue of the Economic Journal (No437) for an interesting discussion of the different notions of convergence. See also the excellent survey by Islam (2003).

[5] In panel-data formulations, the initial levels are typically replaced by a reasonably spaced time-lag (usually, simply t-1), which however captures then only short-run dynamics (mean-reversal –see Islam 2003) and the initial conditions are subsumed in the regional fixed effects included in the model.

returns and the assumption about constant returns to scale. Relaxing either of these assumptions (monotonicity and constant returns) takes us to two distinctively different theoretical traditions and thus two substantively different formulations of the regional growth process, as discussed in the remainder of this section.

3.2 Cumulative Causation and Increasing Returns

Allowing for increasing returns to scale (IRS) opens the possibility of equilibrium divergence, even within the convergence framework (as in the endogenous growth theory, with $b_1 \geq 0$ in the formulation of Eq. 2). Estimates of divergence in this context are often interpreted as reflecting cumulative causation mechanisms (see Petrakos et al. 2005b, and Cibulskiene and Butkus 2007 for relevant discussions), whereby richer regions grow permanently faster than less developed ones. Inversely, evidence of beta-convergence is often taken as a refutation of the cumulative causation story. This is however inaccurate. In Myrdal's (1957) original formulation, the circular process is not about a positive relationship between growth and initial incomes (beta-divergence), but rather about a positive relationship between past and current rates of growth: a circular process of self-perpetuating growth, underlined by institutional, cultural and economic factors, irrespective, in a way, of initial incomes. This is because initial incomes capture only partly the initial advantages in regional conditions and characteristics; and because the latter both *generate* and *maintain* a region's growth advantage.

A somewhat different formulation of this relationship can be derived using Verdoorn's Law. [6] Kaldor's (1970) version of this emphasises the role of demand, in the form of total-economy output growth, in generating increasing returns and productivity enhancing efficiency gains. More recently, in the urban and spatial economics tradition increasing returns derive from the supply-side, through the knowledge- or technology-enhancing role of agglomeration – thus making productivity growth a function, not of the growth in the *volume* of output, but of output *density* (see Ciccone and Hall 1996). Putting together this, admittedly very diverse, family of cumulative growth approaches, leads to the following generic relationship:

$$\Delta(y-l)_{i,t} = c_0 + c_{11}y_{i,t-k} + c_{12}y_{i,t-k-1} + c_2\Delta(l)_{i,t-k} + c_3s_i \qquad (3)$$

[6] Given concerns about the endogeneity of output growth in the Verdoorn equation (Rowthorn 1975), a dynamic specification including past values of output growth may appear more appropriate. Note also that local output growth can be replaced by output growth across a spatial field, thus linking the non-spatial formulation of Vernoorn's Law to NEG's emphasis on market potential (see Angeriz et al. 2008).

where the *t-k* notation has been maintained to allow for dynamic links at longer time-horizons and *s* represents the physical size (e.g., square hectares) of the regional economy. Setting $c_{11} = -c_{12}$ and $c_2 = c_3 = 0$ reproduces the Kaldorian formulation. Setting $c_{11} = -c_{12} = -c_2$ and $c_3 = 0$ brings us instead to the Myrdalian formulation. Finally, setting $c_{12} = c_2 = 0$ and $c_{11} = -c_3$ reproduces the agglomeration economies approach of Ciccone and Hall (1996).

3.3 Non-monotonic Convergence and the Regional Kuznets Curve

Relaxing in turn the monotonicity assumption allows for the possibility that regional growth may vary with the level of national development and each region's relative position within the national economic space. This is essentially a link proposed within the development economics tradition, following the seminal contribution of Kuznets (1955) and its regional adaptation by Williamson (1965) – although recent contributions in urban and spatial economics (Henderson et al. 2001; Duranton and Puga 2004) are also consistent with this. According to Williamson's 'regional Kuznets curve' (henceforth, RKC), regional disparities, originally low for low levels of development, rise sharply as the process of national development accelerates and economic activity concentrates to take advantage of scale and agglomeration economies. In later phases, as connectivity across space improves (e.g., through infrastructure investment or, in NEG terms, declining transportation costs) and congestion diseconomies start biting, new growth opportunities emerge in more peripheral regions and disparities start subsiding.[7]

Traditionally, empirical tests of the RKC hypothesis have examined, across international datasets, the link between national incomes and regional *disparity* (Barrios and Strobl 2006; Ezcurra and Rapun 2006; Francois and Rojas-Romagosa 2008; Persyn and Algoed, 2009). It is possible, however, to derive a RKC-consistent relationship in a regional *growth* formulation, by modelling the latter as a quadratic function of the interaction between national and regional incomes, as follows:

$$\Delta(y-l)_{i,t} = d_0 + d_1[(y-l)_{i,t-k}(y^N - n^N)_t] + d_2[(y-l)_{i,t-k}(y^N - n^N)_t^2] \qquad (4)$$

where $y^N–n^N$ is the log of national income (GDP per capita[8]) and $d_1 > 0$, $d_2 < 0$ to account for the fact that at low (high) levels of national development regional

[7] A weaker version of the RKC hypothesis has been proposed more recently (Higgins and Williamson 2002), which acknowledges that exogenous factors (trade openness, technological progress, etc.) may condition this relationship, making the RKC divergence-convergence path less deterministic.

[8] As national GDP experiences year-to-year fluctuations that are not reflective of (changes in) the level of national development, an alternative formulation of Eq. 4 could replace $y^N–n^N$ with a time-trend, T.

disparities rise (fall), i.e., that growth is first faster and then slower for the more developed regions.

Despite the similarity of the theoretical processes described by the RKC and cumulative causation approaches,[9] it is in fact the neoclassical model that relates more directly to this formulation – under the following rationale. At early stages of national development new technologies are introduced unevenly across space and thus disparities rise: similar economies may still exhibit a tendency for convergence, but differences in technology will lead to conditional only convergence and thus to unconditional divergence. As the national economy matures, however, new technologies diffuse across space and the sub-national economies become more similar, so that absolute convergence kicks-in. Thus, the speed of convergence is itself a quadratic function of the national income

$$b_{1t} = e_0 + e_1(y^N - n^N)_t + e_2(y^N - n^N)_t^2 \tag{5}$$

with $e_0, e_2 < 0$ and $e_1 > 0$ so that

$$\Delta(y-l)_{i,t} = b_0 + e_0(y-l)_{i,t-k} + e_1[(y-l)_{i,t-k}(y^N - n^N)_t] + e_2[(y-l)_{i,t-k}(y^N - n^N)_t^2] \tag{6}$$

which is equivalent to Eq. 4 for $d_1 = e_1$, $d_2 = e_2$ and $e_0 = 0$. There is one major difference, however, between these two models. In the RKC version (Eq. 4), the regional variable of interest is relative, measured as the region's distance from the national level of labour productivity. In the amended neoclassical formulation (Eq. 6) this variable is instead specified in absolute terms, as the assumption is that the national level of development affects the speed of convergence and thus the elasticity of regional growth to (absolute) past levels of productivity.

4 Empirical Results

Informed by the models derived in the previous section, our empirical investigation seeks to unveil the extent to which different theoretical hypotheses and formulations are validated by the CEE experience of regional growth over the last two decades. We start by testing the neoclassical model. As shown in Table 1, the convergence hypothesis is broadly validated by the data. Even without taking into account national differences in growth rates (first row), evidence of

[9] In the initial phases of development, RKC describes a process of divergence which is cumulative and thus consistent with the cumulative causation approach. However, while the former assumes a strict deterministic path towards convergence as national development matures, for the latter 'return to convergence' is neither deterministic nor inevitable –and it is not directly linked to the level of national development.

Table 1 Regional growth in CEE and neoclassical convergence

	Constant		Lagged productivity		
Model	All regions	Top 25%	All regions	Top 25%	R^2
Unconditional convergence					
Cross-country NC	0.219***		−0.092***		0.062
(OLS)	(0.012)		(0.006)		
Within-country NC	0.379***		−0.090***		0.127
(Country FEs)	(0.018)		(0.006)		
NC with common business	0.356***		−0.224***		0.161
Cycle (year FEs)	(0.016)		(0.009)		
Country-and-time independent	0.503***		−0.259***		0.237
(Country and year FEs)	(0.020)		(0.010)		
Conditional convergence					
Conditional: on regional	0.645***		−0.313***		0.209
Characteristics (within FE)	(0.021)		(0.011)		
... add time dummies	0.960***		−0.413***		0.317
	(0.031)		(0.012)		
... and replace with 5-year lag	0.104***		−0.048***		0.096
	(0.008)		(0.005)		
Club convergence					
Club convergence: speed	0.530***	−0.0004	−0.290***	0.040**	0.253
(Includes C&Y FEs)	(0.020)	(0.035)	(0.011)	(0.015)	
Club convergence: steady-state	0.517***	0.0853***	−0.283***		0.251
(includes C&Y FEs)	(0.020)	(0.010)	(0.010)		

Notes: Standard errors in *parentheses*. *, **, *** show significance at 10 %, 5 % and 1 %, respectively

neoclassical convergence across the 190 CEE regions of our sample is obtained. In fact, country differences in growth rates, although significant (see the rise in explanatory power between the models in rows 1 and 2), affect only marginally the obtained speed of convergence (the coefficient drops from 9.2 % to 9.0 %). In contrast, the speed of convergence changes significantly (increasing by 2.5 times) when we account for temporal variations in growth rates, i.e., for the position of the CEE business cycle (third row). Additionally, when we condition regional growth rates on fixed regional characteristics (captured here by a set of regional fixed effects), the estimated speed of convergence increases further, especially when temporal controls are also included (rows 5 and 6, respectively). This evidence is consistent with both convergence processes: the CEE regions converge fast towards their own steady-states (conditional convergence), which are also convergent across space (unconditional convergence).

On closer scrutiny, however, this evidence appears much weaker. First, when lagged productivity is replaced by its 5-year lag (third row from bottom), to account for the fact that in data with year-to-year variation evidence of convergence (in the long-run) may be convoluted with evidence of mean reversal (in the short-run), the convergence coefficient drops by over 800 %, suggesting that much of the evidence on convergence is driven by short-run dynamics.[10] Second, when we allow for variable speeds of convergence across advanced and less advanced regions,[11] evidence of club convergence emerges. Specifically, by interacting club membership with lagged productivity (penultimate row of Table 1), we find that high-productivity regions convergence more slowly to the steady-state (which, in this model appears to be common, as the club dummy is not significant statistically). When we allow for differences in steady-state growth but not in speeds of convergence (last row), the results provide strong evidence of differentiation, with steady-state productivity growth being over 15 % higher for high-productivity regions compared with the rest.

It thus appears that processes of divergence, polarisation and cumulative causation may well be in place. To explore this, in Table 2 we examine more formally the case of cumulative causation. We start with a very simple specification, which follows the Myrdalian argument of circular (self-reinforcing) growth, regressing productivity growth on its 1-year lag. When not including temporal or spatial controls (not shown), the lagged term is highly insignificant and the model has no explanatory power ($R^2 = 0.0002$). Controlling for time-effects produces a circular causation effect which is highly significant statistically but in economic terms rather trivial (see col.1). The results become notably stronger, however, when we include temporal lags of a longer horizon (columns 2–5). The fit of the model increases and all lagged terms are significant irrespective of estimation method (OLS, DVLS, Within, Arellano-Bond) and whether or not we include country and time dummies (columns 2–3), region fixed effects (col.4) or controls for the possible autocorrelation between the fixed effects and the lagged regressors (col.5). Although the evidence is less than overwhelming, some support for the circular causation mechanism is nevertheless obtained.

Departing from what is essentially a simple test of persistence, in columns 6–10 we examine the Kaldorian specification. The obtained Verdoorn coefficient in col.6 is very high[12] and becomes even larger when temporal and national controls are included (col.7), indicating clearly the presence of a CC mechanism relating to increasing returns. The self-reinforcing nature of this mechanism is confirmed by

---

[10] We have also tested the convergence hypothesis using other time-lags, ranging between 2 and 10 years. For lags of 3 years or more the results are consistent with the pattern depicted by the 5-year lag.

[11] Results reported here concern a club defined by regions whose productivity belonged to the top-25 % of their national distribution of regional productivities in each and every of our sample years. The results are very similar under alternative definitions: (a) as above, but using the median or mean as the membership threshold; (b) as in (a) but with the condition applying to the majority of years (or alternatively to at least 1 year) rather than to all years; (c) as in (b) but with the reference value being the national productivity level of each particular year, rather than a point in the distribution of regional productivities across all years.

[12] Previous estimates are in the area of 0.5, but are usually derived from models where the dependent variable measures productivity growth in manufacturing. Our dependent variable measures GVA in all sectors and, in the absence of data on capital, it refers to labour, than total factor productivity growth.

Table 2 Regional growth in CEE and processes of cumulative causation

Model	OLS	OLS	DVLS	Within	xtabond	OLS	DVLS	DVLS	DVLS	DVLS	OLS	DVLS	Within	DVLS	DVLS
Productivity growth t-1	0.0377	0.1490	0.1242	0.1041	0.0864									0.0472	0.0689
	(0.014)	(0.010)	(0.011)	(0.011)	(0.012)									(0.010)	(0.009)
Productivity growth t-2		0.0822	0.0590	0.0489	0.0409									0.0070#	
		(0.007)	(0.007)	(0.008)	(0.008)									(0.009)	
Productivity growth t-3		0.0255	0.0230	0.0143^	0.0092									−0.0448	
		(0.006)	(0.006)	(0.006)	(0.006)									(0.006)	
Output growth						0.8691	1.114		0.7058					0.7680	1.103
						(0.027)	(0.026)		(0.017)					(0.018)	(0.018)
Output growth t-1								0.0591	0.2341	0.1223				0.1398	
								(0.020)	(0.011)	(0.014)				(0.015)	
Output growth t-2									0.0786	0.0652				0.0049#	
									(0.006)	(0.008)				(0.012)	
Output growth t-3									0.0740	0.0660				0.1081	
									(0.006)	(0.008)				(0.011)	
Output density											0.0035#	0.0107	0.3382	−0.0044	−0.004^
											(0.003)	(0.003)	(0.021)	(0.001)	(0.002)
Fixed effects	Year	–	Year and country	Year and region	Year and region	–	Year and country	Year and country	Year and country	Year and country	–	Year and country	Year and region	Year and country	Year and country
Constant	0.121	0.023	−0.002#	0.025	0.021	0.024	0.348	−0.170	−0.003#	0.006#	0.047	0.265	0.334	−0.031	0.221
	(0.013)	(0.001)	(0.006)	(0.005)	(0.005)	(0.003)	(0.015)	(0.015)	(0.005)	(0.006)	(0.005)	(0.020)	(0.031)	(0.005)	(0.011)
R-squared	0.051	0.100	0.175	0.096		0.236	0.406	0.105	0.490	0.193	0.000	0.091	0.142	0.515	0.594

Notes: Standard errors in *parentheses*. All coefficients are significant at the 1 % level, unless otherwise stated. ^ shows significance at 5 %; # shows lack of statistical significance (p-value > 0.10). xtabond is the Arellano-Bond estimator for dynamic panel data the fact that the relationship persists over time (columns 8–10), with productivity growth being positively related to past output growth irrespective of whether or not we control for the contemporaneous relation between the two aggregates.[13]

The Kaldorian formulation has been linked to demand-side processes, whereby market expansion drives technological adaptation and production efficiency (and thus growth – see Angeriz et al. 2008). Evidence of increasing returns, however, could be consistent also with a supply-side story, where productivity growth is driven by agglomeration, and thus knowledge and technology diffusion, rather than by demand-induced scale effects. The regressions in columns 11–13 in Table 2 try to test for this, by making productivity growth a function of the density of output (Ciccone and Hall 1996). As can be seen, output density, while not statistically significant in a simple OLS formulation, appears to be a significant contributor to productivity growth in better specified models. This holds especially true for deviations in the volume of output from its regional and year-specific average (col.13), with a 1 % deviation resulting in a 0.34 % rise in productivity growth.

Overall, our exploration of the cumulative growth hypothesis has provided evidence in support of all alternative views on the process. In the last two columns of Table 2 we attempt to examine the relative validity of each of these, by estimating a model that nests all three interpretations. Despite concerns about definitional correlation among the regressors[14] the performance of the models is very good and the R^2 increases significantly, suggesting that each of the three nested models adds a distinctive piece of information into the analysis. The variables corresponding to the two cumulative causation interpretations are both highly significant and have the correct signs. Of the two, the Kaldorian interpretation appears to produce stronger and more robust results, as the circular mechanism relating to the Myrdalian interpretation seems to die out quickly and even to reverse at t-3. In contrast, the agglomeration-economies interpretation returns a counter-intuitive result, with output density being negatively related to productivity growth. It thus appears that demand-driven cumulative growth, and to a lesser extent persistence, reflect more strongly the process of regional growth in our CEE sample. Controlling for such drivers, as well as for country and temporal fixed

---

[13] Although not in a formal way, this also addresses the question of the direction of causality between output growth and productivity growth (Angeriz et al. 2008), at least in its Granger-causality sense.

[14] On the right-hand-side we have GDP growth, GDP per hectare, and lags of the growth of GVA per worker. However, this does not seem to be a problem in the estimation. When we tested the reduced form of the nested model (as depicted in Eq. 3) the area variable (corresponding to the density argument) was highly insignificant and statistically very different from the estimate obtained for the regional GDP (in the notation of Eq. 3, $c_{11} \neq -c_3$ and $c_3 = 0$); the coefficients for output and lagged output where not statistically different in absolute terms (so that $c_{11} = -c_{12}$, consistent with the Kaldorian formulation); while the coefficient for employment growth was statistically different from the elasticities found for output (i.e., $c_{12} \neq c_2$ – for the Myrdalian formulation equality should hold). Thus, the process described by the Kaldorian interpretation appeared to carry more weight, similar to our findings in Table 2.

effects, wipes out completely any positive influence that agglomeration, and thus supply-side processes, may be exerting on productivity growth.

On balance, then, we see that both processes of convergence and of cumulative causation may be in operation. Although not directly deriving from these, the regional Kuznets Curve hypothesis may thus be particularly relevant in unveiling the conditions under which each of these processes dominates. In Table 3 we explore this, through a set of regressions that examine various specifications of Eq. 4 and its augmented version as given in Eq. 6. Starting from a simple OLS model (first column), it appears that the RKC hypothesis is not validated by the data. The obtained coefficients have the correct signs but they are not statistically significant. Moreover, they remain not significantly different from zero when we add controls for temporal and country-specific fixed effects to test a version of the conditional RKC hypothesis (results not shown). Although when adding region-specific fixed effects (col.2) the relationship becomes significant, the validity of this effect and its relevance to the RKC hypothesis is weakened by the fact that when we replace the national level of development with a time-trend (col.3 – see also footnote 11), the derived pattern is reversed and the obtained estimates suggest a cumulative (exponential) process of divergence over time.[15]

In col.4 we revert to the country-specific measure of national development (GDPpc) and amend the model of col.2 to include the past level of relative productivity as an additional regressor. This tends to stabilise the results and maintain the bell-shaped relationship (which is now consistent across econometric specifications and for both definitions of development – GDPpc and time-trend). Nevertheless, despite the fact that the direct effect of past productivity (in relative-to-national terms) turns out to be inversely related to current growth, for within-sample values the RKC process is not convergent: even for the highest national GDPpc value in our sample, poorer regions continue to grow at a slower pace (implying that national levels of development in CEE have not reached the threshold required for dispersion forces to become dominant). To examine whether this result captures simply a non-linearity in the neoclassical process, rather than a genuine RKC dynamic, in col.5 we test specifically the former, as specified in Eq. 6 (using past productivity measured in absolute terms). The results are radically different: for extremely (out-of-sample) low values of national development the model predicts divergence; while the speed of convergence first increases with the process of national development and then starts to decelerate, leading to a return to divergence for very high values of national development (but within the range of our sample values). Thus, based on the in-sample performance of the two models, it seems that the non-linear neoclassical convergence process provides a more accurate description of the data.

[15] The use of the time-trend as a measure of development imposes the assumption of a common development path across all CEECs. Clearly, our results suggest that relative regional growth and convergence/divergence dynamics have not been influenced by common developments in the CEE region but rather by country- and region-specific factors and characteristics.

Table 3 Regional growth in CEE: Kuznets curve and non-linear processes

Model	(1)	(2)	(3)	(4)	(5)	(6)
Past productivity (x) national income	0.050	0.114***	−0.027***	0.963***	−1.237***	0.222***
	(0.044)	(0.058)	(0.0085)	(0.293)	(0.093)	(0.045)
Past productivity (x) nat'l income squared	−0.016	−0.040***	0.001***	−0.173***	0.208***	−0.060***
	(0.012)	(0.015)	(0.0005)	(0.047)	(0.013)	(0.012)
Past productivity				−1.328***	1.212***	−0.220***
				(0.439)	(0.169)	(0.015)
Output growth						0.830***
						(0.033)
Fixed effects	–	Region	Year and region	Year and region	Year and region	Year and region
Nat'l development measured by	GDPpc	GDPpc	Time-trend	GDPpc	GDPpc	GDPpc
Past reg'l prod/vity	Relative	Relative	Relative	Relative	Absolute	Relative if
Measured in	Terms	Terms	Terms	Terms	Terms	Interacted
Constant	0.039***	0.029***	0.026	−0.106***	0.903***	0.482***
	(0.004)	(0.005)	(0.017)	(0.017)	(0.022)	(0.037)
R-squared	0.002	0.006	0.077	0.085	0.470	0.445

Notes: Standard errors in *parentheses*. *, ** and *** show significance at the 10 %, 5 % and 1 % levels, respectively. Models including fixed effects have been estimated by the within FE estimator

The last model of Table 3 (col.6), nests within it the simplest forms of all three competing interpretations of the regional growth process examined in this chapter (RKC: col.2 of Table 3; cumulative causation: col.7 of Table 2; and neoclassical convergence: row 6 of Table 1). As can be seen, in this model all three approaches are validated: evidence of a regional Kuznets Curve is clearly present (first two rows), as is evidence of neoclassical convergence (conditional on the other processes as well as on regional and temporal fixed effects) and of cumulative causation[16] (third and fourth rows, respectively). Still, despite being an amalgamation of the three approaches (and despite the obtained results being fully robust to the inclusion or exclusion of subsets of right-hand-side variables), this model does not produce a better fit than the augmented neoclassical model of col.5. Importantly, this model outperforms all RKC models presented in Table 3 (the latter explain persistently less than 10 % of the regional and temporal variation of growth rates in our sample) as well as the original convergence models presented in Table 1 (the obtained R^2 for the augmented neoclassical model is 1.5 times higher than that of the conditional neoclassical model in the sixth row of Table 1) and it is as strong

[16] Results reported correspond to the Kaldorian interpretation of cumulative causation. The results are qualitatively identical, however, when we replace the growth of output with a distributed lag structure of the dependent variable, to approximate the Myrdalian interpretation of circular causation.

as the best-performing cumulative causation models of Table 2. We discuss the implications of these findings in the concluding section.

5 Conclusions

There are two key points of departure for the analysis in this chapter. On the one hand, that owing to processes of transition and integration, regional disparities in the CEECs increased substantially over the last two decades and regional productivities and incomes became significantly polarised. On the other, that the tools used for the analysis of these developments, as elsewhere in the literature, have been rather limiting, owing to the dominance, as a key analytical tool, of the neoclassical convergence hypothesis. In this chapter we looked at three distinctive analytical traditions and tried to synthesise them in a model of regional growth that allows not only for convergence-divergence but also for cumulative and non-linear growth paths. Our econometric examination of the regional growth patterns in CEE sought to unveil the relevance and applicability of the competing explanations in our data.

Our results provide evidence in support of all three processes (convergence, cumulative growth and a regional Kuznets curve), but on balance they seem to favour a hybrid explanation, which sees regional growth as a non-linear process which is dependent on the level of national development and produces aggregate divergence and polarisation at later stages of development despite an overall tendency for convergence. This hybrid model is as suitable for describing the regional growth paths observed in CEE as are the three main competing processes combined. There is a very important implication stemming from this observation. Despite prima facie evidence to the contrary, the regional growth process in CEE is not obeying the rule of convergence, either monotonically (neoclassical) or through a bell-curve (regional Kuznets curve). Rather, convergence dynamics are significantly swaddled by processes of cumulative causation (Table 2), especially in early and, later-on, in more advanced stages of national development (Table 3), where the importance of demand-side market-size effects is increasingly heightened. As a result, through the process of national catch-up growth, and despite the tendency for mean-reversal, regional evolutions continue to be on the whole divergent, with a pattern of convergence at the middle- and lower-ends of the distribution and a slower tendency for club formation at the higher end, and thus overall an increasing trend of polarisation.

This finding raises important questions, for both policy and theory, that the present study is not designed to address: namely, will further national development in CEE reverse this process of polarisation and divergence; or is this process rather embedded in the institutions, economic-industrial structures and wider developmental models of these countries? As acknowledged in the Introduction, there is an increasing realisation in the literature that cultural and institutional factors, together with structural economic ones, play an important role in determining how regional growth evolutions, within their national and international context, are shaped by national and international processes of integration and structural change. Whereas it is these factors that ultimately determine absolute and relative regional economic performance, as is further elaborated in many of the chapters of this book, it appears that on aggregate, and in the particular case of the CEE regions in the period between the beginning of transition and the eruption of the global financial crisis, these factors have not contributed to a pattern of harmonious economic development across space. The structural dynamics underlying the sub-national growth process in the region over the last three decades have led to a widening of disparities in regional performance that may be indeed manifesting a more fundamental disparity in regional structures, institutions and potentials. Understanding how the latter constrain the attainment of regional convergence is a continuing challenge for both policy and academic enquiry.

References

Altomonte C, Resmini L (2002) Multinational enterprises as a catalyst for local industrial development. Case Poland Ital J Reg Sci 1(2):29–58

Angeriz A, McCombie J, Roberts M (2008) Estimates of returns to scale and spatial spillovers for EU regional manufacturing. Int Regional Sci Rev 31(62):62–87

Artelaris P, Kallioras D, Petrakos G (2010) Regional inequalities and convergence clubs in the EU new member states. East J Eur Stud 1(1):113–133

Banerjee B, Jarmuzek M (2010) Economic growth and regional disparities in the Slovak republic. Comp Econ Stud 52:379–403

Barrios S, Strobl E (2006) The dynamics of regional inequalities. FEDEA working paper No 2006-01. Foundation for the Study of Applied Economics, Spain

Barro RJ, Sala i Martin X (1991) Convergence across states and regions. Brookings papers on economic activity, no. 1

Bellini E, Ottaviano G, Pinelli D, Prarolo G (2012) Cultural diversity and economic performance: evidence from European regions [This Book]

Capello R, Fratesi U (2012) Globalization and endogenous regional growth [This Book]

Cibulskiene D, Butkus M (2007) The influence of cumulative causation process on regional divergence in Lithuania during 1995–2003. Jahrbuch f˙ur Regionalwissenschaft 27:59–87

Ciccone A, Hall R (1996) Productivity and the density of economic activity. Am Econ Rev 86 (1):54–70

Del Bo C, Florio M, Manzi G (2010) Regional infrastructure and convergence: growth implications in a spatial framework. Transit Stud Rev 17(3):475–493

Diliberto A, Usai S (2012) TFP convergence across European regions: a comparative spatial dynamics analysis [This Book]

Duranton G, Puga D (2004) Micro-foundations of urban increasing returns: theory. In: Henderson JV, Thisse J-F (eds) Handbook of regional and urban economics, vol 4. North Holland, Amsterdam, pp 2063–2117

Ezcurra R, Rapún M (2006) Regional disparities and national development revisited. Eur Urban Reg Stud 13(4):355–369

Ezcurra R, Pascual P, Rapún M (2007) The dynamics of regional disparities in Central and Eastern Europe during transition. Eur Plann Stud 15(10):1397–1421

Fischer M, Stirböck C (2006) Pan-European regional income growth and club convergence: insights from a spatial econometric perspective. Ann Reg Sci 40:693–721

Francois J, Rojas-Romagosa H (2008) Reassessing the relationship between inequality and development. CPB discussion paper No 107, Netherlands Bureau for Economic Policy Analysis

Henderson V, Shalizi Z, Venables A (2001) Geography and development. J Econ Geogr 1 (1):81–105

Herz B, Vogel L (2003) Regional convergence in Central and Eastern Europe: evidence from a decade of transition. In: Hausen C, Resinek M, Schürmann N, Stierle M (eds) Determinants of growth and business cycles: theory, empirical evidence and policy implications. Springer, Berlin, pp 155–179

Higgins M, Williamson J (2002) Explaining inequality the world round: cohort size, Kuznets curves, and openness. Southe Asian Stud 40(3):268–302

Iara A, Traistaru I (2003) Integration, regional specialization and growth differentials in EU accession countries: evidence from Hungary. In: Hausen C (ed) Determinants of growth and business cycles: theory, empirical evidence and policy implications. Verlag für Wissenschaft und Forschung, Berlin, pp 181–202

Islam N (2003) What have we learnt from the convergence debate? J Econ Surv 17(3):309–362

Kaldor N (1970) The case of regional policies. Scot J Polit Econ 17(3):337–348

Kallioras D (2010) Regional inequalities in the New European Union Member-States: is there a "population size" effect? Discussion paper No 16, Department of Planning and Regional Development, Univesrity of Thessaly

Kallioras D, Petrakos G (2010) Industrial growth, economic integration and structural change: evidence from the EU new member-states regions. Ann Reg Sci 45(3):667–680

Kuznets S (1955) Economic growth and income inequality. Am Econ Rev 45(1):1–28

Magrini S (1999) The evolution of income disparities among the regions of the European Union. Reg Sci Urban Econ 29:257–281

Martin R, Sunley P (1998) Slow convergence? The new endogenous growth theory and regional development. Econ Geogr 74(3):201–227

Myrdal G (1957) Economic theory and underdeveloped regions. Harper & Row, London

Nijkamp P, Siedschlag I (2008) European integration, regional growth and cohesion: an overview. International Journal of Public Policy 3:141–145

Niebuhr A, Schlitte F (2009) EU enlargement and convergence: does market access matter? Eastern Eur Econ 47(3):28–56

Paas T, Kuusk A, Schlitte F, Võrk A et al. (2007) Econometric analysis of income convergence in selected EU countries and their nuts 3 level regions, The University of Tartu Faculty of Economics and Business Administration Working paper, 60

Petrakos G (1996) The regional dimension of transition in Eastern and Central European Countries: an assessment. Eastern Eur Econ 34(5):5–38

Petrakos G, Artelaris P (2009) Regional convergence revisited: a WLS approach. Growth Change 40(2):314–331

Petrakos G, Economou D (2002) The spatial aspects of development in Southeastern Europe. Spatium 8:1–13

Petrakos G, Psycharis Y, Kallioras D (2005a) Regional inequalities in the EU new member-states: evolution and challenges. In: Bradley J, Petrakos G, Traistaru I (eds) The economics and policy of cohesion in an enlarged European union. Springer, New York, pp 45–64

Petrakos G, Rodríguez-Pose A, Rovolis A (2005b) Growth, integration, and regional disparities in the European Union. Environ Plann A 37(10):1837–1855

Persyn D, Algoed K (2009) Interregional redistribution, growth and convergence. Vives Discussion Paper series No. 4, Research Centre for Regional Economics, Katholieke Universiteit Leuven, Belgium

Resmini L (2003) Economic integration, industry location and frontiers economies in transition countries. Econ Syst 27:204–221

Resmini L (2007) Regional patterns of industry location in transition countries: does economic integration with the EU matter? Reg Stud 41:747–764

Rey S, Janikas M (2005) Regional convergence, inequality, and space. J Econ Geogr 5(2):155–176
Rodriguez-Pose A, Comptour F (2012) Do clusters generate greater innovation and growth? An analysis of European regions [This Book]
Römisch R (2003) Regional disparities within accession countries. In: Tumpel-Gugerell G, Mooslechner P (eds) Economic convergence and divergence in Europe: growth and regional development in an enlarged Europe. Edward Elgar, Cheltenham, pp 183–208
Rowthorn R (1975) What remains of Kaldor's law? Econ J 85:10–19
Storper M (1997) The regional world: territorial development in a global economy. Guilford Press, New York
Tondl G, Vuksic G (2003) What makes regions in Eastern Europe catching up? the role of foreign investment, human resources, and geography. ZEI discussion paper No B12, Bonn Center for European Integration Studies
Traistaru I, Nijkamp P, Resmini L (2003) The emerging economic geography in EU accession countries. Ashgate, London
Williamson J (1965) Regional inequality and the process of national development a description of the patterns. Econ Dev Cult Change 13(2):3–45

Some Practical Elements Associated with the Design of an Integrated and Territorial Place-Based Approach to EU Cohesion Policy

Philip McCann and Raquel Ortega-Argilés

Abstract This chapter discusses various important elements in the design of place based approach regional development strategies in the case of a reformed Cohesion Policy operating in the context of the Europe 2020 agenda. In particular, we focus on practical ways in which an integrated place-based approach to territorial and social cohesion can be exploited to tailor policy-design to the specific needs and potential of an EU region, in a manner which is consistent with the smart growth, sustainable growth and inclusive growth goals of Europe 2020. The chapter focuses specifically on a set of prerequisites which must be addressed in order to ensure that a place-based approach can be developed within the EU context. These prerequisites are a necessary condition for the correct ex ante design of the strategy; without an explicit consideration of these prerequisites it is very difficult to ensure that sectorally-structured institutions will adopt an integrated approach involving the mobilization of all actors and the extraction of local knowledge. The prerequisites discussed here are not sufficient to ensure the development of an integrated place-based approach, but they are a fundamental and necessary first step in the ex-ante design of a place-based strategy. The issues to be stressed are: firstly, a clear identification of the combined place-specific characteristics of each region and secondly, a clear identification of the territorial context. Finally, some regional

JEL Categories: H1, R11, R58

P. McCann
Department of Economic Geography, Faculty of Spatial Sciences, University of Groningen, 9700AV, Groningen, The Netherlands
e-mail: p.mccann@rug.nl

R. Ortega-Argilés
Department of Global Economics and Management, Faculty of Economics and Business, University of Groningen, 9700AV, Groningen, The Netherlands
e-mail: r.ortega.argiles@rug.nl

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_6,

cases which have recently successfully developed integrated place-based policy frameworks building on a smart specialisation way of thinking are also presented.

Keywords European Cohesion Policy • Smart specialisation

1 Introduction

The following sections discuss various important elements of a place based approach to the design of regional development strategies in a reformed Cohesion Policy operating in the context of the Europe 2020 agenda. In particular we focus on practical ways in which an integrated place-based approach to territorial and social cohesion can be exploited to tailor policy-design to the specific needs and potential of an EU region, in a manner which is consistent with the *smart* growth, *sustainable* growth and *inclusive* growth goals of Europe2020.

The aim of the chapter is not to discuss all aspects of a place-based approach, as these are discussed in detail elsewhere (Barca 2009). As such, this chapter does not discuss matters relating to policy implementation, conditionalities, monitoring and evaluation, or the deliberative process and public debate underpinning such a policy. Instead the chapter focuses specifically on a set of prerequisites which must be addressed in order to ensure that a place-based approach can be developed within the EU context. These prerequisites are a necessary for the correct ex ante design of the strategy; without an explicit consideration of these prerequisites it is very difficult to ensure that sectorally-structured institutions will adopt an integrated approach involving the mobilization of all actors and the extraction of local knowledge. The prerequisites discussed here are not sufficient to ensure the development of an integrated place-based approach, but they are a fundamental and necessary first step in the ex ante design of a place-based strategy.

The chapter is organised as follows. In the next section a background to some of the central elements to the place-based approach is provided. Each of these elements are linked directly to various practical steps which help in undertaking the type of analysis required of policy-makers for designing good place-based development policies. The issues to be stressed are: firstly, a clear identification of the combined place-specific characteristics of each region and secondly, a clear identification of the territorial context. In order to provide such clarity Sect. 3 discusses a straightforward and practical way of identifying a region's dominant place-specific characteristics using a simple diagrammatic device, a *Europe 2020 box*. We provide then four hypothetical examples of regions based on the *Europe 2020 box* diagram and we use these examples to show how a place-based logic helps directly to prioritise and tailor the appropriate types of integrated local development strategies required. We also extend the argument to incorporate rather more complex regions characterized by different types of places, and in addition we discuss examples of integrated development policies based on the linking of different places, rather than integrated policies focused at or within places. The relationship between policies focused at or within places and policies focused on linking places is very important, because it requires a different analytical approach as well as different policy priorities. Section 4 explains the appropriate analytical approaches for examining these different territorial (regional and urban) contexts and the ways in which different analytical approaches can be employed to help define policy priorities. Section 4 also uses examples to show how the territorial scale also influences the appropriate and meaningful outcome indicators which can be adopted. Section 5 briefly discusses some examples of regions which have successfully developed integrated place-based policy frameworks building on a smart specialisation way of thinking, and Sect. 6 outlines some brief conclusions.

2 The Place-Based Approach to Designing Development Policy

The place-based approach to regional development outlined here is based on the observation that places are different, and the ways in which places differ are very specific and particular. The place-based approach stands in marked contrast to the so-called 'space-neutral' approach of the World Bank (2009) which considers that regional place specifics and particularities are basically ad hoc background contextual features, largely akin to wallpaper, with no real explanatory or determining features. The place-based approach advocated by the OECD (2009a, b), the Barca (2009) report, the CAF (2010) report on Latin America, the Fifth Cohesion Report (European Commission 2010), and now also by the US government,[1] is based on the assumption that places really do matter for regional growth and development. Indeed, place characteristics not only define a place, but the geographic and institutional specifics and particularities of a place hold the clues as to the most appropriate pathways for development. An integrated place-based approach argues that dealing with the combined spatial and institutional challenges of a region in a systematic and integrated manner, rather than targeting individual sector-specific issues, is the most powerful form of development policy. As such, it is these very specifics and particularities of places on which the place-based approach builds.

The proposed reformed post-2013 Cohesion Policy regulations published in October 2011 recognise the importance of adopting an integrated approach to place-based development as being an essential aspect of the reforms, whereby all three dimensions of Europe 2020 are embodied in policy design and delivery as consistently as possible. For an integrated place-based approach to work it is necessarily to adopt an explicitly territorial approach to development policy, exactly as both the OECD and European Commission advocate. The advantage of this explicitly territorial approach is that it helps policy-makers to consider a range of possible place-based policy interventions which may well extend well beyond the control of existing sector-based institutional structures. Government departments are typically

[1] See: http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10–21.pdf http://www.whitehouse.gov/blog/2010/06/30/place-based-investments

organized on a sectoral logic, a logic which is not always naturally suited to the delivery of integrated policy solutions of a type needed for responding to the multi-dimensional Europe 2020 challenges facing society. Therefore, the shift towards an integrated and multi-level governance emphasis is a major priority within the EU, and this is necessary to ensure integrated place-based solutions are provided. Indeed, the need to overcome sector-based institutional "silos" in order to provide integrated and place-based development policy solutions is now also regarded as being of the highest importance in current US policy-thinking.[2]

A place-based approach to designing well-tailored and integrated development policy solutions requires first, an identification of the place-specific characteristics of a region. Second, it is necessary to consider the territorial features of a region, in terms of the interrelationships and interdependencies within or between places. This is important because the combination of the place-specific characteristics and the territorial aspects heavily influences the choice of the appropriate analytical approach to be adopted and the institutional challenges to be addressed. These in turn influence the choice of likely policy priorities, the realistic alternative policy goals on offer, and the appropriate outcome indicators to be employed.

In order to help achieve the first stage of the place-based approach, Sect. 3 below provides a simple diagrammatic method to demonstrate how an integrated approach can be a powerful tool for identifying the principal development challenges faced by an EU region, within the context of Europe 2020. The diagrammatic method employed here involves the construction of a box containing individual cells, each of which is different and is defined according to the innovation-environment-demographic characteristics of a place. This simple device allows policy-makers to quickly and easily identify the specific combined place-based features of their particular region. Identifying these features is an important first stage in the policy-design process because it is the particular combination of these place-specific features evident in each particular region which determines the principal development challenges faced by each region. These combined features will also point to the likely and realistic policy options available to the region. In order to demonstrate this we discuss four hypothetical example regions, showing how both the development challenges of smart, sustainable and inclusive growth and the policy responses are all interrelated.

Once the major place-based characteristics of a region have been identified, the place-based development policy logic implies that the territorial dimensions of the regions also need to be considered. This is important because managing authorities need to decide which parts of their region are to be prioritised for development assistance, exactly why this is the case in terms of market failures, bottlenecks etc., and which types and levels of assistance are to be provided for these areas. Obviously, the identification of general thematic policy priorities is driven by broad social wellbeing debates and political discourses. However, embedded within these broad thematic priorities are the specific place-based policy priorities, and these are seen in part to both depend on, and are also be dependent on, the analytical approach

[2] See footnote 1.

adopted. As will be described in detail in Sect. 4, for our purposes here, the different place-based analytical approaches on offer can be defined variously as an *urban* approach, a *regional* approach, an *inter*-regional approach, and a *super*-regional approach. As we will see in Sect. 4, the choice of the appropriate place-based analytical approach to be adopted will depend greatly on the territorial context, where territorial context is understood in terms of the spatial distribution of activities and people, the spatial aspects of the institutional and governance architectures, and the spatial patterns of development challenges to be faced. The appropriate analytical approach will help to best prioritise different competing policy objectives and also to consider how the choice of outcome indicators and conditionalities employed both relate to the development challenges being faced and the policy priorities made.

3 Identifying the Place-Based Characteristics of a Region

The three dimensions of the *Europe2020* agenda, namely *smart* growth, *sustainable* growth and *inclusive* growth, can be integrated into the design of Cohesion Policy at a regional and local level in a straightforward manner which allows for an integrated approach to regional policy design and delivery.

The way this is achieved us by adopting the *place-based* approach which requires policy-makers and managing authorities to *self-evaluate* the dominant characteristics of their own regions which are being targeted by Cohesion Policy. Identifying the major features of each region across all three Europe2020 dimensions will allow policy-makers to best design *integrated* development strategies, in which operational programmes and projects explicitly foster *territorial cohesion* and *social cohesion*.

At this point it is necessary to make clear that when we use the term ‘region’ we are assuming an ‘administrative region’. Moreover, at this stage we also initially assume that an administrative region is largely internally homogenous and that an administrative region can indeed be considered to be a functional region in nature. In other words a rural region is assumed to be primarily rural throughout the administrative regional area and an urban region is assumed to be primarily urban throughout the administrative region. We will discuss more complex and diverse regions later on in Sects. 3.2 and 3.3 and in Sect. 4.

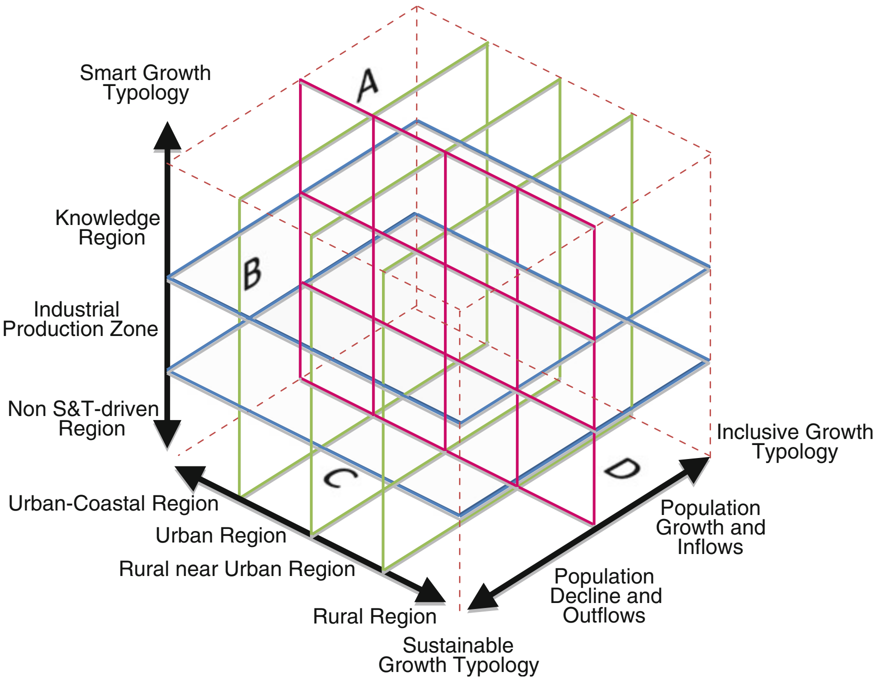
In order to help policy-makers and managing authorities to identify the dominant characteristics of their own regions we employ a three-dimensional box diagram, within which any *individual administrative* region participating in Cohesion Policy programmes or projects can be positioned or situated. The sides of the box reflect the three different dimensions of Europe2020. Each side of the box provides a typology which most concisely captures the major features associated with each of the individual Europe2020 challenges.

For the *smart growth typology*, the most concise framework is provided by the OECD (2011a) regional innovation typology in which regions are grouped into three types, namely *knowledge regions*, *industrial production zones*, and *non-Science and Technology-driven regions*. These three categories reflect the major observed differences in terms of the relationships between knowledge, innovation and regional characteristics. All regions can be broadly categorised into one of these groupings in terms of the role played by knowledge in fostering their local innovation processes.

For the *sustainable growth typology* the classification scheme which most concisely captures the different combinations of environmental and energy challenges faced by different types of regions is based on a consideration of the relationship between the built environment and the natural environment. At its most fundamental level, this gives us four types of regions, namely regions which in nature are primarily *rural* regions, *rural near urban* regions, *urban* regions, and *urban-coastal* regions. This sustainable growth classification scheme closely resembles the OECD (2009a, c, 2011b) regional typology based on the dominant built-environment-natural environment features which uses three types of regions, namely *predominantly urban regions*, *predominantly intermediate regions*, and *predominantly rural regions*. Our definition of a primarily *urban* region obviously corresponds to the OECD definition of a predominantly urban region. Our definition of a *rural* region corresponds to the OECD definition of a predominantly rural region, and here we include mountain regions within this category. Our definition here of a *rural near urban* region corresponds to the OECD definition of *predominantly intermediate* region. We also include a separate a specific category which is an *urban and coastal* region, because of the specific combination of ecological and environmental challenges faced by these regions allied with the particular trade and commercial possibilities of these regions to respond to these challenges.

For the *inclusive growth typology*, the classification scheme which most concisely captures the very different social inclusion issues faced by regions is that which is also adopted by the ESPON (2010) DEMIFER project. This has two broad types of regions, namely regions facing *population decline and population outflows* and regions facing *population growth and population inflows*. Migration is a highly selective phenomenon and mobility is highly correlated with skills and income. In today's society which is becoming both increasingly mobile and also is ageing, differences in migration patterns are engendering major differences in population change in different places, and these have significant impacts on both innovation and environmental issues.

The box diagram below can be used as a device to help policy-makers identify the *combined* place characteristics of their own regions. Given that regions exhibit different features according to these different dimensions of Europe 2020 it is therefore possible to consider each region as reflecting a particular combined set of such place-characteristics. In the box diagram below, each individual axis represents one of the three Europe2020 agenda dimensions. The combination of the *smart* growth, *sustainable* growth and *inclusive* growth typologies allows for 24 possible tripartite types of place characteristics, each of which is reflected by a different cell in the three-dimensional box of regions. The principal features of individual each region can therefore be captured in terms of one of 24 different tripartite sets of place-specific characteristics. Each set of place-specific characteristics is associated with a particular tripartite combination of the principal innovation-environment-demographic characteristics of that region. Using this simple diagrammatic device it is quite straightforward to situate each region within the Europe 2020 box, and therefore for policy-makers to identify the major development challenges they face in their region.



Europe 2020 dimensions: integrated regional typologies

3.1 *Hypothetical Examples of Different Place-Specific Characteristics*

The self-identification approach is an important first stage in the design of a regional policy strategy. Identifying the combined Europe2020 place-specific features of a region promotes awareness of the challenges to be faced and the likely policy options. This awareness allows for alternative *integrated* and *place-based* development policies to be designed, considered and weighed against each other, in the context of a broad understanding of the challenges ahead. The self-identification process also immediately narrows down the possible sets of Cohesion Policy thematic priorities and interventions which can appropriately be employed in each particular case. The combination of a heightened awareness of the regional place-specific features and a narrowing down of likely policy options is important for policy design. This is because integrated local development policies should be tailored to best respond to the specific local combination of smart, sustainable and inclusive growth challenges faced by each particular region.

Here we provide four hypothetical examples of regions, regions **A**, **B**, **C** and **D**, depicted in the diagram above as being situated in different cells of the Europe2020 box. For each region we provide a hypothetical description of how the different aspects of Europe2020 agenda can help inform the Cohesion Policy priorities chosen and the resulting design and implementation of the programmes and projects.

Hypothetical Example 1. Region **A** is a knowledge region, is primarily urban in nature but is also situated on the coast, and faces population growth and population inflows.

For a region such as Region **A**, in *smart growth* terms the major development opportunities will centre on fostering more local innovation and knowledge creation. The actual interventions chosen will depend on the context. However, such a smart growth strategy will further encourage in-migration and local population growth, and this will increase congestion, thereby endangering the capacity of the local economy to maintain its long-term competiveness. In order to offset such adverse congestion effects and to enable *sustainable growth*, possible interventions may involve major infrastructure initiatives for energy provision and wastewater management. In addition, however, population expansion and infrastructure provision in this particular urban-coastal context may also endanger the fragile coastal ecosystems, so these interventions must be tailored to respond to these challenges. Meanwhile, these various natural and physical resource constraints also have *inclusive growth* implications. Land price rises associated with population inflows are likely to lead to increasing spatial and neighbourhood segregation between different income and skills groups within the urban context. Such segregation militates against both *social cohesion* and *territorial cohesion*, and leads to increasingly disconnected and isolated neighbourhoods within the urban-coastal region. The innovation and infrastructure interventions must therefore be integrated with interventions promoting occupational mobility and spatial mobility for all groups. Taken together, the combined features of region **A** therefore imply that while promoting local innovation will be a key driver of Cohesion Policy, such a strategy must also be underpinned by a careful integration of land-use management, infrastructure provision, and public transportation systems, as well as labour-training systems talilored to local needs. The actual range of interventions chosen will depend on thematic priorities adopted and the place-based specifics of the context.

Summary

Major features: Knowledge region; population growth and in-migration: urban + coastal

Major challenges: transport and land-use congestion, social and territorial segregation; environmental damage including marine ecosystem

Major opportunities and place-based policy priorities: multi-sectoral knowledge-enhancing projects; integrated infrastructure, housing and public transport provision

Hypothetical Example 2. Region **B** is an industrial production region, an urban-coastal area, and a region which faces population decline and population outflows.

For a region such as region **B**, in terms of *smart growth*, the region already has major industrial assets on which can be built a smart specialisation strategy based on helping the region's most embedded firms and sectors to diversify technologically into other technologically-related fields of activities. As is always the case, the actual interventions chosen will depend on how the *smart specialization* strategy is tailored to the local context, but the focus should be on both high and medium technology sectors. However, the scale of the industry, the degree of longevity of the actors, and the breadth and depth of the linkages between the firms and other actors in the region will be essential features to be considered in terms of the firms, sectors and technological fields to be prioritised. In terms of *social cohesion* and *territorial cohesion* such prioritisation will also need to be related to the diversity of people whose livelihoods and employment opportunities are related to the targeted sectors firms and technological fields. This *inclusive growth* focus is particularly important in a region facing population decline and outflows, because the growing skills shortages and skills mismatches which need to be overcome by the training and retraining of people whose priority it is to remain living in the area. The changing labour demands of local firms due to their technological diversification strategies must be incorporated into the design of the local skills and retraining programmes. This is necessary in order to ensure that the benefits of development are *inclusive* of all those in the local community and reach the widest range of income groups as possible. Moreover, population outflows and land dereliction also imply that these skills and retraining policies must be accompanied by policies promoting the transformation of land-uses into appropriate spaces for the reconfiguring of work and the fostering of local business investment. This is essential in order to make the policies *sustainable*. Otherwise the local real estate markets will decline and this will have adverse effects on business collateral gearing ratios, thereby reducing the local credit availability for SMEs. The appropriate redesign of public and commercial urban spaces will also involve the reclamation, decontamination and transformation of both marine and urban spaces. Multi-sectoral and multi-level coordination is therefore essential here in order that these policies deliver growth which is *sustainable*, in terms of local investment, local skills, and the built environment.

Summary

Major features: Industrial production zone; population decline and out-migration: urban + coastal

Major challenges: declining transport and land-use usage, dereliction, non-operative real estate markets, skills outflows, declining credit availability, widespread reductions in social and territorial cohesion; environmental damage including marine ecosystem

Major opportunities and place-based policy priorities: smart specialisation policies targeted both at high and medium technology sectors and based on specialized technological diversification strategies in major embedded occupational and technological classes; local labour skills-enhancing programmes in related technologies; integrated land use reclamation and conversion programmes.

Hypothetical Example 3. Region **C** is a non-S&T-driven lagging region, primarily urban region, and a region which faces population decline and population outflows.

For a region such as Region **C**, in *smart growth* terms the major development opportunities are likely to centre on fostering more adoption and adaptation of new technologies within the existing industrial system. The actual interventions chosen will depend on how the *smart specialization* strategy is tailored to the local context. However, a lack of any major local science and technology base means that these interventions should focus primarily on medium technology sectors, with the aim of the strategy being to foster innovation in those key capabilities in which the region has a realistic growth potential. This is designed so as to help bolster the region's long-term resilience and to enhance its competiveness. In terms of *sustainable growth*, in order to encourage the smart specialization strategy, Cohesion Policy priorities may centre on interventions regarding land-reclamation and land-conversion from derelict spaces. This will be undertaken in order to enhance the local environment for the provision of commercial business space and to make the urban region more attractive for inward private sector investment. At the same time, in terms of *inclusive growth*, those interventions aimed at fostering the transfer of 'brownfield' to 'greenfield' sites may need to be allied with transportation interventions, so as to facilitate an improved reconfiguration of mobility patterns promoting more integrated residential spaces. Once again, the actual range of interventions chosen will depend on thematic priorities adopted and the place-based specifics of the context.

Summary

Major features: Non S&T-driven region; population decline and out-migration: primarily urban area

Major challenges: declining transport and land-use usage, dereliction, non-operative real estate markets, skills outflows, declining credit availability, widespread reductions in social and territorial cohesion; environmental damage including marine ecosystem

Major opportunities and place-based policy priorities: smart specialisation policies based focused on medium technology sectors; local labour skills-enhancing programmes in related technologies; integrated land use reclamation and conversion programmes.

Hypothetical Example 4. Region **D** is a rural area, a non-S&T-driven region, which faces population growth and population inflows.

Region **D** is a rural region, typically of low population and low population density, but which now faces population inflows. The local population growth is due to changes in consumer and household preferences in which high skills and income groups increasingly choose to live in high amenity environments. Most of the recent in-migrants have relocated from central core regions, but many still remain in employment activities and occupations which involve interacting regularly with customers in the major cities. In terms of the *smart growth* agenda a key priority will be to enhance the local communications infrastructure via technologies such as broadband and advanced telecommunications systems, as well as upgraded road and rail systems. Increasing local tourism opportunities also imply that cultural and heritage assets must be preserved and enhanced, and protected from land conversions pressures. The increasing demands on local resources provide opportunities to develop and enhance renewable energy sources so as to ensure that local population and income growth is a *sustainable growth* process. At the same time, in order to avoid major energy rebound effects, local land conversion from agricultural to urban uses must be undertaken via coordinated and systematic planning regimes. This is essential in order to allow for the development of efficient local public transportation systems and the avoidance of local heat islands. In addition, population growth combined with resource constraints can lead to rising local living costs, the affects of which are most adversely felt by the lowest income groups. In order to avoid local problems of increasing social and territorial segregation and to make growth as *inclusive* as possible it is necessary to avoid local development taking place in an uncoordinated manner. Skills and training programmes may focus on tourism and natural environmental activities. The design of integrated land use planning policies which incorporate the provision of public services and public transport and communications systems will help to ensure that the local growth is *inclusive*, spreading the benefits of development as widely as possible by providing the broadest access to the new employment opportunities and services to all local people.

Summary

Major features: Non S&T-driven region; population growth and in-migration: rural area

Major challenges: pressure on local resources and land use; social and territorial segregation; economic and geographic isolation

Major opportunities and place-based policy priorities: smart innovation growth policies based on communications infrastructure; preservation and upgrading of heritage and cultural assets; skills enhancement policies focused on tourism and natural environmental arenas; renewable energy policies; social and territorial cohesion focused on integrated land use development and public transport planning.

3.2 Diverse Administrative Regions

Across Europe as a whole and also within each Member State, the number of regions in each cell of the Europe2020 box will differ, depending on the economic, environmental and social geography of each country. For example, more regions in Spain are coastal than in Hungary; more regions are primarily urban in UK than in Bulgaria; more regions are non S + T-driven regions in Poland than in The Netherlands, more regions are industrial production zones in Germany than in Malta, and more regions are knowledge regions in Sweden than in Greece. However, for our purposes the numerical differences are themselves not important, other than to point out that the individual cells of the Europe 2020 box are not filled with equal numbers of regions. Instead, what is important for policy-makers is to identify the *combined* place-specific features of each region, as a first essential step in the design of integrated place-based development policies.

As we have seen above, using this simple diagrammatic device it is quite straightforward to situate each locality within the Europe 2020 box. However, the argument becomes slightly more complicated when we try to situate diverse administrative regions in the Europe 2020 box. The reason is that the ease of doing this depends on the geographical size of the administrative region. Here, the distinction between place and administrative region is useful, because the jurisdictional definitions of administrative regions do not always strictly follow the dominant innovation-environmental-demographic features of different types of places. For example, administrative regions which cover only small geographical areas will generally exhibit just one dominant tripartite set of place characteristics, as represented by a single cell in the box diagram. On the other hand, geographically large administrative regions such as Andalucia, Bavaria or Scotland, will often contain several different types of places, different types of sub-regions, within their jurisdictional boundaries. Each of these sub-regions can be represented by a different cell in the box diagram reflecting the different tripartite combinations of their place characteristics. As such, regional managing authorities who are developing place-based policies within the setting of a large administrative region which contains very different types of places must first decompose their administrative regions analytically into the major different types of places, using the tripartite classification scheme depicted in the box diagram. This ought not to be too difficult in that managing authorities and policy-makers should easily be able to identify the dominant place-based characteristics of their own region as a whole or the different parts of their own region.

3.3 Linking Regions via Integrated Strategies

In the Europe 2020 box diagram above, the individual cells represent the different combined place-specific features of individual regions. In the four hypothetical examples described above the design of the place-based policy strategies focuses on linking each of the Europe 2020 dimensions of development in an integrated manner which is consistent with the place-specific features and challenges of the region. The manner in which the *concept of integration* is understood here is in terms of multi-thematic and multi-dimensional policies which are *linked* and also targeted *at* a place *within* an administrative region. The policy priority here is the vertical coordination of different levels of governance, whereby the higher levels of governance also coordinate policy interventions with lower more local levels of governance.

However, it may be the case that integration also can be understood in terms of *linking between* places or sub-regions within an administrative region. This would imply conceptually breaking down an administrative region into various sub-regions, each with different place-specific characteristics, and then considering how each of the sub-regions relates to each other. Such an understanding would imply somewhat different place-based policies.

Similarly, it may be the case that the appropriate policy is to foster linkages between different administrative regions, some of which may even traverse international borders. This might be appropriate in situations where groups of adjacent administrative regions share sufficiently large common innovation-environmental-demographic characteristics, that policy coordination offers the greatest possibilities for the concentration and targeting of resources. This might also imply coordination across international borders, as well as between administrative regions within an individual Member State. The logic of this approach, irrespective of whether it takes place between administrative regions within a Member State or between administrative regions traversing an international border, is that the linking of regions and the horizontal coordination of policy is the priority.

For example, a possible alternative place-based *smart growth* policy priority for region *B* may be to increase its technological links with region *A*, in a manner which fosters increasing regional exports from region *B* to region *A*. Given that both regions are urban-coastal regions, there may be possibilities for the development of joint environmental and energy platforms for fostering *sustainable* development based on shared learning and shared investment priorities, particularly if both regions share a coastline. On the other hand, the *inclusive* growth strategies designed to foster social inclusion may still operate primarily at the local level, rather than at an inter-regional level, as in the descriptions above.

Similarly, a possible alternative place-based *smart growth* policy priority for region *C* may be a smart specialization strategy aimed specifically at increasing its technological links with region *B*, in a manner which fosters increasing regional exports from region *C* to region *B*, and even to other regions such as region *A*. If both regions share a common freshwater system, then joint environmental strategies may be developed fostering *sustainable* growth and based on shared learning and shared resource needs. In addition, if sufficient technical complementarities exist between the technologies and skills of regions *B* and *C*, it may be possible to develop joint inter-regional integrated skills and re-training policies, the logic of which is linked directly to the smart specialisation strategy. Such an *inclusive* growth approach will represent an interregional inclusive growth strategy directly building on the smart specialization priorities.

Finally, in order to allow for *sustainable* local growth and to avoid the over-exploitation of scare local resources, region *D* may choose to develop a *smart* growth strategy aimed at enhancing knowledge linkages with region *A*. Such an approach might be built around fostering information and communications technologies (ICTs) linkages with regions *A* and *B*, in order to ensure high quality local employment opportunities based on knowledge exchanges with potential key input and output markets. Local skills training once again could be built around these new technology platforms in order to ensure inclusive local growth.

These three examples serve to highlight the fact that the territorial aspects of local development policy may differ depending on the perceived priorities. In some cases, such as the four hypothetical examples sketched out in Sect. 3.1, the emphasis of the policy-design logic was on integrated strategies *at* a place within an administrative region, whereas in the three examples sketched out here, the emphasis of the integrated policy-design logic is in terms of fostering the linkages *between* places within an administrative region, or between administrative regions. In essence these different approaches adopt different *territorial* constructs, reflecting not just different cartographical arrangements, but more fundamentally an understanding that different economic mechanisms can be exploited and influenced in different ways, depending on the nature of the economic linkages between regions. The next section explains how these different territorial approaches influence the policy-design process, the choice of policy priorities, and the different institutional and governance challenges to be addressed.

4 Analytical Approaches to Territorial Cohesion and Policy Design

Administrative regions can be defined in various multi-dimensional ways, as we have just seen with the box diagram above. However, the territorial dimension of regions explicitly influences both the analysis of regions and also the design of Cohesion Policies, because the logic of the policy priorities is altered depending on the territorial context. For example, we have already seen that large administrative regions may need to be decomposed into different sub-regions with varying place-specific combinations, according to the box diagram logic. However, this also implies that the appropriate policy logic and the results/outcome indicators employed, may differ according to the relationship between the innovation-environmental-demographic place-specific characteristics and the institutional architecture of contiguous regions. The reason is that the policy logic depends on the spatial dimensions of both the challenges to be faced and also the likely solutions.

The spatial dimensions of either the challenges to be overcome, or their likely solutions, may differ from being highly localised issues to widely distributed issues. Similarly, it may be that the spatial dimensions of the challenges faced differ from the spatial dimensions of their likely solutions. For example, it may be the case that the major obstacles to development are highly localised, such as traffic congestion bottlenecks or specific institutional bottlenecks, whereas the adverse impacts of these localised bottlenecks may be very geographically widespread, adversely affecting the network coordination abilities of stakeholders in many different places. Possible solutions may be localised decongestion policies or alternatively a more widely distributed redesign of the network system may be preferable. Alternatively, it may be that the social and territorial challenges are very localised, such as social exclusion or spatial segregation, and the most effective solutions might be perceived to be localised integration and skills policies. However, what is important here is that for a genuine territorial approach to be exploited in a place-based Cohesion Policy design process, the explicitly spatial dimensions of all potential policy challenges and solutions must be considered and weighed against each other.

The place-based approach is based on the principle that the promotion of *territorial cohesion* and *social cohesion* are best ensured by adopting regional development policies which are integrated and multi-sectoral in terms of themes, multi-level in terms of governance, and tailored to the context. However, achieving these policy-design features in different types of regions requires us to think slightly differently in different contexts. This will automatically raise questions regarding the institutional issues to be addressed, because as we have seen, many solutions may involve coordination and cooperation across not only different levels of governance, but also across different jurisdictional boundaries. Analysing regional development challenges from a territorial perspective ensures that multi-level governance and institutional coordination issues become central elements in the policy-design process, both in terms of the challenges faced and the possible solutions on offer.

In each case the exact way that the territorial dimension influences the priorities of the policy-design logic will obviously differ according to the particular territorial and multi-dimensional context. However, the simplest way of understanding this is to split the territorial policy-design logic into four broad typologies, namely an *urban* approach, a *regional* approach, an *inter-regional* approach, and *super-territorial* approach.

1. An *urban* approach in focuses on *intra-urban* and *inter-neighbourhood* issues and is the most geographically concentrated of the four territorial typologies. This represents the micro-territorial scale, and the important elements here are the interdependencies and interrelationships within a place.
2. A *regional* approach operates at a broader territorial scale than an urban approach and the regional approach emphasises *intra-regional*, *inter-urban*, *rural*, *urban–rural*, or *polycentricity* issues. This represents the meso-territorial scale, and the important elements here are the interdependencies and interrelationships between places and also between places and their hinterlands.
3. An *inter-regional* territorial approach operates at a very broad territorial scale and focuses on *intra-national* or *inter-national* (INTERREG) issues. This represents the macro-territorial scale. As with a regional approach, the important elements here are the interdependencies and interrelationships between places and also between places and their hinterlands.
4. *Super-territorial* approaches, such as the Baltic Sea and Danube Region macro-regional strategies, operate at the broadest international scale and involve a highly integrated, multi-sectoral and multi-thematic approach applied across a very wide-ranging multi-territorial spatial context. This represents the super-territorial scale, and here the important elements are the interdependencies and interrelationships both between places and also within places.

These differences in territorial scale also influence the policy priorities and the policy design logic, which can be summarised as:

- An *urban* approach to development policy in essence represents a *highly-diversified and multi-sector* approach applied at a *particular location*.
- A *regional* approach to development policy is represents a *more narrowly-defined* and *less diversified multi-sector* approach applied across a *multi-locational, but relatively compact, spatial structure*.
- An *interregional* approach to development policy is typified by a *narrowly-defined and specialised approach* applied across a *broad multi-locational spatial structure*.
- A *super-territorial* approach to development policy represents a *highly-diversified and multi-sector* approach applied across *many different locations*.

In order to compare these different approaches, a useful analogy is to consider a *node-and-network* structure. An urban approach which emphasises issues at or within a particular place can be considered as focusing primarily only on specific nodes, treating these very localized issues as being paramount. In contrast, a regional approach which emphasises the interactions between places can be considered as treating the links between the nodes as being at least as important as what happens within the nodes. With this analogy in mind it is now possible to examine the various dominant features of each of these four different territorial analytical approaches by considering (1) a synthesis of the major characteristics (2) the place-based context (3) the network and agglomerations issues (4) the principal aspects of social cohesion (5) the principal aspects of territorial cohesion (6) the mobility issues (7) the institutional capabilities (8) the transport issues (9) the governance issues, and (10) the major issues to be dealt with via conditionalities.

We can consider the dominant features of the four major territorial analytical approaches in terms of each of these ten issues.

4.1 Urban Approach

For the urban approach, the dominant features are: (1) *Synthesis of Major Characteristics:* spatial concentration of activity and people, multi-thematic approach, multi-sector integration (2) *Place-based context*: The territorial structure underpinning the development policy design-logic is a mono-spatial context. (3) *Networks and agglomeration*: The dominant logic of the system relates to spillovers and interdependencies within a place. In terms of network effects, the major features are the concentration of activity within the individual node and interdependencies operating within the node. The policy aims at the fostering of positive local agglomeration effects. (4) *Social Cohesion*: Social Cohesion in this urban context is understood in terms of intra-urban inter-neighbourhood effects, between different income groups, between different ethnic groups, and between different social groups, within the same place. Social exclusion is observed in terms of spatial segregation. Territorial Cohesion and Social Cohesion cannot be separated at the urban level. (5) *Territorial Cohesion*: In an environment characterised by multiple local interdependencies and spillovers territorial cohesion is essentially the same as social cohesion. (6) *Concept of Mobility*: Mobility issues are primarily related to occupational mobility rather than geographical mobility, in order to avoid segregation and social exclusion due to unemployment or ongoing working poverty. (7) *Institutional Capabilities*: A holistic approach to service-delivery needs to be fostered within intra-urban institutions, so as to allow for a highly diversified multi-sectoral policies to be implemented within and between intra-metropolitan jurisdictions. (8) *Transport Priorities*: The promotion of sustainable public transportation facilities for urban commuting, including buses, light rail, trams, is the priority for ensure sustainable intra-urban mobility. (9) *Governance Priorities*: The multi-level governance priority centres on the fostering of cross-sectoral coordination within jurisdictions which is vertically integrated and highly cross-thematic in nature. (10) *Conditionalities*: Conditionalities focus on issues relating to information provision, outcome indicators, and evaluation processes necessary for ensuring the multi-level institutional reforms required for delivering cross-sectoral vertically-integrated policies.

4.2 Regional Approach

For the regional approach, the dominant features are: (1) *Synthesis of Major Characteristics* Spatial diversity of activity and people, thematic concentration of approach, narrower cross-sector objectives (2) *Place-based context*: The territorial structure underpinning the development policy design-logic is a multi-spatial context. (3). *Networks and agglomeration*: The dominant logic of the system relates to interactions and interdependencies *between* places in a system. In terms of network effects, the major features are activities distributed across nodes and the resulting interdependencies operating between the nodes. The policy aims at the fostering of positive transmission effects and the enhancement of connectivity between places. (4) *Social Cohesion*: The social cohesion priority aims at maximising the local development potential of all income groups in non-core or lagging regions and building on the untapped potential of all places in the national economy. (5) *Territorial Cohesion*: The major emphasis for territorial cohesion is on the fostering of horizontal governance coordination between jurisdictions. In this territorial context, territorial cohesion and social cohesion are distinct and separate. Although they overlap in many aspects, the actual relationship between them depends on the nature of the programme. (6) *Concept of Mobility*: The mobility priorities in this territorial construct are primarily related to fostering the geographical mobility of people, goods and services between places. (7) *Institutional Capabilities*: A holistic approach to cross-institutional service-delivery needs to be fostered within adjacent intra-regional jurisdictions, so as to allow for narrowly focused multi-sector policies to be delivered across and between intra-regional jurisdictions. (8) *Concept of Transport*: The transport priorities focus mainly on removing network bottlenecks so as to enhance the efficient mobility of people, goods and services between places, via high speed trains or key road infrastructure. (9) *Governance*: The multi-level governance priority centres on the fostering of horizontal and coordination between sub-regional areas within administrative regions in order to provide for cross-thematic policies which are more narrowly defined than in the urban context but applied over larger territorial regimes. (10) *Conditionalities*: Conditionalities focus on issues relating to information provision, outcome indicators, and evaluation processes necessary for ensuring the multi-level institutional reforms required for delivering multi-sector policies in a coordinated manner in a cross-jurisdictional territorial context.

This regional-type of approach which emphasises interactions and interdependencies between places can also be applied over much larger territorial contexts. For example, the major features of the *inter-regional* approach are also largely the same as the major features of the regional approach. The major difference between the *inter*-regional approach and the regional approach is that while the analytical framework is largely the same, in the case of the inter-regional approach it is applied more narrowly over a much larger territorial structure traversing jurisdictional boundaries of administrative regions. As such, cross-jurisdictional coordination is an essential element of this approach. These jurisdictional boundaries may even traverse national borders.

Finally, the major features of the super-territorial approach are that multi-sector and multi-thematic integrated principles of the urban approach are applied to an extremely broad multi-territorial spatial structure via a multi-level governance system involving both vertical and horizontal coordination across national borders.

4.3 Territorial Cohesion, Policy Interventions and the Choice of Appropriate Results/Outcome Indicators

In this section we adopt the explicitly territorial pattern of thinking described in Sect. 4 which distinguishes an urban approach from a regional approach, and which is based on the difference between interdependencies and interrelationship which operate primarily at a locality from those which operate primarily between localities. This explicitly territorial approach is intended to help policy-makers consider alternative scenarios for policy interventions from quite different spatial perspectives, even for the same development challenges, as well as for different development challenges.

One element of a place-based regional policy is the issue of outcome indicators. As was mentioned at the beginning of this note it is not the intention of this note to discuss the issues relating evaluation, monitoring, conditionalities, or the use of outcome indicators, all of which are key components of a place-based approach, as these are discussed in detail elsewhere (Barca 2009; Barca and McCann 2011). However, following this territorial pattern of thinking it also becomes clear that the likely outcomes of certain policy interventions are more appropriately understood in an urban context while others are better understood at a regional context. This is also implies that the outcome or results indicators chosen and the interpretation of these indicators should also be seen in a territorial light.

By way of examples, Table 1 provides some examples of Cohesion Policy thematic priorities and interventions in the context of Europe 2020, and these examples are based on the Barca and McCann (2011) listings of results/outcome indicators. In each case the table indicates the likely territorial scale over which interventions implemented under these priorities are likely to operate. As such, this determines the territorial scale over which the intended outcomes or results are likely to be evident. An awareness of this also determines the territorial scale over which the appropriate outcome or results indicators chosen are to be interpreted.

The territorial implications of the use and design of these results/outcomes indicators becomes evident if we consider a couple of examples. For example, if

Table 1 Examples of results/outcome indicators and the appropriate territorial scales

EU priority	Europe 2020 objective	Cohesion policy thematic priorities	Territorial scale of policy intervention
Smart growth	Improving the conditions for innovation, research and development	Strengthening research and technological development	Regional
Smart growth	Improving the conditions for innovation, research and development	Promoting innovation and smart specialization	Urban or regional
Smart growth	Improving the conditions for innovation, research and development	Enhancing accessibility to and use and quality of information and communication technologies	Urban or regional
Smart growth	Improving the conditions for innovation, research and development	Removing obstacles to the growth of SMEs	Regional
Smart growth	Improving education levels	Improving the quality and performance of education and training systems at all levels and increasing participation in tertiary or equivalent education	Urban or regional
Sustainable growth	Meeting climate change and energy objectives	Supporting in all sectors the shift towards a low-carbon, resource efficient and climate resilient economy	Urban or regional
Sustainable growth	Meeting climate change and energy objectives	Promoting renewable energy sources	Urban or regional
Sustainable growth	Meeting climate change and energy objectives	Upgrading Europe's energy network	Regional
Sustainable growth	Meeting climate change and energy objectives	Promoting sustainable transport	Urban or regional
Sustainable growth	Meeting climate change and energy objectives	Correcting and preventing unsustainable use of resources	Urban or regional
Sustainable growth	Meeting climate change and energy objectives	Removing bottlenecks in key network infrastructures	Regional
Inclusive growth	Promoting employment	Increasing labour market participation of women and men, reducing structural unemployment and promoting job quality	Urban or regional
Inclusive growth	Promoting employment	Developing a skilled workforce responding to labour market needs and promoting lifelong learning	Urban or regional
Inclusive growth	Promoting social inclusion and reducing poverty	Promoting social inclusion and combating poverty	Urban or regional
Inclusive growth	Promoting social inclusion and reducing poverty	Combating social exclusion of deprived neighbourhoods	Urban
Inclusive growth	Promoting social inclusion and reducing poverty	Preventing social exclusion linked to migration	Urban or regional

we take the case of the smart growth priority “Enhancing accessibility to and use and quality of information and communication technologies”, there is no reason to suppose that the outcomes or results of an intervention in this arena should primarily be limited to cities. As such, this thematic priority and the associated outcome indicators are better understood at the regional level rather than at the urban level. In contrast, if we consider the thematic priority “Combating social exclusion of deprived neighbourhoods”, the dense concentrations of poverty and patterns of income in segregated urban neighbourhoods mean that the likely interventions and outcomes of these interventions, particularly in many EU-15 countries in particular, are primarily understood as being urban issues. As we see from Table 1 many policy interventions can be understood equivalently in either an urban or a regional context, but this is not always so. In each specific case, the choice as to the appropriate territorial scale for the place-based analytical approach depends on both the geographical nature of the context and also extent of the interdependencies which are intended to be influenced by the policy interventions.

5 Place-Based Regional Innovation Strategies and Smart Specialisation

In EU-15 countries, innovation is to be a major thematic priority in the reformed Cohesion Policy. In particular, the proposed regulations published in October 2011 include conditionalities requiring a region having to develop a regional innovation and smart specialisation strategy. Smart specialisation strategies are framed in order to obtain the most efficient innovation results with the most effective way of spending public resources in line with the principles of the Europe2020 strategy. The idea of smart specialisation is based on the notion that regions cannot achieve everything in science technology and innovation, so it is crucial to follow a thoughtful process of goal-setting, prioritisation, and the concentrating of resources in certain domains of expertise based on the needs and available resources of each region. These strategies should concentrate the funding on a smaller number of priorities which are better linked with the objectives of smart, sustainable and inclusive growth defined by Europe 2020 and also by the Innovation Union flagship initiative, and also focus on the importance of results and the monitoring of progress towards agreed objectives in order to maximise policy returns.

Many EU regions have already been proactive in developing such policies and here we document various examples of European regions which over recent years have developed such policies in line with the objectives of Europe 2020. These policy strategies share various features and patterns but also exhibit differences related with the historical and economic environment of the location. Certain policy development aspects related to the design, implementation, monitoring or diagnosis of the policies, as promoted by these regions can be considered as good examples of what nowadays we understand as smart specialisation policy strategies. The regions highlighted here are: Flanders, Belgium; Navarra, Spain; Lower Austria, Niederösterreich, Austria; Skåne, Sweden; Berlin-Brandenburg, Germany; West Midlands, United Kingdom; Lahti, Päijät-Häme, Finland; Silesia, Slaskie, Poland; Limburg, Netherlands; Nord-Pas-de-Calais, France; Emilia Romagna, Italy.

In order to usefully classify these example regions, as above we use the OECD (2011a) innovation classification scheme, but here we use the categories broken down into the more detailed sub-categories than simply the three broad categories employed above. In the following tables the classification is adapted:

Region	OECD detailed innovation classification	Country
1. Flanders	Medium-tech manufacturing and service provider region	BE
2. Navarra	Medium-tech manufacturing and service provider region	ES
3. Lower Austria (Niederösterreich)	Traditional manufacturing region	AT
4. Skåne County, South Sweden	Knowledge and Technology Hub	SE
5. Berlin (I) – Brandenburg (II)	(I) Knowledge intensive city/capital district, (II) Structural inertia/(de)industrialising region	DE
6. West Midlands	Medium-tech manufacturing and service provider region	UK
7. Lahti (Paijanne Tavastia)	Knowledge and technology hub (Southern Finland)	FI
8. Silesia (Podkarpackie)	Structural inertia or de-industrialising region	PL
9. Province of Limburg (NL)	Service and natural resource region in knowledge-intensive country (Western Netherlands) and knowledge and technology hub (Southern Netherlands	NL
10. Nord-Pas de Calais	Medium-tech manufacturing and service provider region	FR
11. Emilia Romagna	Traditional manufacturing region	IT

Each of the regions highlighted here display interesting aspects of placed based smart specialisation strategies for economic transformation, and here we focus on several particular aspects, namely identification and prioritisation.

Some regions have followed an evidence-based process of the discovery of the industrial engines of regional economic growth based on an identification the *embedded sectors* in the region (technological domains) combined with an analysis of the *related industries and activities* that complement and improve the regional economic structure. The aim of this discovery process is not to identify the hottest topics in the region's R&D but rather to identify the domains where new research and development and innovation projects will complement the region's other productive assets, so as to create future domestic capability and interregional competitive advantage. This is the logic on which policy prioritisation is to be based, exploiting existing resources. This discovery process also explicitly builds on a region's historic strengths and existing assets. Interesting examples are the cases of Skåne (which used network analyses in order to identify the level of relatedness among the industries located in the region), Limburg (in which BAK Basel Economics were involved in mapping the regions at the TTR-ELAt region), Lower Austria (which used a detailed SWOT analysis approach).

In terms of prioritisation, smart specialisation strategies imply that the prioritisation process should avoid fragmentation and duplication or imitation.

Table 2 Policy prioritisation in smart specialisation

Region	Sector of origin	Priorities
Lower Austria	Construction sector	Energy efficiency Core competencies Energy efficient construction and refurbishment (passive house, zero energy house) Healthy interior environments
Silesia	Mining sector	Technologies for the energy and mining sector, e.g. coal combustion technologies, clean coal technologies, fuel cells, renewable energy sources, carbon storage, gas processing technologies, recognition and protection of coal reserves, ICT, nanotechnologies and medical technologies
Skåne	Shipbuilding industry	Moving media in Malmö Western harbour
West Midlands	Automotive sector	Energy and the environment, healthcare and health sciences, and process industries

It should be based on the accumulation of critical mass on certain regional domains of expertise, combined with other activities with potential growth development. The prioritisation process should not be based in one single sector instead it should be a cross-fertilisation process. In some cases, it could involve strategies beyond the regional level (inter-regional development strategies or international development strategies). In some cases the prioritisation is based on processes of restructuring within the regional of the economy. Table 1 provides some interesting examples of this approach from Lower Austria, Silesia, and West Midlands (Table 2).

However, this is not the only possible policy prioritisation approach. In some cases, policy processes are based on the priority of the modernisation and better development of embedded industries. Examples here include Navarra (which prioritises the basic sectors linked to the economic growth and development of their region), and Flanders (which prioritises the transportation-logistics services based on their privileged geographical location in Europe). In some regions policy prioritisation is based on processes of technological diversification from the embedded industries to a related set of industries linked with the embedded ones, and examples here are Navarra, Nord-pas-de-Calais and Skåne, and West Midlands. Some regions have also prioritised on the basis of a combined approach involving different priorities including the development and modernisation of certain sectoral environments. Examples of regions which have used this combined approach are Navarra (mapping of basic, strategic and future commitments), and Nord-pas-de-Calais (which groups priorities into strongly dependent, growth potential and research excellence spheres). Various regions such as Lower Austria and Nord-pas-de-Calais have aimed at the construction of new and supporting infrastructures to help the industries in their development and to facilitate the Triple Helix collaboration. These policies typically involve the creation of different *technopols* and technology and knowledge parks and centres. Finally, some regions have prioritised around the need to foster connectivity and inter-regional and intra-regional knowledge flows. Examples here include Limburg and the TTR-ELAt strategy, the Berlin-Brandenburg strategy, or the international development strategies employed by Skåne.

Almost all of these regions have already developed evaluation systems and expertise which allow the strategy to be monitored, although the domains (sectors/fields being monitored are different): Lower Austria (Technopol program, five clusters); Navarra (Moderna plan, basic/strategic/future commitments set of sectors); Flanders (Pact2020, six clusters); Berlin-Brandenburg (InnoBB, five future fields); Lahti (three lines of expertise); Nord Pas de Calais (Poles of Competitivite and Tecnopols, different stages in the prioritisation); Silesia (eight technological areas and four development pathways); Skåne (five cluster development).

6 Conclusions

The chapter has examined various practical pre-requisites which while not being sufficient, must necessarily be considered when implementing a placed-based approach. As we see, in order to foster an integrated approach all of the characteristics of a region must be considered simultaneously. Programming design and the design of funding arrangements must be developed in tandem precisely because all regions are a specific mixture of Europe 2020 characteristics. As we saw in Sects. 2 and 3 place-based policies by definition will be different between different places, and there is no "one-size-fits-all" policy solution. Integrated and place-based solutions require these differences to be taken on board explicitly, and as we saw in Sect. 3, EU regions can be neatly characterised into one of 24 broad types according smart, sustainable and inclusive growth dimensions of the Europe 2020 agenda. Obviously, within each of these 24 types there are many sub-categories, but as a starting point for place-based policy design, it is essential to consider the major combined features of the region. As we saw in Sect. 4, however, it is also necessary to consider the appropriate analytical approach for us to adopt when designing the policy, whether a regional or an urban approach. While these two approaches are closely related they are indeed rather different, in that the challenges relating to the nature of the interdependencies, the institutional issues and multi-level governance arrangements, the nature of cohesion issues faced, the concepts of mobility and economic geography employed, and the nature of the conditionalities to be addressed, all differ somewhat. In particular, these differences also impact on the choice of results/outcome indicators to be employed. Finally, while these pre-requisites may appear at first to be rather daunting, in actual fact many EU regions are already very experienced in these ways of designing and operating development policies, under the broad banner of smart specialization strategies. Following the publication in October 2011 of the proposed regulations for the reformed post-2013 Cohesion Policy it is clear that other regions now need to follow suit.

References

Barca F (2009) An agenda for a reformed cohesion policy: a place-based approach to meeting European union challenges and expectations. Independent report prepared at the request of the European Commissioner for Regional Policy, Danuta Hübner, European Commission, Brussels

Barca F, McCann P (2011) Methodological note: outcome indicators and targets-towards a performance oriented EU cohesion policy and examples of such indicators are contained in the two complementary notes on outcome indicators for EU2020 entitled meeting climate change and energy objectives and improving the conditions for innovation, research and development. See: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/performance_en.htm, Accessed 1 October 2011

CAF (2010) Desarrollo Local: Hacia un Nuevo Protagonismo de Las Ciudades y Regiones, Corporación Andina de Fomento, Caracas

ESPON (2010) DEMIFER: demographic and migratory flows affecting European regions and cities, applied research 2013/1/3, Deliverable 3: typology of regions, ESPON 2013 Programme, ESPON Luxembourg

European Commission (2010) Investing in Europe: fifth cohesion report on economic, social and territorial cohesion. European Commission, Brussels

OECD (2009a) How regions grow: trends and analysis. Organisation for Economic Cooperation and Development, Paris. ISBN 978–92–64–03945–2

OECD (2009b) Regions matter: economic recovery, innovation and sustainable growth. Organisation for Economic Cooperation and Development, Paris. ISBN 978–92–64–057651–8

OECD (2009c) OECD regions at a glance 2009. Organisation for Economic Cooperation and Development, Paris. ISBN 978–92–64–05582–7

OECD (2011a) Regions and innovation policy. Organisation for Economic Cooperation and Development, Paris

OECD (2011b) OECD regions at a glance 2011. Organisation for Economic Cooperation and Development, Paris

Part II
Institutions and Culture

Cultural Diversity and Economic Performance: Evidence from European Regions

Elena Bellini, Gianmarco I.P. Ottaviano, Dino Pinelli, and Giovanni Prarolo

Abstract We investigate the relationship between diversity and productivity in Europe using an original dataset that covers the NUTS three regions of 12 European countries. In so doing, we follow the empirical methodology developed by Ottaviano and Peri (J Econ Geogr 6:9–44, 2006). The main idea is that, as cultural diversity may affect both production and consumption through positive or negative externalities, the estimation of both price and income equations is needed to identify the dominant effect. Based on this methodology, we find that diversity is positively correlated with productivity. IV estimates confirm the results, suggesting that causation runs from the former to the latter.

Keywords Cultural Diversity • Productivity • European regions

1 Introduction

Diversity is increasingly at the core of public debates and a central issue for policy-making in the EU. The falling costs of mobility across EU regions and nations are making the agglomeration of economic activities increasingly important in shaping not only economic development *per se* (D'Costa et al. 2013 in chapter "Agglomeration and Labour Markets: The Impact of Transport Investments on Labour Market Outcomes"), but also the current and future socio-cultural map of Europe.

JEL code: R11, R12, O52

E. Bellini • D. Pinelli
FEEM,

G.I.P. Ottaviano
London School of Economics and Bocconi University

G. Prarolo (✉)
University of Bologna
e-mail: giovanni.prarolo@unibo.it

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_7,

From an economic point of view, the key question is whether a culturally diversified society is more or less efficient than a culturally homogenous one.[1] On the one hand, cultural diversity generates potential costs as it may entail racism and prejudices resulting in open clashes and riots (Abadie and Gardeazabal 2003), as well as conflicts of preferences leading to a suboptimal provision of public goods (Alesina et al. 1999, 2004). On the other hand, cultural diversity creates potential benefits by increasing the variety of goods, services and skills available for consumption, production and innovation (Lazear 1999; William and O'Reilly 1998; Ottaviano and Peri 2005, 2006; Berliant and Fujita 2004). These opposite outcomes deeply influence the shape of regional development: the standard approach to the economics of agglomeration usually does not touch upon the issue of heterogeneous agents along the dimension of cultural traits. How diverse individuals interact and behave in the context of dense spaces, such as many European regions, is crucial if we want to understand the future development of our economies. With this respect, the chapter by Bakens and Nijkamp (2013) "paves the road" for the study of these issues.

While keeping in mind the importance of spatial and agglomeration issues for the study of regional development (in particular for Europe), we nonetheless do not address the interplay of cultural diversity with them.[2] What we do, instead, empirically is to net out the pure effect of agglomeration forces through an instrumental variable approach and estimate the direct, causal effect of diversity on productivity of European regions.

Empirical analyses show that cultural diversity can be associated with both positive and negative outcomes. An early paper by Easterly and Levine (1997) shows that richer diversity is associated with slower economic growth.[3] Despite strong criticism (see for example Arcand et al. 2000), the results by Easterly and Levine have been confirmed by a number of studies. In particular, Alesina and La Ferrara (2005) find that going from perfect homogeneity to complete heterogeneity (i.e., the index of fractionalisation going from 0-there is just one group-to 1-each individual forms a different group) would reduce a country yearly growth performance by 2 %. Angrist and Kugler (2003) find a small but significant negative impact of migration on employment levels in the EU. La Porta et al. (1999) and Alesina et al. (2003) argue that higher levels of diversity might result in suboptimal decisions on public good provision, consequently damaging the growth performance in the long run. They show that diversity is negatively correlated with measures of infrastructure quality, illiteracy and school attainment, and positively correlated with infant mortality.

[1] Apart from cultural diversity, culture per se, in the form of social capital, has proven to be an important driver of economic activity at the regional level, as Crescenzi et al. (2013) discuss in chapter The 'Bright' Side of Social Capital: How 'Bridging' Makes Italian Provinces More Innovative.

[2] For a theoretical model on the interplay between cultural diversity and urban economics, see Ottaviano and Prarolo (2009).

[3] Easterly and Levine (1997) use a fractionalisation index of diversity calculated from the Atlas Narodov Mira (1964).

However, the conclusion that diversity has a negative effect on the economy needs to be further qualified. Collier (2001) argues that diversity has negative effects on productivity and growth only in non-democratic regimes. Alesina and La Ferrara (2005) find that diversity has a more negative effect at lower levels of income (implying that poorer countries suffer more from ethnic fragmentation). Easterly (2001) constructs an index of institutional quality aggregating data from Knack and Keefer (1995) on contract repudiation, expropriation, rule of law and bureaucratic quality. He finds that the negative effect of ethnic diversity is significantly mitigated by 'good' institutions.

Moreover, a number of studies relating diversity to urban agglomeration suggest that diversity can have also positive economic consequences. Earlier works by Jacobs (1961) and Bairoch (1985) see diversity as the key factor of success of a city: the variety of commercial activities, cultural opportunities, aspects, inhabitants, visitors as well as the variety of tastes, abilities, needs and even obsessions are the engine of development (Jacobs, 1961, p. 137). More recently, Florida (2002) argues that diversity contributes to attract knowledge workers thereby increasing the creative capital of cities and the long-term prospect of knowledge-based growth (Gertler et al. 2002). In the economic literature, Glaeser, Scheinkman and Shleifer (1995) examine the relationship between a variety of urban characteristics in 1960 and urban growth (income and population) between 1960 and 1990 across US cities. They find that racial composition and segregation are basically uncorrelated with urban growth. However, segregation seems to positively influence growth in cities with large non-white communities. Alesina and La Ferrara (2005) use the basic specification of Glaeser, Scheinkman and Shleifer (1995) to estimate population growth equations across US counties over 1970–2000. Consistently with their result at the country level discussed above, they find that diversity has a negative effect on population growth in initially poor counties and a less negative (or positive) effect for initially richer counties. Ottaviano and Peri (2006) find that, on average, US-born citizens are more productive in a culturally diversified environment. This is robust to the use of instrumental variables, thus implying a causal relationship between diversity and productivity. Manacorda, Manning and Wadsworth (2006) and D'Amuri, Ottaviano and Peri (2008) find results similar to Ottaviano for the UK and Germany, respectively. A recent paper by Carrasco, Jimeno and Ortega (2008), based on micro data of the Spanish labour market, find no evidence of a negative impact of immigration on natives' employment and wages.

The existing literature is either based on cross-country analyses or focused primarily on the US, with limited evidence for individual EU countries. A comprehensive analysis of the effects of diversity on productivity across EU countries is, nonetheless, still missing. Our aim is to take a first step in this direction. In so doing, we assemble a new dataset covering the NUTS 3 regions of 12 countries of the EU15 (Austria, Belgium, Denmark, France, former Western Germany, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden and the United Kingdom). Then, we follow the empirical methodology developed by Ottaviano and Peri (2006) in the case of US cities. The main idea is that, as cultural diversity may affect both production and consumption through positive or negative externalities, the estimation of both price and income equations is needed to identify the dominant effect. Based on this methodology, we find that diversity is positively correlated with productivity. Moreover, through instrumental variable estimations, we find evidence that causation runs from the former to the latter. These results for EU regions are broadly consistent with those found by Ottaviano and Peri (2005, 2006) for US cities.

The rest of the chapter is organised as follows. Section 2 describes the dataset. Section 3 presents some facts about the diversity of EU regions. Section 4 introduces the analytical framework and discusses the results of the econometric analysis. Section 5 concludes.

2 The Dataset

The dataset includes demographic, economic and geographical data for over 900 European regions from 12 countries of the EU15 (Austria, Belgium, Denmark, France, former Western Germany, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden and the United Kingdom).[4] Data are collected at NUTS 3 level and refer to two different points in time: 1991 (1990 for Finland and the Netherlands) and 2001 (2000 for Finland and the Netherlands; 1999 for France). The choice of reference years is constrained by the availability of Census data in each country (more on this below).[5]

Data for GDP, employment (three sectors), unemployment, active population and area are from Eurostat's Cronos REGIO database. For GDP, employment and area the REGIO data are of good quality and complete. The series for unemployment and active population includes missing values. In those cases, data at NUTS 3 level were constructed by sharing out NUTS 2 data (using data from the same REGIO or, alternatively, data kindly provided by Cambridge Econometrics when REGIO data were missing).

Information on non-tradable prices is proxied by restaurant prices derived from the Michelin Guides of each country for the reference years. By exploiting the rating system of Michelin we have constructed price indexes that refer to restaurants of comparable quality across countries and cities. In particular, the restaurant price for each region is calculated by averaging the prices of all *deux fourchettes* restaurants reported in the guide for that region. Restaurant prices exclude fixed-price menus.

Demographic data are from the National Statistical Institutes of each country (mostly from national Census Surveys or Registry data). Data cover population by gender, age (0–14; 15–39; 40–64; 65 or over), marital status (unmarried, married,

[4] The dataset has been developed at Fondazione Eni Enrico Mattei with support from the European Commission, 6th RTD Framework Programme, Contract no SSP1-CT-2003–502491 (PICTURE).

[5] In the following, when needed we will simply refer to 1991 and 2001 data.

divorced, widow); and level of education (basic or not educated, secondary school, degree or higher education – consistent with the ISCED classification of the OECD) and citizenship (autochthonous, other EU countries, other European countries, Africa, America, Asia, Oceania, unknown) or country of birth (only for the UK and Ireland). For education and citizenship, the national sources report data using classifications that are often inconsistent across countries and in any case at different levels of detail. The categories we use (see above) are obtained by grouping the raw data (by geographical area in case of citizenship; by number of years in education for education) to obtain a classification that is the most detailed and that at the same time is consistent across countries. For education in 1991, however, for certain countries, there were too many missing observations and the information available to re-classify data was too incomplete to reconstruct internally consistent data.

'Cultural diversity', the key variable in our empirical analysis, has been constructed using the shares of foreigners in each regions. For most of the countries in our analysis, foreigners are foreign citizens, while for UK and Ireland they are foreign born citizens. In the empirical section we will discuss how this difference could affect our results. Obviously, a complete representation of the cultural diversity in a population would require data concerning a wide range of 'individuals' characteristics: their (cultural) background and personal histories, their current and past economic and social statuses, the environment in which they live, their psychological features, etc. However, this would amount to a complex statistical distribution and to a coverage that cannot be achieved for our kind of study. The two indexes we use (see Ottaviano and Pinelli 2007 for a review of indexes of diversity) are the simple share of foreigners and the *fractionalisation* index, the latter being the complement to one of the Herfindal Index of concentration across groups. Given a population of L_c individuals divided in $i=1,\ldots, M$ cultural groups, the fractionalisation index can be calculated as:

$$d_c = 1 - \sum_{i=1}^{i=M} \left(\frac{L_{ci}}{L_c} \right)^2 \tag{1}$$

where L_{ci} is the number of individuals that in region c belong to group i. The index (known in biology as the Simpson index of diversity) measures the probability that two individuals randomly extracted belong to different groups. The index varies between 0 and 1 and increases with both the number of groups and the evenness of the distribution of individuals across groups.

3 Diversity in European Regions

We can now use the database presented in Sect. 2 to discuss the main features of the European landscape of diversity and how this has changed over the period 1991 to 2001. There is a clear change in the distribution patterns of migration from 1991 to

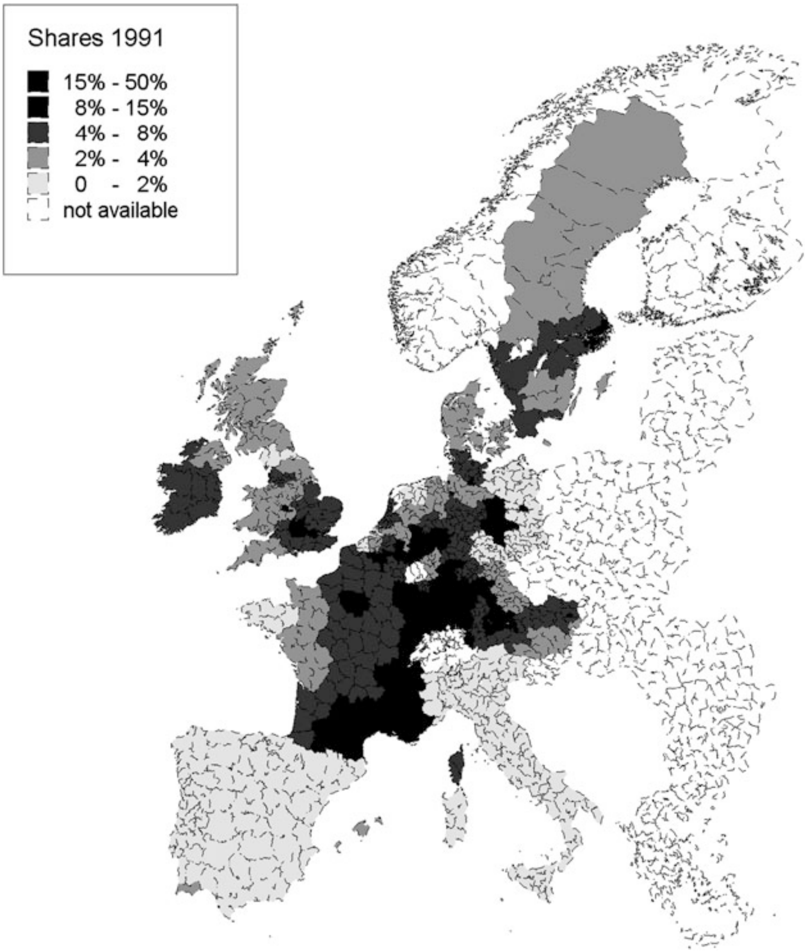


Fig. 1 Shares of foreigners in European regions, 1991

2001. Figures 1 and 2 show the percentage of foreigners in European regions respectively in 1991 and 2001.[6] In 1991 (Fig. 1), diversity characterised only regions in the core of Europe: Paris and Lyon in France, Belgium, the Netherlands, Germany's large cities and the south of the UK. Regions of Spain, Italy, Austria and Nordic countries were fairly homogenous. In Italy and Spain the percentage of residents with foreign citizenship was below 2 % everywhere. The situation has rapidly changed in the 1990s. In 2001 (Fig. 2) most Austrian regions have reached a percentage of foreigners higher than 8 % and the percentage of foreigners in most regions of Italy and Spain is between 4 and 8 %. Overall, the share of foreigners increased from 5.6 % in 1991 to 6.9 % in 2001.

[6] Here and in what follows, we will refer to 'foreigner' as 'foreign-born' in the UK and Ireland, and 'with foreign citizenship' elsewhere.

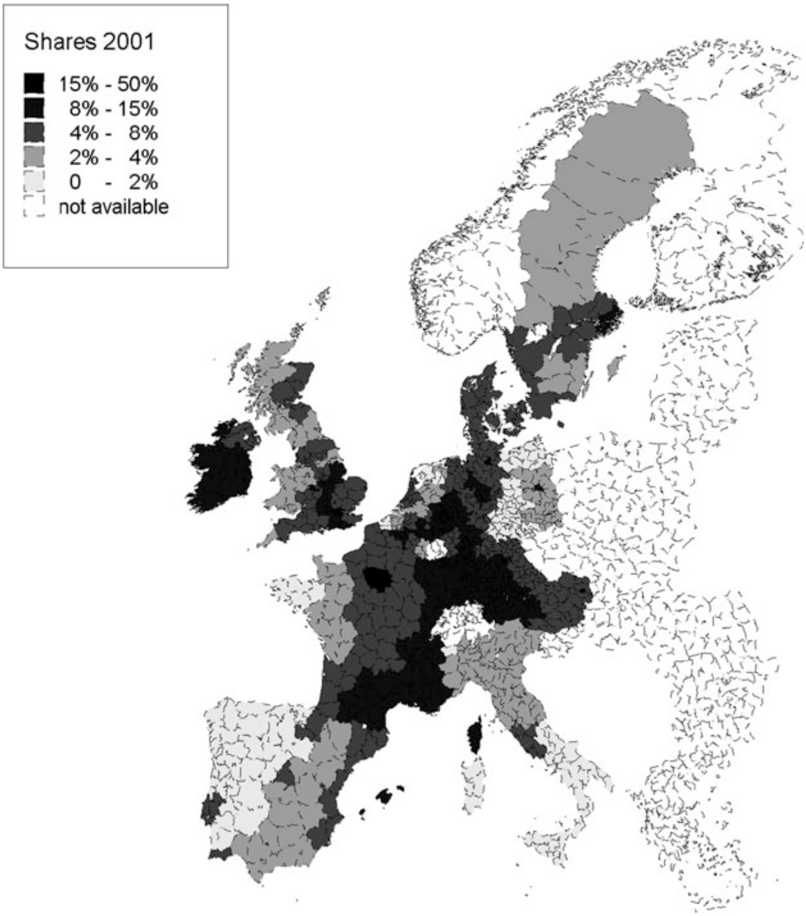


Fig. 2 Shares of foreigners in European regions, 2001

The change in migration patterns from 1991 to 2001 is confirmed in Table 1 showing the most and the least diverse EU regions in 1991 and 2001 ranked according to the Simpson index of diversity (fractionalization) discussed in Sect. 3. The top of ranking included only French and UK urban regions in 1991, joined in 2001 by Brussels and surroundings, as well as Wien following the immigrant inflows from Eastern Europe after 1989. In 1991 the bottom of the ranking included only rural Italian and Spanish regions (which showed practically no diversity). The picture is different in 2001. Some degree of diversity also characterises the most homogenous regions and some of the Italian and Spanish regions have been replaced by rural regions in France and Belgium.

Table 1 Most and least diverse European regions, 1991 and 2001

Most diverse	1991 Simpson	1991 Share of foreigners (%)		2001 Simpson	2001 Share of foreigners (%)
Inner London (UK)	334	27.8	Inner London (UK)	409	33.6
Seine-Saint-Denis (FR)	261	24.1	Seine-Saint-Denis (FR)	315	27.9
Outer London (UK)	230	18.0	Outer London (UK)	304	22.9
Paris (FR)	228	21.7	Paris (FR)	243	21.9
Bruxelles (BE)	223	28.6	Hauts-de-Seine (FR)	208	18.1
Hauts-de-Seine (FR)	190	17.4	Val-de-Marne (FR)	203	19.4
Val-de-Marne (FR)	166	17.6	Val-d'Oise (FR)	191	17.8
Val-d'Oise (FR)	162	15.7	Bruxelles (BE)	182	27.1
Rhône (FR)	136	13.8	Wien (AT)	181	16.4
Leicestershire (UK)	136	9.1	Berkshire (UK)	175	13.1
Least diverse	**1991 Simpson**	**1991 Share of foreigners (%)**		**2001 Simpson**	**2001 Share of foreigners (%)**
Taranto (IT)	1	0.1	Benevento (IT)	5	0.4
Terni (IT)	1	0.1	Vandée (FR)	5	0.4
Albacete (ES)	1	0.1	Taranto (IT)	4	0.6
Badajoz (ES)	1	0.1	Oristano (IT)	4	0.3
Jaen (ES)	1	0.1	Ypres (BE)	4	0.3
Ciudad Real (ES)	1	0.1	Enna (IT)	4	0.4
Zamora (ES)	1	0.1	Tâmega (PT)	4	0.5
Isernia (IT)	1	0.1	Brindisi (IT)	4	0.4
Campobasso (IT)	1	0.1	Eeklo (BE)	4	0.2
Chieti (IT)	0	0.0	Dixmude (BE)	2	0.6

Source*:* Authors' calculation based on national Censuses data for population by *country of birth* for Ireland and the UK and *citizenship* for the other countries (see Sect. 3)
Notes*:* Data are for 1991 and 2001 except for the Netherlands (1990 and 2000) and France (1991 and 1999). Finnish and some German regions are excluded (1991 data are not available)

4 Analytical Framework and Empirical Results

To structure the empirical analysis, we rely on the analytical framework developed by Ottaviano and Peri (2006), who model an open system of cities in which 'diversity' affects both the productivity of firms and the satisfaction of consumers through localised external effects. Both the model and the identification procedure of the impact of diversity on city dwellers build on Roback (1982).

A priori, diversity could have both an amenity or a disamenity effects on a region's residents, as well as productivity could be enhanced or harmed by the diversity of the workforce. When workers and firms are mobile, the sign and the dominant nature of the net effect of diversity can be inferred by looking at the joint effects of diversity on wages and land rents: if both effects are positive, diversity has a positive (net) impact on productivity. The underlying logic is simple enough.

Consider two otherwise identical regions that differ only in terms of diversity and a situation in which wages are higher in the more diverse region. Higher wages are compatible with two very different scenarios. In the first scenario firms in the more diverse region pay higher wages to workers because these are more productive. In the second scenario firms in the more diverse region pay higher wages to workers because these have to be compensated for their aversion to diversity. This second scenario can be, however, ruled out if also land rents are higher in the more diverse region as workers averse to diversity would never pay a premium in order to reside in a more diverse environment.

We now present the results of the empirical analysis, which is carried out in three steps. We first estimate the wage equation. As wage data for European regions and cities are scattered and not available at NUTS 3 level, we use GDP per capita as a proxy.[7] Under the model assumption of free firm mobility the two measures are equivalent, as profits are equalised across regions and income differentials are entirely driven by wage differentials.

Second, we estimate the rent equation. EU-wide comparable data for land rents at city level are not available (and data for a close proxy such as house prices are only available for a restricted number of major cities). However, rents *de facto* capture non-tradable good prices (see footnote 9), which we proxy by the average prices (in logs) of *two-forchettes* restaurants as detailed in Sect. 2.[8]

Third, since our independent variable (diversity) is potentially endogenous, we perform instrumental variables (IV) estimations in order to net out the (possibly positive) effect running from wages and the (possibly negative) effect running from rents to diversity. In fact, since migrants are more mobile than natives (Schündeln 2007), an exogenous positive shock in wages should attract relatively more migrants than natives, leading to an increase in diversity. Conversely, an exogenous positive shock in the land rent should repel relatively more migrants leading to less diversity.

The very poor quality of education data for 1991 does not allow us to perform panel estimations with reliable controls. We therefore perform a cross section analysis using only 2001 data and exploiting 1991 data to construct the instruments. The basic equation is thus the following:

$$ln\, y_c = D_r + \beta div_c + \varphi^I \boldsymbol{X}_c + e_c \qquad (2)$$

where c indexes the NUTS 3 province. We exclude formerly Eastern German provinces, which leaves us with 844 NUTS 3 observations for 171 NUTS 2: the average number of provinces included in a region is around five. As discussed, the

[7] REGIO also contains data for 'Compensation of employees' but scattered and only available at NUTS 2 level.

[8] Where data availability makes computation possible, the correlation between restaurant prices and house prices is typically large and positive. For example, in a sample of 12 major Italian cities such correlation was roughly 70 % in 2001.

dependent variable *(ln y_c)* is GDP per capita (in logs) in the income regression and restaurant prices in the price regression. The key regressor is the province's diversity (*div_c*). We use two measures of diversity: the Simpson index (see Sect. 3) and the simple share of foreigners in total population, in some specifications coupled with the Simpson index calculated only among foreigners, as in Ottaviano and Peri (2006).[9]

In all regressions, we introduce NUTS 2 region fixed effects, D_r. Region fixed effects (D_r=1 for the all the NUTS 3 regions belonging to a specific NUTS 2 region; 0 otherwise) control for characteristics, such as institutions and other NUTS 2-specific variables, that apply to all provinces (NUTS 3) in the same region. This means that when NUTS 2 fixed effects are introduced, only the provincial deviations from the NUTS 2 baseline must be explained. This is very important in our case. As discussed in Sect. 2, a potential problem affecting our estimates may derive from differences in the citizenship rules, which would make our diversity measures incomparable across regions. NUTS 2 dummies would however clear from those differences, at least to the extent that citizenship rules affect equally the NUTS 3 in the same NUTS 2 (i.e., it affects the NUTS 2 baseline, but not provincial deviations from it). This is a reasonable assumption, given that citizenship rules are set at national level and their application does not vary between NUTS 3 within the same NUTS 2. As stated above, for the UK and Ireland the resident population is classified by country-of-birth rather than by citizenship. As this goes beyond differences in citizenship rules, we run the regressions for the full sample and excluding the UK and Ireland.

We also include a set $\boldsymbol{X_c}$ of standard control variables (see Temple 1999 for a review of the literature on income and growth regressions) such as the share of agriculture in total employment (*$agri_c$*) to control for differences in industrial structure and the share of inhabitants with at least secondary education (*edu_c*) to control for differences in human capital endowments. The density of population *($dens_c$)* is introduced to control for those 'non-pecuniary' externalities that derive from sheer proximity of economic actors.[10] Market potential *($mpot_c$)* controls for the 'pecuniary' externalities that derive from the agglomeration of economic activities, as highlighted by the new economic geography literature (see Redding and Venables 2004, Ottaviano and Pinelli 2006). Descriptive statistics and the correlation matrix of the variables used in regressions are presented in Annex I.

In what follows we firstly discuss the results for the three steps, and then we compare them with those obtained when dropping the UK and Ireland out of the regressions.

[9] As explained in Sect. 2, population is classified by citizenship in all countries apart from the UK and Ireland for which we use the 'country of birth'.

[10] Local external effects can be positive, due to easier non-market interactions leading to technological externalities (see Ciccone 2002; Ciccone and Hall 1996) or negative, due to higher congestion and consequent waste of resources that make interactions difficult.

Table 2 Wage regressions-OLS

Dep. variable: log(GPDpc)	(1)	(2)	(3)	(4)
Simpson index	5.73068***		3.66791***	
	[1.60738]		[0.76727]	
Share of foreigners		4.82905***		3.21975***
		[0.82310]		[0.50362]
Simpson index among foreigners		0.94970***		0.34025**
		[0.14410]		[0.16062]
Share of agriculture			−0.00816**	−0.00648
			[0.00364]	[0.00414]
Human capital			0.02387***	0.02138***
			[0.00436]	[0.00490]
Density			0.00001*	0.00002***
			[0.00001]	[0.00001]
Market potential			−0.00001	−0.00001
			[0.00001]	[0.00001]
NUTS 2 dummies	yes	yes	yes	Yes
Observations	787	787	679	679
R-squared	0.80	0.82	0.88	0.89

Robust standard errors in *brackets*. Observations are weighted for working population
*significant at 10 %; **significant at 5 %; ***significant at 1 %

4.1 First Step: Income Regressions

Table 2 shows the results of the basic income regressions, estimated with OLS. Robust standard errors are also reported in brackets as heteroskedasticity often characterises cross-regional analyses. Columns (1) and (2) report the results of specifications where the only regressors are the overall Simpson index and the share of foreigners plus the Simpson index calculated among foreigners only. The diversity indices are positive and strongly significant, suggesting positive correlations with the log of GDP per capita, which is our wage proxy. NUTS 2 dummies already explain a lot of variation: regressions including NUTS 2 dummies alone show R-squared around 0.69. Nonetheless, the inclusion of diversity indices significantly adds explanatory power. In columns (3) and (4) we replicate the two previous estimations adding some of the controls described above. The coefficients of diversity indices are a bit smaller but still strongly significant and the coefficients of controls show expected signs, except for the one of market potential that is, however, not significant. The share of agriculture has a negative coefficient, which is significant in column (3), consistently with most findings in literature (see, for example, Bivand and Brunstad 2003). The human capital variable has a positive and strongly significant coefficient, consistent with the growth literature (Temple 2001) and with usual findings in regional economics, such as Maré and Fabling (2013) in chapter "Productivity and Local Workforce Composition". Finally, the density of population has a positive coefficient hinting at positive agglomeration effects as in Ciccone and Hall (1996) and Ciccone (2002). Market

Table 3 Restaurant prices regressions-OLS

Dep. variable: log(RestPrice)	(1)	(2)	(3)	(4)
Simpson index	1.64327***		1.80630***	
	[0.33582]		[0.48860]	
Share of foreigners		1.14742***		0.90911***
		[0.27731]		[0.31238]
Simpson index among foreigners		0.24145***		0.18509**
		[0.08272]		[0.08788]
Share of agriculture			−0.00486	−0.00510*
			[0.00299]	[0.00294]
Density			0	0.00001
			[0.00001]	[0.00001]
Market potential			−0.00001	0
			[0.00001]	[0.00001]
NUTS 2 dummies	yes	yes	yes	Yes
Observations	686	686	630	630
R-squared	0.96	0.96	0.97	0.97

Robust standard errors in *brackets*. Observations are weighted for working population
*significant at 10 %; **significant at 5 %; ***significant at 1 %

potential is not precisely estimated, probably because of its relatively high correlation (more than 0.3) with population density.

Under the realistic assumption of limited labour mobility, in the light of Fig. 2 such results would point at a positive effect of diversity on 'firms' productivity. Nevertheless, in the presence of labour mobility, higher wages in more diverse regions could simply reflect aversion to diversity rather than a genuine effect on productivity. To rule this possibility out, in what follows we study the relationship between diversity and local prices (see *Second Step* below).

4.2 Second Step: Price Regressions

Table 3 shows the results of the prices regressions following the same structure of Table 2. All regressions have large explanatory power once we control for NUTS 2 fixed effects. This implies that most of the variation in restaurant prices is not explained by local NUTS 3 characteristics but possibly by less 'local' determinants. However, coefficients are positive and significant for all the diversity measures, thus revealing a (small) positive relation between diversity and land rents. This small effect is consistent with low labour mobility. The coefficients of control variables are never significantly different from zero. The exception is the coefficient on the share of agriculture. This is negative and marginally significant in specifications (3) and (4), confirming that a higher specialisation in agriculture is negatively associated with productivity.

Following our identification strategy, the positive signs of the diversity measures' coefficients rule out the possibility that diversity acts as a consumption disamenity and, thus, point out a positive correlation between diversity and productivity. Short of a randomized experiment, we cannot be sure that this positive correlation reveals a causal link from the former to the latter due to possible reverse causation. We address this concern through instrumental variables (IV) in *Third Step* below.

4.3 Third Step: Instrumental Variables

The idea is to substitute our potentially endogenous diversity measures by a set of proxies correlated with diversity within regions in 2001 but not otherwise correlated with the residuals of regressions (11) and (12). To this end, we use the 'shift-share methodology' firstly applied by Card (2001) and, more recently, by Saiz (2003) and Ottaviano and Peri (2006). The key idea is that new migrants tend to settle close to where old migrants of the same provenience already reside. Accordingly, the predicted end-of-period composition of a region's population in 2001 can be computed on the basis of its beginning-of-period composition in 1991, by attributing to each group in the region its average growth rate in the country to which the region belongs to from 1991 to 2001. Alternatively, one could use the average growth rate in the EU as a whole. We prefer the national growth rate as there are differential behaviours of different ethnic groups that are country-specific. These differential behaviours arise from the heterogeneous nature of 'European countries' populations: different languages, different colonial history of sending countries and different cultures. In Figs. 3 and 4 (where "Rest" indicates foreigners from Oceania and unknown origin) we report the composition of foreign population for years 1991 and 2001, respectively. It is easy to spot patterns that are more likely to be country-specific than widespread across European countries.

The procedure used to build our instrument is based on Card (2001). The underlying hypothesis is that in any NUTS 3 region the current number of immigrants from a certain ethnic group can be decomposed in two components: an exogenous supply-push component that depends on the historical distribution of immigrants and the total inflow from the sending country; a residual component that reflects differences from the historical pattern, which could depend on productivity shocks. It follows that for any ethnic group we can construct the *predicted* actual population as the product of the historical population by the overall (nation-wide in our case) growth rate of the specific ethnic group under consideration. With the predicted actual populations of all ethnic groups we can build an instrument for the actual Simpson index that is independent from current region-specific productivity shocks. We exploit here the change shown in Sect. 3 between the historical geographical distribution of migrants (as shown by 1991 data) and the settling of new migrants in the period 1991–2001 (which would reflect current socio-economic developments, including productivity shocks). Since our instrument depends on the historical distribution, we can assume that it is exogenous to current productivity shocks.

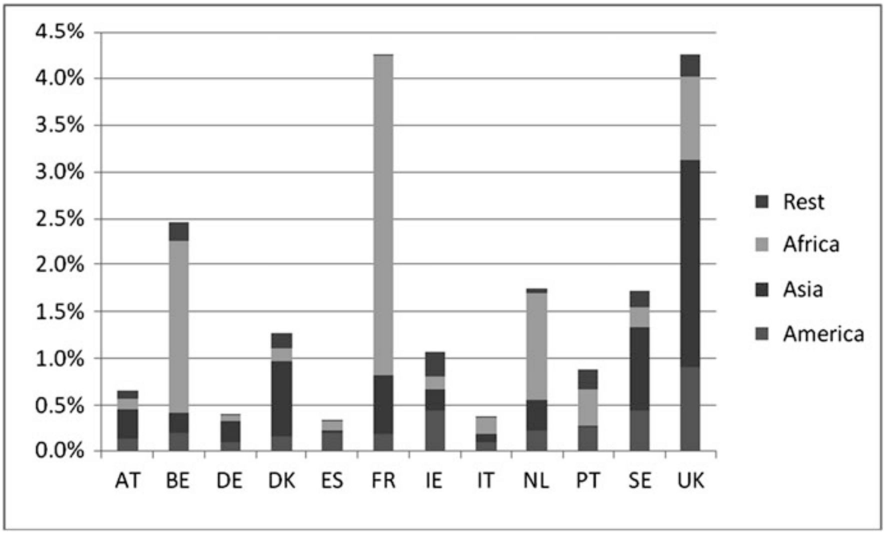


Fig. 3 Composition of foreign population in 1991

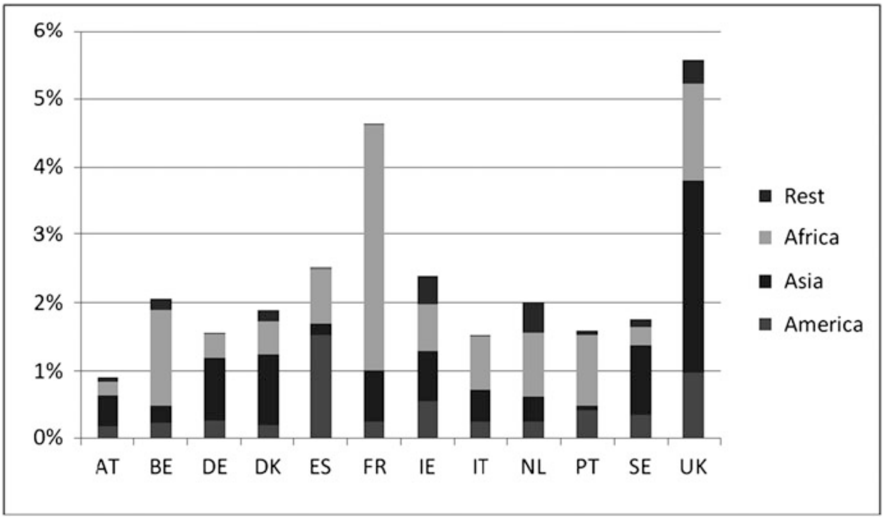


Fig. 4 Composition of foreign population in 2001

To construct the predicted shares of foreigners, we start by building a nation-wide growth rate g_{nf} for each ethnic group f (Europe, America, Asia, Africa and Rest of the World) in each country n. Formally:

$$g_{nf} = \left(N_{nf2001} - N_{nf1991}\right)/N_{nf1991} \tag{3}$$

where N_{nf2001} and N_{nf1991} are the number of inhabitants of ethnic group f in country n in years 1991 and 2001, respectively. Then, for each NUTS 3 region c belonging to country n, we construct the *predicted* number of inhabitants in year 2001 as:

$$N'_{cf2001} = N_{cf1991}(1 + g_{nf}) \tag{4}$$

Then, for each NUTS 3 region, we construct the total *predicted* population in 2001, N'_{c2001}, by summing the N'_{cf2001} across all ethnic groups f. Finally, we follow expression (1) to compute d'_c as the *predicted* Simpson Index for each NUTS 3 region.

An analogous procedure is applied to compute the *predicted* share of foreigners in each NUTS 3 region, defined as total population minus the autochthonous.

The results of the IV estimations of wages and prices regressions are reported in Table 4. In the wage regressions both the coefficients of the Simpson Index and of the share of foreigners turn out to be smaller than the OLS estimates in Table 2 but still strongly significant. The smaller magnitude of coefficients implies that the positive ‘pull effect’ of migrants going to places with higher wages has been netted out by the instrument. In fact, the first stage regression shows large F-tests and not negligible R-squared, indicating that instruments are not weak. The coefficients of the Simpson Index among foreigners also remain positive but lose their significance. Finally, the controls show similar values and are always significant as in Table 2, with the exception of market potential.

In price regressions the previously described ‘pull effect’ should be negative: we expect that, other things being equal, migrants decide to settle in places where prices are lower. This seems indeed to be the case when we instrument the Simpson Index: its coefficient becomes bigger, meaning that, if any, the effect of the price level on diversity is negative, implying a downward bias in the estimate in specification (3) of Table 3. Turning to the last specification, the coefficient on the share of foreigners is not significantly different from the OLS regression and the coefficient on the Simpson Index among foreigners is not significant.

Overall, the values and the pattern of significance of the IV results trace the ones of the OLS, pointing at a positive causal relationship between diversity and productivity possibly accompanied by an amenity effect of diversity on consumption.

4.4 Results Without the UK and Ireland

We now check the robustness of our findings excluding from the regression the UK and Ireland, for which we had to use a different identity marker than for the other countries (i.e., ‘country-of-birth’ rather than ‘citizenship’). With respect to the OLS estimates, the results of the income regressions are very similar to the regressions

Table 4 Instrumental variable regressions

Dep. variable	log(GPDpc)		log(RestPrice)	
	Second stage			
Simpson index	2.44676***		2.24448***	
	[0.70145]		[0.72176]	
Share of foreigners		2.30350***		0.82383**
		[0.50461]		[0.41814]
Simpson index among foreigners		0.1233		0.02727
		[0.19007]		[0.12543]
Share of agriculture	−0.01017***	−0.00956**	−0.00417	−0.00697**
	[0.00359]	[0.00385]	[0.00307]	[0.00281]
Education	0.02184***	0.02123***		
	[0.00381]	[0.00453]		
Density	0.00002***	0.00002***	0	0.00001**
	[0.00001]	[0.00001]	[0.00000]	[0.00000]
Market potential	−0.00001	−0.00001	−0.00001	0
	[0.00001]	[0.00001]	[0.00001]	[0.00001]
NUTS 2 dummies	yes	yes	yes	Yes
Observations	555	555	508	508
R-squared	0.90	0.97	0.93	0.93
	First stage			
Predicted simpson index	0.52830***		0.52040***	
	[0.09187]		[0.09399]	
Predicted share of foreigners		0.61680***		0.63185***
		[0.07202]		[0.07524]
Predicted simpson index among foreigners		0.64574***		0.69154***
		[0.03884]		[0.04032]
Partial R-squared	0.43	0.67 ǀ 0.51	0.41	0.70 ǀ 0.48
F-test	33.07	37.9 ǀ 173.7	30.65	36.4 ǀ 166.0

Robust standard errors in *brackets*. Observations are weighted for working population
*significant at 10 %; **significant at 5 %; ***significant at 1 %

that include the UK and Ireland (shown in Table 2). In column 3, the coefficient of the Simpson Index is now 3.685, at the same confidence level. In column 4, the coefficients for the share of foreigners and the Simpson Index among foreigners are 3.226 and 0.349, respectively, at the 1 % and 5 % significance level. For the price regressions (see Table 3), results are also confirmed. In column 3 the coefficient of the Simpson Index is now equal to 2.104. In column 4, the coefficients of the share of foreigners and the Simpson Index among foreigners are equal to 1.163 and 0.429, respectively, and significant at 1 % level.

With respect to IV estimates, (see Table 4) results are also very similar to the regressions that include the UK and Ireland. For the income regressions, in column 1 the coefficient of the Simpson Index is 2.436, significant at the same 1 % level. In column 2, the coefficients on the share of foreigners and the Simpson Index among foreigners are respectively 2.295 and 0.122, significant at the 1 % level the former, not significant the latter. For the price regressions, the specification with the Simpson Index (column 3) delivers a coefficient of 2.319, significant at a 1 %

confidence level. The specification in column 4 shows also very similar results. The coefficients for the share of foreigners and the Simpson Index among foreigners are 1.089 and 0.227, respectively, significant at the 5 % confidence level the former, not significant the latter.

Overall, excluding the UK and Ireland from the regressions does not seem to change our results. This would imply that the use of country-of-birth rather than citizenship may well affect the NUTS 2 baseline, but apparently not the NUTS 3 deviations from it, which makes NUTS 2 dummies sufficient to correct the estimate. Indirectly, it provides some evidence that national differences in citizenship rules are unlikely to affect our results significantly.

5 Conclusions

In a world increasingly interconnected and global, the economic development of regions depends on the overlapping effects of many forces, whose joint study is far from established. An example of paramount importance concerns the increasingly free movement of diverse people within EU borders and its impacts on the balance of agglomeration and dispersion forces in regions, the cross-interaction of diverse people in the workplace and the aggregation of their preferences in the context of public policies, just to name a few. Our contribution in this chapter is to control for agglomeration forces and partially open up the black box of ‘workers’ interactions, focusing on the partial effect that the cultural diversity of inhabitants of a region may have on productivity. In so doing, we have used an original dataset covering the NUTS three regions of 12 countries of the EU15 to supplement national studies with an overview of the relationship between diversity and economic performance across a large set of European regions. Our results confirm earlier findings (mostly focused on the U.S.), according to which diversity is positively correlated with productivity and that causation seems to run from the former to the latter.

Acknowledgement We thank Andrea Ichino, Francesco D’Amuri and Giovanni Peri as well as seminar participants at FEEM and the University of Bologna for comments. An early version of the paper was presented at the final ENGIME Conference on ‘Multicultural Cities: Diversity, Growth and Sustainable Development’, Rome, November 2004. We thank all the participants for their comments. We thank the European Commission (“SUS. DIV Network of Excellence, Contract No. CIT3-CT-2005–513438”) and the Volkswagen Foundation for financial support.

Annex 1

Descriptive statistics of the variables used in the regressions

Variable	1991 Obs	Mean	Std. dev.	Min	Max	2001 Obs	Mean	Std. dev.	Min	Max
Restaurant prices (log, in Euro 2000)	716	4.266	1.036	1.984	6.044	739	4.165	0.656	2.797	5.225
Wages (log of GDP per capita, in Euro 2000)	842	9.621	0.384	8.007	10.846	844	9.997	0.345	8.797	11.298
Simpson index	663	0.026	0.033	0.000	0.334	787	0.034	0.036	0.002	0.409
Share of foreigners (%)	844	0.052	0.045	0.000	0.286	844	0.063	0.046	0.004	0.336
Simpson index among foreigners	663	0.475	0.187	0.099	0.769	787	0.444	0.196	0.070	0.765
GDP in agriculture (% share, x100)	167	5.835	4.502	0.022	27.628	779	3.267	3.069	0.000	19.948
Education (% share with tertiary education, x100)						801	12.977	8.352	2.220	54.323
Density (inhabitants per square kilometer)	844	452	1,044	3	20,488	844	460	1,044	3	20,246
Market potential (average GDP per capita of surrounding regions)	844	12,202	4,595	2,329	30,273	844	18,806	6,935	3,620	46,695

Correlation matrix of the variables used in the regressions (2001)

	Restaurant prices (log, in Euro 2000)	Wages (log of GDP per capita, in Euro 2000)	Simpson index	Share of foreigners (%)	Simpson index among foreigners	GDP in agriculture (% share, x 100)	Education (% share with tertiary education, x100)	Density (inhabitants per square kilometer)
Restaurant prices (log, in Euro 2000)	1							
Wages (log of GDP per capita, in Euro 2000)	0.45	1						
Simpson index	−0.09	0.39	1					
Share of foreigners (%)	0.51	0.62	0.63	1				
Simpson index among foreigners	−0.80	−0.24	0.27	−0.45	1			
GDP in agriculture (% share, x 100)	−0.50	−0.56	−0.22	−0.50	0.34	1		
Education (% share with tertiary education, x100)	−0.57	0.18	0.55	0.14	0.54	0.06	1	
Density (inhabitants per square kilometer)	0.13	0.43	0.59	0.52	0.03	−0.31	0.34	1
Market potential (average GDP per capita of surrounding regions)	0.68	0.55	0.22	0.56	−0.47	−0.51	−0.15	0.33

References

Abadie A, Gardeazabal J (2003) The economic costs of conflict: a case study for the Basque country. Am Econ Rev 93(1):113–132

Alesina A, La Ferrara E (2005) Ethnic diversity and economic performance. J Econ Lit 43:762–800

Alesina A, Baqir R, Easterly W (1999) Public goods and ethnic division. Q J Econ 111 (4):1243–1284

Alesina A, Devleschawuer A, Easterly W, Kurlat S, Wacziarg R (2003) Fractionalization. J Econ Growth 8:155–194

Alesina A, Baqir R, Hoxby C (2004) Political jurisdictions in heterogenous communities. J Polit Econ 112(2):384–396

Angrist JD, Kugler AD (2003) Protective or counter-productive? Labour market institutions and the effect of immigration on EU natives. Econ J 113:302–331

Arcand J-L, Guillaumont P, Guillaumont Jeanneney S (2000) How to make a tragedy: on the alleged effect of ethnicity on growth. J Int Dev 12:925–938

Atlas Narodov Mira (1964) Moscow: Glavnoe upravlenie geodezii i kartografii

Bairoch P (1985) De Jéricho à Mexico: Villes et economie dans l'histoire. Editions Gallimard, Paris

Bakens J, Nijkamp P (2013) Migrant heterogeneity and urban development: a conceptual analysis. In: Crescenzi R, Percoco M (eds) Geography, institutions and regional economic performance. Springer-Verlag Berlin Heidelberg

Berliant M ,M Fujita (2008) Knowledge creation as a square dance on the Hilbert Cube. International Economic Review 49(4):1251–1295

Bivand RS, Brunstad RJ (2003) Regional growth in Western Europe: An empirical exploration of interactions with agriculture and agricultural policy. In: Fingleton B (ed) European regional growth. Springer, New York, pp 351–375

Card D (2001) Immigrant inflows, native outflows, and the local labour market impacts of higher immigration. J Labour Econ XIX:22–64

Carrasco R, Jimeno JF, Ortega AC (2008) The effect of immigration on the labor market performance of native-born workers: some evidence for Spain. J Popul Econ 21:627–648

Ciccone A (2002) Agglomeration effects in Europe. Eur Econ Rev 46:213–227

Ciccone A, Hall R (1996) Productivity and the density of economic activity. Am Econ Rev 87:54–70

Collier P (2001) Implication of ethnic diversity. Econ Policy 32:129–166

Crescenzi R, Gagliardi L, Percoco M (2013) Social capital and innovation in Italy. In: Crescenzi R, Percoco M (eds) Geography, institutions and regional economic performance. Springer-Verlag Berlin Heidelberg

D'Amuri F, Ottaviano G, Peri G (2010) The labor market impact of immigration in Western Germany in the 1990s. European Economic Review 54(4):550-570

D'Costa S, Gibbons S, Overman H, Pelkonen P (2013) Agglomeration and labour markets: the impact of transport investments on labour market outcomes. In: Crescenzi R, Percoco M (eds) Geography, institutions and regional economic performance. Springer-Verlag Berlin Heidelberg

Easterly W (2001) Can institutions resolve ethnic conflict? Econ Dev and Cult Change 49 (4):687–706

Easterly W, Levine R (1997) Africa's growth tragedy: policies and ethnic division. Q J Econ 111 (4):1203–1250

Florida R (2002) The rise of the creative class (tr. it., L'ascesa della nuova classe creativa, Milano, ed. Arnoldo Mondadori Editore SpA)

Gertler MS, Florida R, Gates G, Vinodrai T (2002) Competiting on creativity: placing ontario's cities in North American context, institute for competitiveness and prosperity, Ontario Ministry of Enterprise

Glaeser EL, Scheinkman JA, Shleifer A (1995) Economic growth in a cross-section of cities. J Monet Econ 36(1):117–144

Jacobs J (1961) The death and life of great American cities. Vintage, New York

Knack S, Keefer P (1995) Institutions and economic performance: cross-country tests using alternative institutional measures. Econ Polit 7:207–227

La Porta R, Lopez de Silanes F, Shleifer A, Vishny R (1999) The quality of government. J Law Econ Organ 15(1):222–279

Lazear EP (1999) Globalisation and the market for team-mates. Econ J 109:C15–C40

Manacorda M, Manning A, Wadsworth J (2012) The impact of immigration on the structure of wages: theory and evidence from Britain. Journal of the European Economic Association 10 (1):120–151

Maré D, Fabling R (2013) Productivity and local workforce composition. In: Crescenzi R, Percoco M (eds) Geography, institutions and regional economic performance. Springer-Verlag Berlin Heidelberg

Ottaviano GIP, Pinelli D (2007) Measuring diversity: a cross-disciplinary comparison of existing indices. Ph.D. thesis, University of Bologna

Ottaviano GIP, Peri G (2005) Cities and cultures. J Urban Econ 58:304–337

Ottaviano GIP, Peri G (2006) The economic value of cultural diversity: evidence from US cities. J Econ Geogr 6:9–44

Ottaviano GIP, Pinelli D (2006) Market potential and productivity: evidence from Finnish regions. Reg Sci Urban Econ 36:636–657

Ottaviano GIP, Prarolo G (2009) Cultural identity and knowledge creation in cosmopolitan cities. J Reg Sci 49(4)

Redding S, Venables A (2004) Economic geography and international inequality. J Int Econ 62:53–82

Roback J (1982) Wages, rents and the quality of life. J Polit Econ 90:1275–1278

Saiz A (2003) Immigration and housing rents in American cities. Federal Reserve Bank of Philadelphia, Working paper 03–12

Sassen S (1994) Cities in a world economy. Pine Forge Press, Thousand Oaks

Schündeln M (2007) Are immigrants more mobile than natives? Evidence from Germany, IZA Discussion papers 3226

Temple J (1999) The new growth evidence. J Econ Lit 37:112–156

Temple J (2001) Generalization that aren't, evidence on education and growth. Eur Econ Rev 45:905–918

Williams, KY, O'Reilly CA (1998) Demography and Diversity in Organizations: A Review of 40 Years of Research. In: Staw BM, Cummings LL (eds) Research in Organizational Behavior, vol. 20, JAI Press, Greenwich, CT, pp 77–140

The ‘Bright’ Side of Social Capital: How ‘Bridging’ Makes Italian Provinces More Innovative

Riccardo Crescenzi, Luisa Gagliardi, and Marco Percoco

Abstract Existing studies have failed to reach a consensus on the impact of social capital on local innovative performance: some empirical analyses emphasize a positive effect while others speak about a ‘dark side’ of social capital. This chapter aims to shed new light on the differential role of ‘bonding’ and ‘bridging’ social capital in innovation dynamics. The spatial economic analysis of the innovative performance of the Italian provinces suggests that social capital is an important predictor of innovative performance. However, only ‘bridging’ social capital-based on weak ties-can be identified as a relevant driver of the process of innovation while ‘bonding’ social capital is shown to be non-significant for innovation.

Keywords Innovation • Social Capital • Knowledge Transfer • Regional Development

1 Introduction

Social capital as determinant of a successful economic outcome has received a significant attention in the past decades gaining wide acceptance in the economic literature. Standard economic theories have largely failed to explain the persistence of economic differentials among countries and regions, stimulating an in-depth analysis of “soft” factors as complementary and fundamental ingredients for growth

---

JEL classification: O31, O33, R15

R. Crescenzi (✉) • L. Gagliardi
Department of Geography and Environment & SERC, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom
e-mail: r.crescenzi@lse.ac.uk, l.gagliardi@lse.ac.uk

M. Percoco
Department of Policy Analysis and Public Management and CERTeT, Università Bocconi, Via R. Sarfatti, 25 20100 Milan, Italy
e-mail: marco.percoco@unibocconi.it

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_8,

and development (Banerjee and Duflo 2005; Bellini et al. in chapter "Cultural Diversity and Economic Performance: Evidence from European Regions" and Bauernschuster et al. in chapter "Explicitly Implicit: How Institutional Differences Influence Entrepreneurship" of this Book). In this context, social capital has emerged as an important explanation for a wide range of phenomena: from economic growth (Knack and Keefer 1997) to political participation (Di Pasquale and Glaeser 1999), development trap (Woolcock 1998), institutional performance (La Porta et al. 1997) or the spread of secondary education (Goldin and Katz 1999).

However, the analysis of the link between social capital and the generation of innovation – in its turn a crucially important driver of economic growth – has remained relatively unexplored by 'mainstream' economic literature. Economists of innovation and economic geographers have recently tried to fill this gap in the understanding of the impact of social capital on economic performance opening the way to new insights into the mechanics of social capital in the economy (Cohen and Fields 2000; Hauser, et al. 2007; Kallio et al. 2009; Laursen and Masciarelli 2007; Patton and Kenney 2003; Sabatini 2009; Tura and Harmaakorpi 2005).

In particular, this stream of literature has contributed to a better conceptualisation of 'social capital' shedding light on its 'multidimensionality' and suggesting that different dimensions may impact upon the economy in very different ways. It is the intensity and typology of network relations among innovative actors that matters for innovation. The characteristics of such networks and the social incentives for their formation qualify the way in which valuable knowledge is exchanged and re-combined linking together individuals, groups and geographical areas (Audretsch and Feldman 2004) further stimulating relational proximity and preventing stagnation and lock in (Boschma 2005). In this context the traditional debate regarding the optimal level of social capital seems to be less pertinent: the effectiveness of social capital doesn't lie in the density of relationships within the network but in their intensity and in the extension of their "radius of trust" (Fukuyama 1995). The wider is the radius of trust of the network relationships among knowledgeable individuals the greater the likelihood of complementary knowledge exchange. This, in turn, implies that, in an innovation enhancing perspective, the potential negative role of social capital is mainly related to the existence of closed networks that lead to the exchange of redundant knowledge.

This chapter aims to explore the nexus between social capital and innovation by looking at the dichotomy between bonding and bridging social capital, where the first is based on strong ties and closed networks reinforced by deep emotional involvement and the second is, instead, related to weak ties linking otherwise disconnected communities. The empirical analysis looks at Italian provinces, one of the most intensively studied cases in the literature on social capital (Guiso et al. 2004; Ichino and Maggi 2000; Putnam 1993) but – to the best of our knowledge – largely unexplored in terms of the link between social capital and innovation. The existing empirical literature has made reference to a broad conceptualisation of social capital (associational activities, political participation, institutional thickness and trust) with a limited attempt to clarify the channels through which it affects the innovative performance of regions. Conversely, this chapter aims to develop a coherent definition of social capital by explicitly addressing the mechanisms behind its effect on innovation. In doing this – and in line with the conclusion reached by Hauser et al. (2007) – the chapter is focused on the network dimension of social capital, and provides some empirical evidence on its spatial patterns and its association with innovation in Italy.

The results suggest that social capital is a fundamental driver of innovation in Italian provinces if and only if it operates as a channel for the exchange of non redundant and complementary knowledge.

The chapter is organized as follows: we first provide and overview of the economic literature on the role of social capital, highlighting the specific meaning of the term with reference to innovation. In Sect. 3 we discuss the methodology and the dataset. Section 4 presents some descriptive statistics, the spatial analysis and the main results. Finally some conclusions are drawn underlining the fundamental role of social capital as a determinant of local innovative performance.

2 Social Capital as a Determinant of Innovation

The aim of this section is to look at the vast literature on the economic impact of social capital in order to develop a suitable ‘working definition’ and an appropriate conceptual framework for its analysis in relation to the process of innovation.[1]

A fundamental vagueness is still characterizing the definition of social capital (Guiso et al. 2010). Coleman (1988) argued that it coincides with the social structure of a society facilitating the actions of individuals. Putnam (1993) identified social capital in terms of trust-based relations and groups. Fukuyama (1995) suggested that social capital has to be intended in terms of trust, civicness and network relations. However, none of the above definitions has made it possible neither to develop a comprehensive measure of social capital nor to overcome the traditional debate on its potential ‘dark side’ (i.e. the low innovative dynamism of some high social capital regions as emphasized by Florida 2002). From the methodological perspective, several difficulties exist in the operationalisation of the concept. As Solow (1999) emphasized in his critique to Fukuyama (1995): if social capital is something more than a fuzzy concept it has to be somehow measurable. However, we are still far from dealing with a universal measure of social capital. Different aspects were alternatively emphasized and different measures were proposed: from civic cooperation to collective action, from trust to political participation, groups and networking. The analysis of the link between social capital and innovation calls for a more rigorous definition of the term in order to single out the channels through which social capital may potentially affect innovation.

The qualification of social capital with respect to local innovative performance builds on the so called “relational turn” of economic geography (Boggs and Rantisi 2003) and challenges the under-socialized nature of the past approach to innovation

[1] This section heavily relies on Crescenzi et al. (2012)

that systematically neglected the social dimension of innovation processes. This drawback becomes apparent when looking at the traditional mainstream economic theory of innovation based on firm-level knowledge production function approach (Griliches 1979) in an a-spatial and atomistic fashion.

The re-discovery of the concept of social capital as a fundamental determinant of innovation followed the theoretical contributions of Granovetter (1985) and Coleman (1988). Innovation started to be progressively considered as a social process embedded in the local social environment and systematically affected by the strength and the intensity of social ties. Regions can be seen as systems of relations located within certain geographical contexts in which different economic actors are systematically engaged in interactive processes of collective learning (see chapter "Firm Capabilities and Cooperation for Innovation: Evidence from the UK Regions" by Iammarino et al. in this Book; Cooke and Morgan 1998; Kostiainen 2002).

The emphasis on the social dimension of innovation led to the definition of innovation prone regions (Rodriguez-Pose 1999), social filters (Rodríguez-Pose and Crescenzi 2008; Rodriguez-Pose and Comptour in chapter "Evaluating the role of clusters for innovation and growth in Europe" of this book) innovative milieux (Breschi and Lissoni 2001; Camagni 1995), learning regions (Florida 1995; Morgan 1997), regional systems of innovation (Cooke et al. 1997). In all these cases the focus is on the network dimension, supposed to be able to foster innovative capabilities facilitating the diffusion of valuable and non redundant knowledge and preventing stagnation and lock in (Boschma 2005).

According to the aforementioned literature, the link between social capital and innovation lies exactly in the concepts of networking and embeddedness (Granovetter 1985). Relational networks linking together individuals, groups, firms, industries with different knowledge bases are a critical precondition for knowledge creation and transfer. In this context innovation is emerging from a cumulative process embedded in the social context and systematically affected by dynamics of interactive learning, stimulating the exchange and re-combination of knowledge (Asheim 1999; Lundvall 1992). Moreover social structures, in particular in the form of social networks, systematically affect innovative outcomes since they determine the flow and quality of information exchanged (Granovetter 2005).

Social capital is then a crucial factor for community development since it stimulates inter-personal interactions and the circulation of valuable knowledge (Tura and Harmaakorpi 2005). If we accept this simple statement, then social capital can be thought to be an input into an ideal knowledge production function.

However, the idea of "relations as central units of analysis" (Boggs and Rantisi 2003) is still questionable. Significant criticisms are associated to the existence of robust empirical evidence in support to this preponderant role of relations and untraded interdependences (Markusen 1999; Overman 2004). This shortfall becomes even more relevant when looking at the mechanisms driving this potential effect. Capello and Faggian (2005) emphasized the role of relational capital as crucial ingredient in the creation and diffusion of innovation looking at knowledge spillovers as key transmission channels to account for the effect of networking and social relations on innovative performance. Kallio et al. (2009) suggested that the link between the social dimension and the emergence of an innovative outcome lies in the local absorptive capacity enabling the diffusion of knowledge within the regional system of innovation. Other authors argued that social capital has only a second order effect and that it is mediated by the increasing returns on the investments in human (Bourdieu 1986; Dakhli and De Clercq 2004; Gradstein and Justman 2000) or physical capital (Becker and Diez 2004; Cainelli et al. 2005; Fritsch and Franke 2004). Conversely, this chapter investigates at the effect of social capital on innovation by looking at the innovative potential of network exchanges (Hauser et al. 2007): the characteristics of these networks clarify the mechanisms underlying the impact of social capital on innovation. Innovation is more likely to be found “in the structural holes between dense social networks” (Burt 2004; Granovetter 2005). By looking at social capital as a fundamental component of the socio-institutional environment shaping the process of innovation, this chapter contends that differences in the nature of social networks, rather than the density of their linkages, offer a potential explanation for the non-linear relation between social capital and innovation (Hauser et al. 2007).

The so called “weak ties hypothesis” proposed by Granovetter (1973) is crucial in this context. Relationships between people can be characterized by either frequent contacts and deep emotional involvement or sporadic interactions with low emotional commitment. The former category is generally identified as ‘strong ties’ – such as the relationships within families or close friends – while the latter is associated with the definition of ‘weak ties’ linking individuals characterized by loose acquaintances. Contextualising Granovetter’s argument into the analysis of innovation, ‘weak ties’ can be seen as the source of novel information and responsible for the diffusion of ideas (Granovetter 1982; Rogers 2005), while ‘strong ties’ increase the risk of exchanging redundant knowledge simply because they connect knowledge seekers with other individuals that are more likely to deal with ‘known’/ familiar information and knowledge (Levin and Cross 2004).

In other words, weak ties are fundamental in spreading information because they operate as a bridge between otherwise disconnected social groups (Ruef 2002). Weak ties serve as a bridging mechanism between communities within the same society, while strong ties function as a bonding device within homogeneous groups potentially hampering the degree of sociability outside the closed social circle (Beugelsdijk and Smulders 2003). Bonding social capital (Rodriguez-Pose and Storper 2006; Storper 2005) is likely to affect negatively innovation because it may work in favour of small groups lobbying for preferential policies and protection of the status quo hampering risky, innovative activities (Dakhli and De Clercq 2004; Knack and Keefer 1997; Portes and Landolt 1996). Conversely, bridging social capital, by lowering transaction costs, may contribute to the building of an environment congenial for innovation investment. As effectively pointed out by Putnam, the primary use of bonding social capital is to ‘get by’, while that of bridging social capital is to ‘get ahead’, implying that an over reliance on bonding social capital can generate a disincentive in creating connections “outside one’s own immediate network or social circle and into new areas of information and opportunity” (Cooke et al. 2005).

This implies that the 'dark side of social capital' lies in the typology of the ties and in the radius of trust of the network rather then in the intensity of the relationships among knowledgeable individuals: we need to look for the 'right' typology, rather than for the optimal 'quantity' of social capital if we aim to enhance local innovative performance.

In this chapter we focus on the relevance of social capital for the production of innovation in Italian provinces. The case of Italy is of potential interest because of both the considerable spatial variation in development and cultural traits and the availability of a large body of specific literature.

Putnam (1993) has in fact proposed the hypothesis that one of the main reasons for the persisting differences in development between North and South of Italy is due to the quality of institutions and social capital which in turn are the outcomes of historical accidents, i.e. areas in which independent city-states (the so-called *Repubbliche Comunali*) were more diffused are also the areas in which the level of trust and government effectiveness are higher. Recently, Guiso et al. (2008) and Percoco (2010a, b) have provided empirical support to this idea, although their main focus was on the explanation of income and productivity levels. In a similar context, Guiso et al. (2004) found a positive association between industrial development and social capital. Similarly, De Blasio and Nuzzo (2010), using microdata from the Survey of Household Income and Wealth conducted by the Bank of Italy, report that social capital increases the probability of being an entrepreneur.

The existing literature explicitly addressing the link between social capital and innovation is more limited. Some recent studies are largely qualitative (Ramella and Trigilia 2009). There are few recent papers applying a quantitative methodology to the analysis of the link between social capital and innovation in Italy. Some of them (Cainelli et al. 2005) looks at peculiar geographic areas such as the industrial districts arguing that the extensive horizontal relationships among local economic actors generate positive network externalities favouring the exchange of valuable knowledge and fostering the innovative performance of local firms. Others (Arrighetti and Lasagni 2010) adopt a firm based perspective in order to address the role of social and institutional factors on the probability of firms to innovate and their willingness to invest financial resources in innovation related activities. By analysing the effect of social conditions on the propensity to innovate of Italian firms they suggest that innovative firms tend to cluster in provinces characterized by relatively higher levels of "positive social capital" – interpreted as civicness and high social interactions – and lower levels of "negative social capital" generally associated with opportunistic behaviour due to the coexistence of groups lobbying for specific interests.

3 Methodology and Sources of Data

Our empirical analysis is based on the Knowledge Production Function (KPF), formalised by Griliches (1979, 1986) and Jaffe (1986). However, this chapter adopts a place based perspective and is focused on Italian provinces (NUTS

3 level) as units of observation. This specification, building on previous research in the field (Audretsch 2003; Audretsch and Feldman 1996; Crescenzi et al. 2007, 2012; Feldman 1994; Fritsch 2002; O’hUallachain and Leslie 2007; Ponds et al. 2010; Varga 1998), is particularly coherent with the main purpose of our analysis because it allows us to focus upon the territorial dynamics of innovation introducing social capital as a determinant of regional innovative performance.

The modified Knowledge Production Function takes the following form:

$$Patents\_growth_{iT-t} = \beta_0 + \beta_1 patent_{i,t-T} + \beta_2 soccap_{i,t-T} + \beta_3 grad_{i,t-T} + \beta_4 privrd_{i,t-T} + \beta_5 X_{i,t-T} + \delta_i + \varepsilon_i$$

Where $Patents\_growth_{i,T-t} = \frac{1}{T}\ln\left(\frac{Patents_{i,t}}{Patents_{i,t-T}}\right)$ is the logarithmic transformation of the ratio of patent applications in region i at the two extremes of the period of analysis (t-T,t). Among the independent variables $soccap_{i,t-T}$ is our variable of interest and represents the measure(s) of social capital in each province i at time (t-T); $patents_{i,t-T}$ is the log of the number of patents per million inhabitants at the beginning of the period of analysis (t-T); $privrd_{i,t-T}$ is private expenditure in R&D as percentage of regional GDP at (t-T); $grad_{i,t-T}$ is the number of graduates in respect to regional population at time (t-T); $X_{i,t-T}$ is the matrix of additional controls (i.e. regional sectoral composition, population density and female unemployment) at (t-T); Finally, δ_i represents macro-regional dummies for southern, central and northern Italy and ε_i is the error term. A detailed description of the main variables is reported in Table A.1 in Appendix.

Regional Innovative Performance – Patents data coming from OECD are used as a proxy for innovation. We construct our measures of innovation using the log transformation of the growth rate of patents in the time interval 2001–2007. Patent statistics can be considered a good measure of innovative output providing comparable information on inventors across a broad range of technological sectors. The main limitations of this measure are the differentiated propensity to patent of different sectors and the non-patentability of many inventions (Crescenzi et al. 2007). In fact, differences in the number of patents among provinces may be an indicator for differences in industrial specialization. If sectors differ structurally in terms of propensity to innovate or to patent, then those differences will be reflected into differentials in terms of number of patents (or their growth). To overcome this limitation, in our empirical approach, we will control for the sector structure of the economy.

Initial patent intensity – The initial patent intensity in each province is used as a proxy of the existing technological capabilities and the distance from the technological frontier. It also controls for differences in the patenting propensity often related to pre-existent differences in sector specialization.

Social Capital – Building on our conceptual framework we look at social capital emphasizing the component related to the networking activity, but trying to distinguish such networks with respect to their effect on the circulation of information.

Table 1 Principal components analysis (PCA)

	PCA: social capital	PCA: bonding	PCA: bridging
	(1)	(2)	(3)
	PC1	PC1	PC1
Eigenvalues	2.33353	1.078	1.52095
% of explained variance	0.5834	0.5390	0.7605
Variables			
Blood donations	0.5429		0.7071
Voluntary associations	0.5688		0.7071
Weekly lunch	−0.0663	0.7071	
Adult children	−0.6143	0.7071	

Note: Only principal components with eigenvalues > 1 are retained

As previously mentioned this implies a crucial distinction between networks based on weak ties, or bridging social capital, and networks based on strong ties, or bonding social capital.

We measure social capital by means of several variables in order to take into account both its bridging and bonding dimensions. Subsequently, through a principal component analysis, these variables are combined into a composite measure of social capital (Table 1, Column 1) and into two additional separate measures one for bonding and one for bridging social capital respectively (Table 1, Columns 2 and 3). Due to the characteristics of social capital in Italy and its specific spatial pattern, the composite indicator constructed through principal component analysis provides a preliminary evidence of the dichotomy between the bonding and bridging dimensions. In order to develop a deeper analysis of this evidence the composite indicator is then sub-divided into its bridging and bonding components.

Data on family characteristics are used as proxies for bonding social capital based on strong ties (Beugelsdijk and Smulders 2003; Levin and Cross 2004; Ruef 2002) and data on voluntary associations as a proxy for bridging social capital based on weak ties operating as forms of horizontal relations, fostering networks of civic engagement (Arrighetti and Lasagni 2010; Beugelsdijk and van Schaiik 2005). In order to capture the strength of family ties two indicators are included in the analysis: the number of families having lunch at least once per week with relatives and close friends (per 100 households) and the number of young adult individuals living with their parents (per 100 young adults).

Strong family ties are assumed to imply geographical proximity of adult children: young adults tend to stay longer with parents and the relationships within the family are particularly strong and based on repeated interactions. Family members tend to gravitate around the main core creating a system of nested families and a larger family size (Alesina and Giuliano 2010).

At this point, it should be mentioned that the characteristics of the family are at the heart of the hypothesis on the importance of social capital in Italian development since the very seminal work by Banfield (1958) who advanced the idea that low propensity to cooperate is generally associated to, among other things, the strength of family ties. In particular, Banfield (1958) argued that underdevelopment is a result of a low propensity to cooperate which, in turn, produces high transaction costs. This development trap is the outcome of strong family ties (the so-called “amoral familism”), high uncertainty and a highly unequal distribution of income and wealth. So far, we do not have conclusive empirical evidence supporting Banfield’s hypothesis, however, some pieces of evidence seem to confirm at least partially this theory. Alesina and Giuliano (2010), in fact, find that strong family ties are associated to low levels of generalized trust. Similarly, Giavazzi et al. (2010) relate family types to female labor market participation rate in European regions, whereas Duranton et al. (2009) relate past family structures to a number of contemporary outcomes.

Bridging social capital based on weak ties is instead measured using two of the traditional indicators adopted in the economic literature as proxies for social capital. Blood donations and participation in voluntary associations are assumed to be proxies for the participation of individual in activities with positive social externalities and as an indicator for altruism (Cartocci 2007).

The number of families having lunch at least once per week with relatives[2] and the number of young adults living with parents[3] are used to define a composite indicator of bonding social capital while blood donations and voluntarism concur to define the composite indicator for bridging social capital. We further defined a comprehensive measure of social capital encompassing both the bonding and bridging dimension that is used in the first stage of the analysis in order to detect the overall effect of social capital on innovation before going into details.

We finally included a spatial lag of our composite measure of social capital in order to control for potential spillovers effect. All the spatially lagged variables are constructed based on a standard queen contiguity spatial weighting matrix.

Innovation inputs – Private R&D as a share of regional GDP and the number of graduates over the total population are used as proxies for the key inputs of the ‘standard’ regional Knowledge Production Function. On account of limited data availability our R&D measure is available only at regional level (NUTS 2) while the number of graduates is available for each province (NUTS3).

Controls – Our specification of the knowledge production function includes controls for population density at province level, labour market characteristics in terms of female unemployment rate and sector structure approximated by the Herfindhal Index.

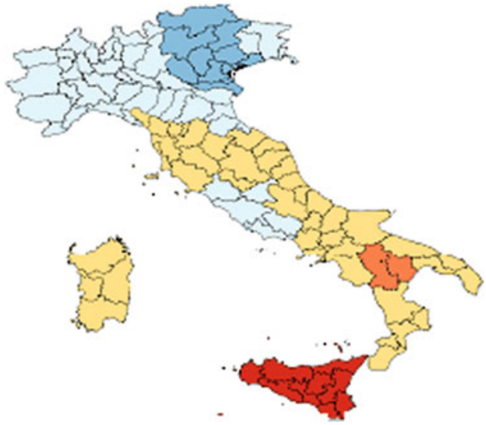
The Herfindhal Index is defined using data on employment for three sectors: agriculture, industry and services and it is interpreted as a measure of specialization.

We further add some controls to take into account spatial correlation. In particular we defined the spatial lag of population density as a measure of accessibility. Macroregional dummies are inserted to control for time invariant characteristics and other sources of spatial correlation.

[2] Per 100 families

[3] Per 100 young adults

Fig. 1 Bonding social capital



4 Empirical Results

Preliminary evidence on the potential link between the innovative performance of Italian provinces and the characteristics of the local social environment can be analyzed by looking at the descriptive statistics and the correlation between the relevant variables.

The principal components analysis (Table 1) shows that our composite measure of social capital-jointly accounting for both the bonding and bridging dimensions of the concept-attributes opposite signs to the two components. Variables used as proxies for bridging social capital enter the composite indicator with a positive weight, while the variables used as proxy for bonding social capital show a negative sign. This implies that our composite measure of social capital explicitly takes into account the characteristics of the Italian context, characterized by a predominance of strong ties (interpreted here as family ties) in Southern regions and a higher level of bridging social capital and community involvement in Central and Northern regions.

In order to clarify how this dichotomy in the characteristics of the social structure between Northern and Southern regions affects local innovative dynamism, the composite measure of social capital is sub-divided in its two main components: bonding and bridging social capital respectively (Table 1).

The well known North–south dichotomy in Italy is clearly reflected into the characteristics of the social environment. Figures 1 and 2 visualise the spatial distribution of bridging and bonding social capital in Italian provinces. Bridging social capital seems to be systematically higher in Northern Italy and in part of the Central regions while Southern provinces are characterized by the predominance of bonding social capital. On the other side, the geographical distribution of innovation (Fig. 3) seems to be very similar to the distribution of bridging social capital providing some preliminary support in favour of our hypothesis regarding the crucial role of weak ties as complementary preconditions for innovation.

Fig. 2 Bridging social capital

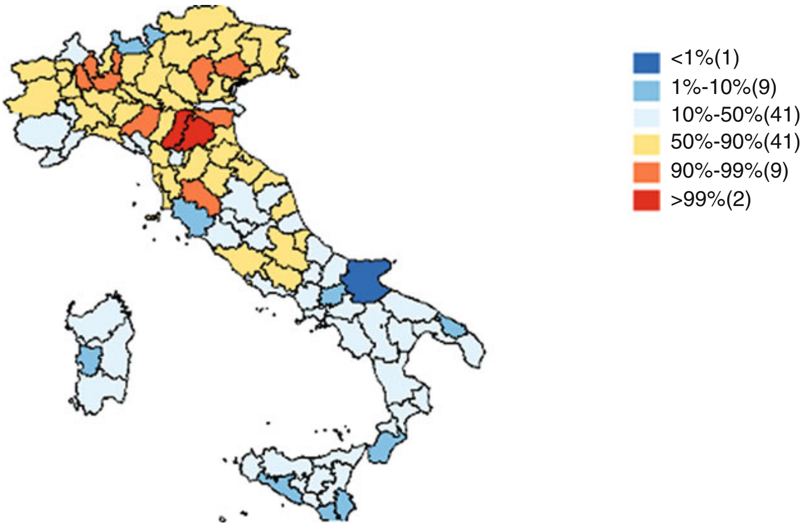
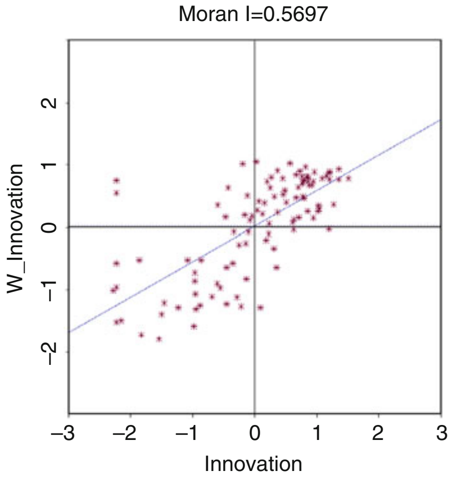


Fig. 3 Innovation

A more in-depth analysis of the spatial structure of both bonding and bridging social capital and innovation is necessary in order to clarify this point. Figures 4, 5, and 6 report the Moran’s I spatial correlation index for innovation, bonding and bridging social capital respectively. Bonding and bridging social capital shows a clear pattern of spatial concentration. For innovation the magnitude of the coefficient is lower, but there is still evidence of a significant level of spatial concentration.

Deepening the analysis of the spatial patterns of social capital in Italy, the Local Spatial Autocorrelation Index (LISA) is reported and it generally supports the hypothesis regarding the spatial concentration of bonding and bridging social capital in Italian provinces.

Fig. 4 Moran's I innovation



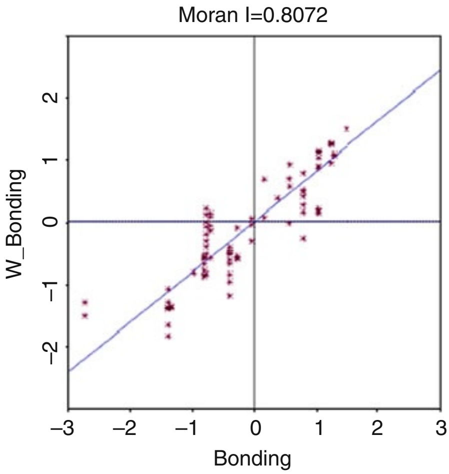


Fig. 5 Moran's I bonding social capital

Figures 7 and 8, reporting respectively the cluster map for bonding and bridging social capital, classify the areas with respect to the predominant typology of spatial correlation. Both bridging and bonding social capital are characterized by the predominance of the high-high and the low-low clusters with an opposite geographical distribution. Provinces with the highest level of bridging social capital tend to be concentrated in the North while those showing lowest scores are clustered in Southern Italy. Symmetrically, bonding social capital characterizes Southern provinces while areas showing the lowest level remain concentrated in the North. Conversely, intermediate clusters (high-low and low-high) do not show any clear pattern, confirming the strong spatial concentration of bonding and bridging social capital. The LISA index for innovation, reported in Fig. 9, shows a spatial pattern

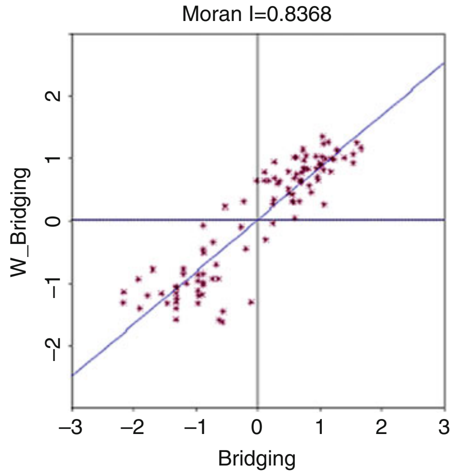


Fig. 6 Moran’s I bridging social capital

Fig. 7 LISA bonding social capital

similar to bridging social capital, with provinces characterised by the most dynamic innovative performance clustered in the North and those with the lowest scores concentrated in the South[4].

[4] Note that two provinces (Sondrio and Pistoia) despite being located in highly innovative areas show low innovative performance probably due to their agricultural vocation while one province (Enna) in spite of being located in a low innovative area is characterized by a good performance. This last case can be explained by the localization in this territory of the high tech cluster of the “Etna Valley”.

Fig. 8 LISA bridging social capital

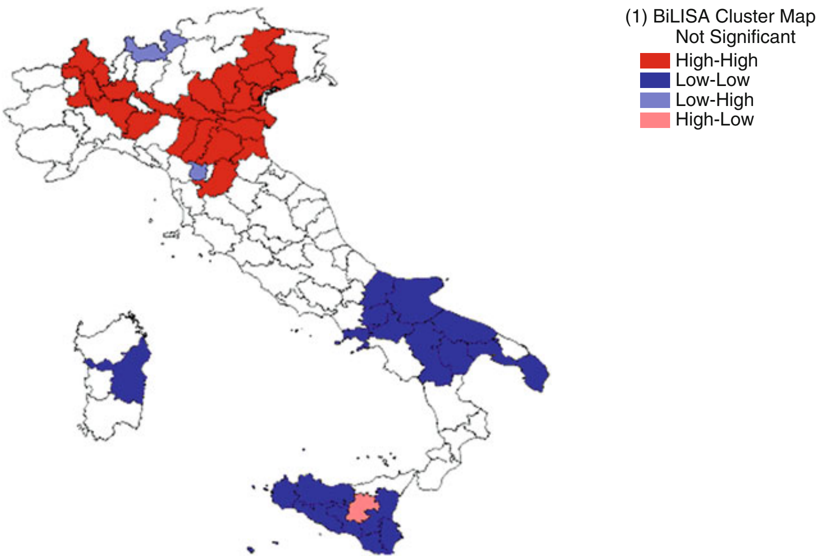


Fig. 9 LISA innovation

This preliminary evidence in support of a similar spatial concentration pattern between innovation and bridging social capital and a complementary negative spatial correlation between innovation and bonding social capital seems to be further confirmed by the multivariate LISA shown in Figs. 10 and 11 respectively. Figure 10 is characterized by the predominance of the high-high and the low-low cluster suggesting that innovation is spatially concentrated in areas characterized by

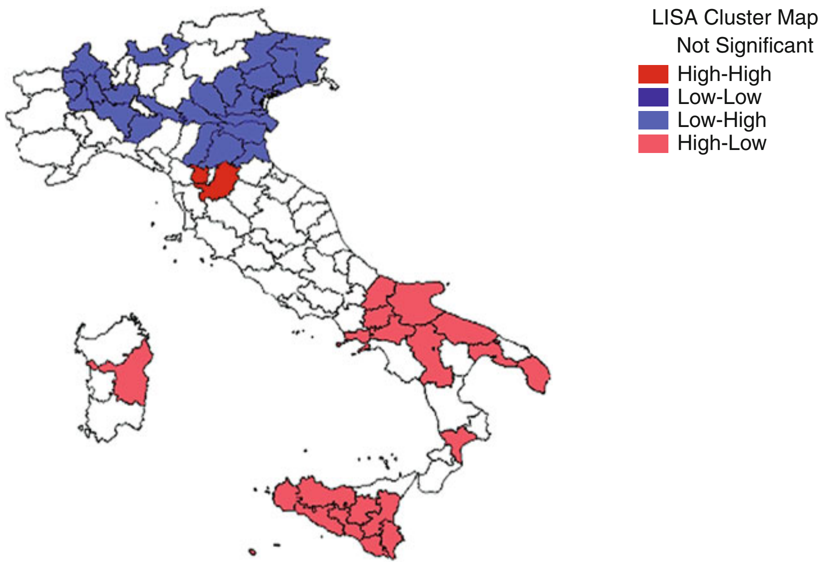


Fig. 10 Multivariate LISA innovation-bonding

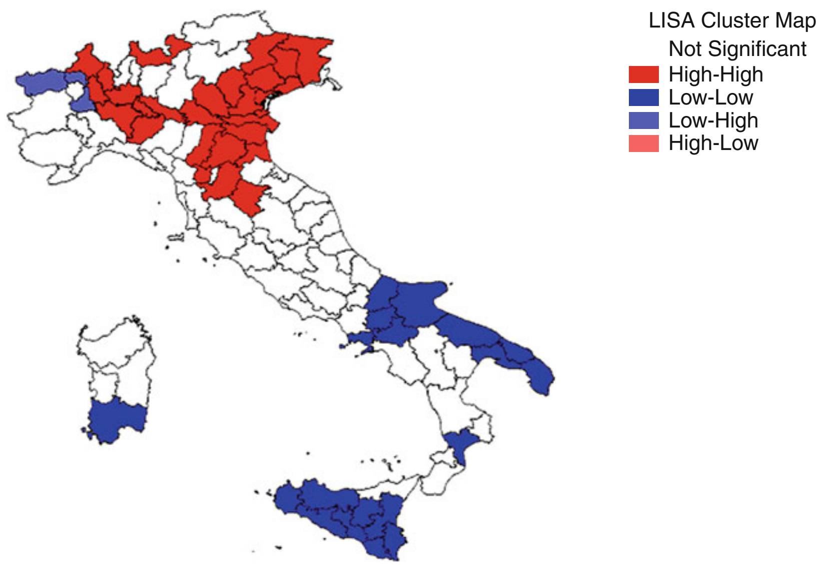


Fig. 11 Multivariate LISA innovation-bridging

higher levels of bridging social capital. On the contrary, Fig. 11 shows a predominance of the high-low and low-high clusters supporting the hypothesis of a divergent spatial distribution of innovation with respect to bonding social capital.

The evidence regarding the sign and the magnitude of the correlation between innovation and bonding/bridging social capital deserves a deeper analysis.

Table 2 Descriptive statistics

Macroregion	Variable	Obs	Mean	Std. Dev.	Min	Max
North	Patents growth	45	0.060	0.063	−0.114	0.238
	Bonding	45	−0.937	0.572	−2.838	−0.293
	Bridging	45	0.994	0.575	−0.651	2.054
Centre	Patents growth	24	0.058	0.068	−0.072	0.200
	Bonding	24	0.363	0.682	−0.748	1.063
	Bridging	24	−0.077	0.858	−1.478	1.318
South	Patents growth	28	0.039	0.141	−0.207	0.339
	Bonding	28	1.029	0.486	0.154	1.549
	Bridging	28	−1.492	0.616	−2.668	−0.236

Table 3 Correlation matrix

	Patents_growth	Bonding	Bridging
Patents_growth	1.0000		
Bonding	0.0079	1.0000	
Bridging	0.2079	−0.5990	1.0000

Data reported in Table 2 suggest a substantial congruence between local innovative outcome and the bridging component of social capital as well as a systematic negative association between innovation and bonding social capital. This descriptive evidence further supports the main hypothesis of our analysis: social capital has a beneficial effect on the innovative performance of local areas when it is based on the existence of weak ties between otherwise disconnected communities. Complementary, a strong predominance of bonding social capital is associated with lower innovative performance. The two sided effect of social capital on innovation is further confirmed by the correlation matrix reported in Table 3.

The systematic correlation between bridging social capital and innovative performance in Italian provinces supports the initial hypothesis of this chapter and calls for further investigation through standard econometric analysis.

Table 4 reports the estimation results for the place based Knowledge Production Function. In the basic version we just control for capital and labour and the initial level of patenting intensity in each region (Table 4 Col.1). The initial number of patents per million of inhabitants is statistically significant at 1 % level and negatively associated to our dependent variable. The sign of the coefficient can be justified through a convergence trend in patenting due to either the crisis of traditionally successful innovative areas (such as the industrial districts) or the emergence of new successful players.

Some controls for population density, the labour market characteristics, the sector structure and the spatial lag of population density (used as proxy for accessibility) are progressively included in the model (Table 4, column 2). Neither the level of female unemployment, used as proxy for the efficiency of the local labour market, nor the Herfindhal index, used as an indicator for sector specialization, are statistically significant. Conversely, population density seems to be positively associated to innovation with a significance level of 5 %. On the contrary the spatial lag of population density shows a significant negative effect at 10 % level.

Table 4 Estimation results

	(1)	(2)	(3)	(4)	(5)	(6)
Dep var: patents growth	OLS	OLS	OLS	OLS	OLS	OLS
Initial patent intensity	–0.0407***	–0.0396***	–0.0824***	–0.0830***	–0.0872***	–0.0828***
	(0.0104)	(0.0107)	(0.00924)	(0.00977)	(0.00970)	(0.0102)
Private R&D	0.0373***	0.0376***	0.0215**	0.0210**	0.0179	0.00991
	(0.00992)	(0.0104)	(0.00959)	(0.00879)	(0.0119)	(0.0108)
Graduates	0.0766	0.107*	0.0396	0.0406	0.0257	0.0199
	(0.0488)	(0.0528)	(0.0530)	(0.0542)	(0.0564)	(0.0561)
Female unempl.		0.0140	0.00649	0.00682	0.00912	0.00734
		(0.0137)	(0.0115)	(0.0110)	(0.0108)	(0.0114)
Herfindal index		–0.00216	–0.00139	–0.00128	–0.000655	–0.000553
		(0.00209)	(0.00202)	(0.00187)	(0.00196)	(0.00202)
Pop density		0.0311**	0.0324***	0.0329***	0.0326***	0.0303***
		(0.0146)	(0.0102)	(0.0108)	(0.0110)	(0.0104)
Spatial lag		–0.0385*	–0.0383**	–0.0396**	–0.0299**	–0.0231
Pop density		(0.0212)	(0.0147)	(0.0157)	(0.0135)	(0.0155)
Social capital			0.0472***	0.0413*	0.0467**	
			(0.00520)	(0.0211)	(0.0221)	
Spatial lag				0.00720	0.00321	
Social capital				(0.0231)	(0.0249)	
Nord					0.00612	0.0407
					(0.0394)	(0.0784)
Centro					0.0352	0.0612
					(0.0289)	(0.0424)
Bonding social capital						–0.00582
						(0.0148)
Bridging social capital						0.0477***
						(0.0119)
Constant	0.428**	0.599**	0.491**	0.499**	0.406*	0.319
	(0.165)	(0.229)	(0.192)	(0.200)	(0.228)	(0.214)
Observations	97	97	97	97	97	97
R-squared	0.181	0.253	0.456	0.457	0.474	0.461

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$
Robust standard errors in *parentheses*

In column 3 we control for our measure of social capital which is highly significant at 1 % level and positively correlated to innovation in each province. In the interpretation of this finding it is necessary to bear in mind the characteristics of our measure of social capital. The composite indicator constructed through principal components analysis already takes into account the characteristic Italian dichotomy between bridging and bonding operationalizing it through the attribution of a negative sign to bonding social capital and a positive sign to the bridging component. The positive sign associated to social capital in our regression suggests that the bridging component plays the crucial role and that provinces characterized by significant levels of cooperation and associational activities are more prone to innovation. In the model estimates reported in column 3 the spatial lag of social capital is included in order to control for potential neighbouring effects. The regressor is not statistically significant; however its inclusion affects the estimation of other coefficients suggesting the presence of spatial autocorrelation. In order to further control for this potential neighbourhood effect and spatial correlation macro-regional dummies[5] are included into the model. The measure of social capital remains positively associated to innovation with a significance level of 5 %.

In the final step of the analysis, the two fundamental components of social capital (i.e. bonding social capital based on strong ties, and bridging social capital based on weak ties) are assessed separately by splitting our aggregate measure of social capital into two separate regressors (Table 4, column 5). The bridging component remains statistically significant at 1 % level and positively associated to innovation while the bonding social capital is not statistically significant. These results confirm that the positive and significant impact of social capital on innovation is based on the mechanism of weak ties rather than strong ties. Weak ties make it possible to access non-redundant information, favouring the transfer and re-combination of valuable knowledge and fostering the innovative performance of Italian provinces. Conversely, bonding social capital, based on strong ties, is not a statistically significant determinant of innovation.

5 Conclusions

Soft factors – such as social capital – have gained progressive importance in the economic literature. This chapter, has focused on the link between innovation and social capital by looking at the networking and associational dimension of social capital and exploring the mechanisms for the diffusion and the circulation of valuable knowledge.

The effect of social capital on innovation is shaped by its capability to facilitate the exchange of complementary knowledge between individuals. This implies that networks and ties bridging individuals belonging to heterogeneous epistemic communities (as opposed to homogeneous like-minded groups), are conducive to innovation because they allow the access to non-redundant information.

Our results suggest that social capital based on weak ties is a fundamental determinant of innovation: it is the quality of social capital (and not its total endowment) that affects its correlation with innovation.

Further research is needed in order to deepen the understanding of the mechanisms driving the correlation between local social characteristics and innovative performance and to overcome the challenges related to the analysis of the causal link behind the effect of social capital on innovation.

[5] Moran's I over the residual is calculated in order to test for the existence of spatial correlation. Controlling for the spatial lag of population density and social capital and adding macroregional dummies the coefficient of the Moran I decrease, from 0.25 to 0.085, and becomes statistically insignificant. The p-value further confirms the rejection of the null of spatial correlation in the residuals.

Appendix

Table A.5 Variables list

Variables	Description		Source	Year
Patents growth	Logarithmic transformation of the ratio of patent applications per million inhabitants in region *i* at the two extremes of the period of analysis (t-T,t)		OECD RegPat database	2001–2007
Patents (level in 2001)	Logarithm of the level of patent applications per million inhabitants at the beginning of the period of analysis (t-T)		OECD RegPat database	2001
Private R&D	Logarithm of private expenditure in R&D as percentage of regional GDP at (t-T)		ISTAT Indicatori ricerca e innovazione	2001
Graduates	Logarithm of the number of graduates in over 24 population at time (t-T)		EUROSTAT Regional database	2001
Female unemployment	Logarithm of the number of unemployed women in total female labour force		OECD Regional database – regional labour market TL3 database	2001
Sectoral shares and Herfindal index	Sector employment/total employment ratio defined for agriculture, industry and services. Herfindal calculated as the sum of the square of these ratios.		OECD Regional database – regional labour market TL3 dataset	2001
Population density	Logarithm of the population in respect to local surface		OECD Regional database-demographic statistics TL3 dataset	2001
Social capital	BRIDGING	Blood donations (Number of blood donations per 100 residents)	Cartocci (2007)	2001
		Voluntary associations (Number of voluntary associations per Kmq)	Cartocci (2007)	2001
	BONDING	Weekly Lunch (Number of families having lunch at least once per week with relatives and close friends per 100 households)	ISTAT Rilevazione "Parentela e Reti di solidarietà"	2001
		Adult children (Number of young adult individuals living with parents per 100 young adults)	ISTAT Rilevazione "Parentela e Reti di solidarietà"	2001

References

Alesina A, Giuliano P (2010) The power of the family. J Econ Growth 15(2):93–125
Arrighetti A, Lasagni A (2010) Capitale sociale, contesto istituzionale e performance innovative delle imprese. Sci Reg 10:5–34
Asheim B (1999) Interactive learning and localised knowledge in globalising learning economies. GeoJournal 49:345–352
Audretsch DB (2003) Innovation and spatial externalities. Int Reg Sci Rev 26:167–174
Audretsch DB, Feldman MP (1996) R&D spillovers and the geography of innovation and production. Am Econ Rev 86:253–273
Audretsch DB, Feldman MP (2004) Knowledge spillovers and the geography of innovation. In: Henderson JV, Thisse J-F (eds) Handbook of regional and urban economics, vol 4. Elsevier, Princeton, pp 2713–2739
Banerjee AV, Duflo E (2005) Growth theory through the lens of development economics. In: Durlauf SN, Aghion P (eds) Handbook of economic growth, volume 1A. Elsevier Science, Amsterdam
Banfield EC (1958) The moral basis of a backward society. Free Press, Chicago
Becker W, Diez J (2004) R&D Cooperation and innovation activities of firms. Evidence for the German manufacturing industry. Res Policy 33:209–223
Beugelsdijk S, Smulders S (2003) Bridging and bonding social capital: which type is good for economic growth? Paper presented at European Regional Science Association, Jyvaskila
Beugelsdijk S, van Schaiik T (2005) Social capital and growth in European regions: an empirical test. Eur J Polit Econ 21:301–324
Boggs JS, Rantisi NM (2003) The ‘relational turn’ in economic geography. J Econ Geogr 3:109–116
Boschma RA (2005) Proximity and innovation: a critical assessment. Reg Stud 39(1):61–74
Bourdieu D (1986) The forms of capital. In: Richardson J (ed) Handbook of theory and research for the sociology of education. Greenwood, New York
Breschi S, Lissoni F (2001) Localised knowledge spillovers vs. innovative milieux: knowledge “tacitness” reconsidered. Pap Reg Sci 90:255–273
Burt RS (2004) Structural holes and good ideas. Am J Sociol 110:349–399
Cainelli G, Mancinelli S, Mazzanti M (2005) Social capital, R&D and industrial district, Feem working papers n 744584
Camagni RP (1995) The concept of “innovative milieu” and its relevance for public policies in European lagging regions. Pap Reg Sci 74:317–340
Capello R, Faggian A (2005) Collective learning and relational capital in local innovation processes. Reg Stud 39:75–87
Cartocci R (2007) Mappe del tesoro: atlante del capitale sociale in Italia. Il Mulino, Bologna
Cohen S, Fields G (2000) Social capital and capital gains: an examination of social capital in Silicon Valley. In: Kenney M (ed) Understanding Silicon Valley. Stanford University Press, Stanford
Coleman JS (1988) Social capital in the creation of human capital. Am J Sociol 94:S95–S120
Cooke P, Morgan K (1998) The associational economy: firms regions and innovation. Oxford Economic Press, Oxford
Cooke P, Uranga MG, Etxebarria G (1997) Regional innovation systems: institutional and organisational dimensions. Res Policy 26:475–491
Cooke P, Clifton N, Oleaga M (2005) Social capital, firm embeddedness and regional development. Reg Stud 39(8):1065–1077
Crescenzi R, Rodriguez-Pose A, Storper M (2007) The territorial dynamics of innovation: a Europe–United states comparative analysis. J Econ Geogr 7:673–709
Crescenzi R, Rodríguez-Pose A, Storper M (2012) The territorial dynamics of innovation in China and India. J Econ Geogr (forthcoming)

Dakhli M, De Clercq D (2004) Human capital, social capital and innovation: a multi-country study. Entrep Reg Dev 16(2):107–128
De Blasio G, Nuzzo G (2010) Historical traditions of civicness and local economic development. J Reg Sci 50(4):833–857
Di Pasquale D, Glaeser E (1999) Incentives and social capital: are homeowners better citizens? J Urban Econ 45(2):354–384
Duranton G, Rodriguez-Pose A, Sandall R (2009) Family types and the persistence of regional disparities in Europe. Econ Geogr 85(1):23–47
Feldman MP (1994) The geography of innovation. Kluwer, Boston
Florida R (1995) Towards the learning region. Futures 27(5):527–536
Florida R (2002) The rise of the creative class: and how it’s transforming work, leisure, community, and everyday life. Basic Books, New York
Fritsch M (2002) Measuring the quality of regional innovation systems: a knowledge production function approach. Int Reg Sci Rev 25:86–101
Fritsch M, Franke G (2004) Innovation, regional knowledge spillovers and R&D cooperation. Res Policy 33:245–255
Fukuyama F (1995) Trust: the social virtues and the creation of prosperity. Free Press, New York
Giavazzi F, Schiantarelli F, Serafinelli M (2010) Attitudes, policies and work, mimeo
Goldin C, Katz L (1999) Human capital and social capital: the rise of secondary schooling in America, 1910–1940. J Interdiscipl Hist 29:683–723
Gradstein M, Justman M (2000) Human capital social capital and public schooling. Eur Econ Rev 44:879–891
Granovetter MS (1973) The strength of weak ties. Am J Sociol 78:1360–1380
Granovetter M (1982) The strength of weak ties: a network theory revisited. In: Marsden PV, Lin N (eds) Social structure and network analysis. Sage, Beverly Hills, pp 105–130
Granovetter M (1985) Economic action and social structure: the problem of embeddedness. Am J Sociol 91(2):481–510
Granovetter M (2005) The impact of social structure on economic outcomes. J Econ Perspect 19 (1):33–50
Griliches Z (1979) Issues in assessing the contribution of research and development to productivity growth. Bell J Econ 10:92–116
Griliches Z (1986) Productivity, R & D, and the basic research at the firm level in the 1970s. Am Econ Rev 76:141–154
Guiso L, Sapienza P, Zingales L (2004) The role of social capital in financial development. Am Econ Rev 94(3):526–556
Guiso L, Sapienza P, Zingales L (2008) Long term persistence, mimeo
Guiso L, Sapienza P, Zingales L (2010) Civil capital as the missing link, NBER Working Paper 15845
Hauser C, Tappeiner G, Walde J (2007) The leraning region: the impact of social capital and weak ties on innovation. Reg Stud 41(1):75–88
Ichino A, Maggi G (2000) Work environment and individual background: explaining regional shirking differentials in a large Italian firm. Q J Econ 115:1057–1090
Jaffe A (1986) Technological opportunity and spillovers of R&D: evidence from firms’ patents, profits and market value. Am Econ Rev 76:984–1001
Kallio A, Harmaakoorpi V, Pihkala T (2009) Absorptive capacity and social capital in regional innovation system: the case of the Lahti region in Finland. Urban Stud 47:303–319
Knack S, Keefer P (1997) Does social capital have an economic payoff? A cross-country investigation. Q J Econ 112(4):1252–1288
Kostiainen J (2002) Urban economic development policy in the network society. Tekniikan akateemisten liitto, Tampere
La Porta R, Lopez de Silanes F, Shleifer A, Vishny RW (1997) Trust in large organizations. Am Econ Rev 87(2):333–338
Laursen K, Masciarelli F (2007) The effect of regional social capital and external knowledge acquisition on process and product innovation, DRUID working paper

Levin D, Cross R (2004) The strength of weak ties you can trust: the mediating role of trust in effective knowledge transfer. Manage Sci 50:1477–1490

Lundvall B (1992) Introduction. In: Lundvall B-Å (ed) National systems of innovation. Towards a theory of innovation and interactive learning. Pinter Publishers, London

Markusen A (1999) Fuzzy concepts scanty evidence, policy distance: the case for rigour and policy relevance in critical regional studies. Reg Stud 33:869–884

Morgan K (1997) The learning region: institutions innovation and regional renewal. Reg Stud 31:491–504

O'huallachain B, Leslie T (2007) Rethinking the regional knowledge production function. J Econ Geo 7(6):737–752

Overman H (2004) Can we learn anything from economic geography proper? J Econ Geogr 4:501–516

Patton D, Kenney M (2003) Innovation and social capital in Silicon Valley, BRIE Working Paper 155

Percoco M (2010a) Path dependence, institutions and the density of economic activities: evidence from Italian cities, FEEM Working Paper N. 110

Percoco M (2010b) Geography, institutions and urban development: Italian cities 1300–1861, Università Bocconi, mimeo

Ponds R, van Oort F, Frenken K (2010) Innovation, spillovers and university–industry collaboration: an extended knowledge production function approach. J Econ Geogr 10(2):231–255

Portes A, Landolt P (1996) The downside of social capital. Am Prospect 94(26):18–21

Putnam R (1993) Making democracy work: civic traditions in modern Italy. Princeton University Press, Princeton

Ramella F, Trigilia C (2009) Le strategie dell'innovazione. Indagine sui brevetti europei delle imprese italiane. Econ Polit Ind 36(2):199–213

Rodriguez-Pose A (1999) Innovation prone and innovation averse societies: economic performance in Europe. Growth Change 30:75–105

Rodríguez-Pose A, Crescenzi R (2008) R&D, spillovers, innovation systems and the genesis of regional growth in Europe. Reg Stud 42(1):51–67

Rodriguez-Pose A, Storper M (2006) Better rules or stronger communities? On the social foundations of institutional change and its economic effects. Econ Geogr 82(1):1–25

Rogers EM (1995) Diffusion of innovations, 4th edn. The Free Press, New York

Ruef M (2002) Strong ties, weak ties and Islands: structural and cultural predictors of organizational innovation. Ind Corp Change 11:427–449

Sabatini F (2009) Il Capitale Sociale nelle Regioni Italiane: un'Analisi Comparata. Rev Polit Econ 99:167–220

Solow R (1999) Notes on social capital and economic performance. In: Dasgupta P, Serageldin I (eds) Social capital: a multifaceted perspective. World Bank, Washington, DC

Storper M (2005) Society, community and economic development. Stud Comp Econ Dev 39 (4):30–57

Tura T, Harmaakorpi V (2005) Social capital in building regional innovative capability. Reg Stud 39:1111–1125

Varga A (1998) University research and regional innovation. Kluwer, Boston

Woolcock M (1998) Social capital and economic development: toward a theoretical synthesis and policy framework. Theor Soc 27(2):151–20

Explicitly Implicit: How Institutional Differences Influence Entrepreneurship

Stefan Bauernschuster, Oliver Falck, Robert Gold, and Stephan Heblich

Abstract Implicit institutions are shaped by societal norms and values. We expect them to impact an individual's decision to become an entrepreneur. Exploiting a natural experiment in Germany's recent history, we compare individuals born and raised in the former socialist East Germany to their West German counterparts. Our results show that the socialist regime shaped attitudes which are negatively associated with entrepreneurship. An analysis of East Germans who moved to West Germany after the fall of the Berlin Wall confirms that the socialist legacy not only runs through the channel of a less developed economic environment but indeed through implicit institutions.

Keywords Entrepreneurship • Socialism • Occupational Choice

JEL Code: A13, L26, P30

S. Bauernschuster (✉)
Ifo Institute and CESifo, Poschingerstr. 5, Munich 81679, Germany
e-mail: bauernschuster@ifo.de

O. Falck
Ifo Institute, University of Munich (LMU) and CESifo, Poschingerstr. 5, Munich 81679, Germany
e-mail: falck@ifo.de

R. Gold
Max Planck Institute of Economics, Entrepreneurship, Growth, and Public Policy Group, Kahlaische Str. 10, Jena 07745, Germany
e-mail: gold@econ.mpg.de

S. Heblich
Division of Economics, University of Stirling, CESifo, IZA and SERC, Stirling FK9 4LA, UK
e-mail: stephan.heblich@stir.ac.uk

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_9,

1 Introduction

In 1956, Soviet Union's premier Nikita Khrushchev, when addressing Western ambassadors, said: "We will bury you." At the time, such a threat did not seem farfetched; after all, the USSR had just won the space race by launching Sputnik and the West was, indeed, running scared. There was a general consensus that the central planning taking place in the Soviet Union would produce persistently high growth rates (Moore 1992). However, after some 30 years, it became clear that this fear, at least, was baseless. As nicely set out by Audretsch (2007), the socialist planned economy did well at large-scale mass production but lost ground when it came to the creativity necessary for new ideas and growth-enhancing innovation. The socialization of profits worked against Schumpeter's pioneer rent as motivation for entrepreneurial action and hence individuals lacked any incentive to build on existing knowledge and develop new ideas. Accordingly, technological progress leading to economic growth was comparatively slow, which eventually led to the system's collapse. In the long run, the free market economy proved to be superior, not the least because it provided individual freedom for entrepreneurial activity.

The general history of these two economic systems is amongst others a lesson in the importance of institutions to entrepreneurship. Delving a little deeper and following North (1990), at least two different kinds of institutions can be discerned: First are the explicit institutions in the form of laws, e.g. regarding property rights and individual freedom. Second are the implicit institutions, in the form of prevailing values and norms, which e.g. help determine an individual's risk attitude or capacity for opportunity recognition. The collapse of the former socialist countries led to a change in the explicit institutions, with many of these countries heading in the direction of a market economy. A paramount example of this development can be seen in reunified Germany, where the former socialist eastern part of the country fully adopted the explicit institutions of the West. But did the implicit institutions transform as well? A large body of literature suggests that societal norms and values develop over time, are quite persistent, and change only gradually over the course of one or two generations (Halaby 2003; Alesina and Fuchs-Schündeln 2007; Rainer and Siedler 2009).

We follow these conclusions and proceed on the assumption that some effects of the socialist regimes outlived the systems' collapse. The experience of several decades of socialism is nowadays part of the cultural heritage in numerous European countries. Apart from the explicit effects of central planning on the economic structure in these states, e.g. the prevalence of state-owned, large-scale oriented industries in the 1990es whose privatization became a major task in the transition process, at least two generations of economic actors were brought up, socialized and educated according to socialist ideology. Hence we assume that the cultural heritage of socialism still affects economic action in the former socialist countries in Europe, eventually having an impact on the economic performance of former socialist regions (c.f. Tabellini 2010) as compared to regions that were governed by liberal democratic regimes – at least from the year 1945 on.[1] Obviously, the case of Germany is well suited to assess regional differences of that kind.

Subsequently, we will focus on regional differences in entrepreneurship that might result from cultural differences between regions. Particularly, differences in social norms and values (i.e. implicit institutions) regarding e.g. the role of the individual as compared to the role of the state or the acceptance of risk taking behavior could affect the regional level of entrepreneurship. Given the importance of entrepreneurship for innovation and growth, differences in the level and the quality of entrepreneurship stemming from the cultural imprint of socialism might translate into significant differences in the economic performance between regions (c.f. Di Liberto and Usai, chapter "TFP Convergence Across European Regions: A Comparative Spatial Dynamics Analysis" in this volume). This might be of particular relevance for European transition economies. If the heritage of implicit institutions shaped by socialism hampers entrepreneurship, this could eventually become an obstacle for economic convergence (c.f. Monastiriotis, chapter "Regional Growth in Central and Eastern Europe: Convergence, Divergence and Non-Linear Dynamics" in this volume). Again, the German case provides a unique opportunity to investigate those relationships. Since the breakdown of socialism, East German regions are in a transition process, striving to catch up with West German regions. We hypothesize that within this process, East German regions face an obstacle to entrepreneurship resulting from 40 years of socialist rule that goes beyond the economic effects of central planning.

Accordingly, this chapter's goal is to assess the influence of the socialist regime in the German Democratic Republic (GDR) on implicit institutions in East Germany which in turn affect the decision to become an entrepreneur. Given the intuitive, complex interaction between explicit and implicit institutions, it does not seem feasible to design policies aimed at enhancing entrepreneurial activity, e.g. in former socialist countries, without disentangling these institutions and their effects. In order to do so, we draw upon Germany's recent history, which provides us with a natural experiment to study the effects of implicit institutions on entrepreneurship. In today's reunified Germany, citizens from both the former German Democratic Republic (GDR) as well as from the Federal Republic of Germany (FRG) face a similar explicit institutional framework for entrepreneurs. However, the legacy of a divided Germany may result in persistent differences in the implicit institutions. In particular, we expect that the experience of a socialist environment, along with an education and socialization according to communist values, will continue to influence the attitudes of individuals who were raised in the (former) GDR (cf. Mortimer and Lorence 1979; Hout 1984). As the implicit institutions prevailing in the GDR were hostile toward a market economy, it seems plausible to suppose that they also influence economic decisions such as the choice to become an entrepreneur.

[1] For a more general discussion of cultural impacts on regional economic performance see Elena Bellini, Ottaviano, Pinelli and Prarolo (chapter "Cultural Diversity and Economic Performance: Evidence from European Regions" of this volume) who identify a positive effect of cultural diversity on productivity in West European regions.

Therefore, we suspect that individuals who were born and raised in East Germany possess less entrepreneurial spirit than their fellow citizens in West Germany.

It is a challenge to analysis as well as to policy design that these effects cannot be found in macro-level data. Using the German Social Insurance Statistics, we compare the number of firm foundations in East German regions to that of West German regions and show that start-up rates are persistently higher in East Germany. The reason for macro-level data not revealing the negative effects of the socialist regime on entrepreneurship is quite obvious: The conditions and opportunities available in each part of the country differ significantly, which can be explained by an ongoing catch-up process resulting from the implementation and development of market structures after reunification.

Yet the individual motivation to become an entrepreneur as it is influenced by implicit institutions should still vary despite the catch-up process and all the economic incentives offered to affect the same. Everything else equal, we assume East Germans to possess less entrepreneurial spirit than their West German counterparts. Using micro-level data from the German General Social Survey (ALLBUS), a socioeconomic survey conducted on a sample of the German micro-census, we compare East and West Germans on a set of different norm and value variables and find substantial differences. Moreover, we can show that these differences are strongly associated with the probability of being an entrepreneur. Further, we would like to make the point that regional heterogeneity is not the only channel for our results. There is more to the legacy of the socialist regime than "just" a poor and less developed economic environment. This is demonstrated by looking at the subsample of individuals who were born in East Germany but migrated to West Germany after the fall of the Iron Curtain. Even this selective subsample shows a clear effect of the treatment with socialist norms and values which are negatively associated with entrepreneurship.

The remainder of this chapter is organized as follows: Section 2 develops the role of explict and implicit institutions in entrepreneurial activity in more detail. Section 3 introduces the German history of separation and reunification as natural experiment. Section 4 describes the differences in entrepreneurial activity between East and West Germany based on macro-level data. Section 5 uses micro-data to analyze the effect of the socialist regime on a whole set of societal norms and values and shows that these implicit institutions are negatively associated with entrepreneurship to a substantial degree. Section 6 concludes.

2 Entrepreneurship and Institutional Incentives

2.1 What Drives Entrepreneurship?

The decision to start a business and become an entrepreneur is influenced by various factors,[2] not least by personal characteristics. As Kihlstrom and Laffont

[2] For an overview, see Parker (2009).

(1979) show, these include an individual's risk attitude as well as his or her motivation (Schumpeter 1912) and skills (Lazear 2005), along with the ability to spot niches in the market (Kirzner 1973), raise financial resources (Michelacci and Silva 2007; Guiso et al. 2004), and networking ability (Sanders and Nee 1996; Stuart and Sorenson 2005).[3]

The decision to start a business is also influenced by external characteristics based in the institutional framework. This is made obvious in Saxenian's (1994) comparison between the evolution of Silicon Valley in California and Route 128 in Boston, Massachusetts. Much of Silicon Valley's greater success compared to Route 128 is the result of institutional factors. Apparently, the absence of legal restrictions on job mobility and the resulting diffusion of knowledge led to a vertically disintegrated, entrepreneurial business culture in Silicon Valley, where new ideas quickly result in new firms (cf. Gilson 1999; Klepper 2009). Public research facilities, leading to increased knowledge flows, and public funding are another example of institutional factors that stimulate entrepreneurship. By contrast, an extensive welfare system could affect individual risk-aversion and (leisure) time preferences in ways that lessen the incentive to become an entrepreneur (Parker and Robson 2004; Fölster 2002).[4]

Both personal and external factors affect an individual's decision-making process, but their relative importance depends on the person's psychological makeup. For instance, entrepreneurial individuals are expected to be more risk accepting, self-confident, and independent (cf. Blanchflower and Oswald 1998; Camerer and Lovallo 1999). Accordingly, individuals with strong entrepreneurial intentions are likely to overcome financial or other constraints, whereas less entrepreneurial individuals might be discouraged more easily. An individual's self-image is strongly influenced by his or her education and overall socialization (Halaby 2003; Falck et al. 2009), which determine how the person understands the prevailing social norms and habits, which in turn shape the individual's view of *who he or she is* and what the individual and others *should* or *should not* do (Bernhard et al. 2006). Accordingly, these non-codified social obligations act as implicit institutions that do not explicitly prescribe individual behavior, but nevertheless have a crucial impact on economical decisions and actions (North 1991). Particularly, they affect an individual's tendency to have a rather entrepreneurial or a rather bureaucratic job orientation (Miller and Swanson 1958).

---

[3] Anyhow, networking does not only have an individual but also a regional component. Regional networks may support entrepreneurs as long as entrepreneurs manage to build up and maintain trust (Sanders and Nee 1996); i.e. the availability of social capital in a region affects entrepreneurship as well (Bauernschuster et al. 2010). See also Crescenzi et al. (chapter "The 'Bright' Side of Social Capital: How 'Bridging' Makes Italian Provinces More Innovative" in this volume); who assess the impact of social capital on regional innovative output for the case of Italy.

[4] For a detailed analysis of the interactions between personal characteristics, institutional environment and space with respect to academic entrepreneurship see Abreu and Grinevich (chapter "Academic Entrepreneurship and the Geography of University Knowledge Flows in the UK" in this volume).

2.2 Implicit Institutions and Entrepreneurship

We define implicit institutions as the mindsets individuals develop by being exposed to their society's norms, values and traditions. With regard to entrepreneurship, these implicit institutions might influence an individual's desire to be an entrepreneur, as well as his or her risk attitude and capacity for opportunity recognition. Consider an individual growing up in an environment of freedom, liberalism, and self-realization. This person might never have read about Schumpeter's (1912) entrepreneurial virtues, but he or she will certainly have a better understanding of them, even if not explicitly, than will an individual growing up in an egalitarian society where competition and individual self-realization are proscribed (cf. Alesina and Fuchs-Schündeln 2007 or Bauernschuster et al. 2012). In an environment where self-reliance and self-realization are not rewarded by the expectation of future profits, most people would rather work 9 to 5 for predictable wages. The resulting increase of risk aversion and decrease of individual incentive will eventually crowd out the entrepreneurial spirit altogether.

It is this situation that describes the business environment prevalent in the socialist countries of the Eastern Bloc. These planned economies had no room for entrepreneurial activity and their suppressive political regimes favored communist ideals and egalitarianism over liberalism and individuality. Private property was nationalized and for nearly 50 years were people raised and educated according to socialist values—a period long enough to develop the belief that conformity was the norm, individuality a form of deviance.[5] Thus, implicit institutions rejecting entrepreneurship were established and internalized over a fairly long period, making them unlikely to vanish over night, regardless of how the rest of the world changed. Everything else equal, individuals from East Germany should consequently show a lower propensity to become an entrepreneur than their fellow citizens from the West, thus confirming the hypothesis that implicit institutions do exert influence on the individuals' occupational choice.

3 German Reunification as a Natural Experiment

German separation and its termination through reunification can be viewed as an exogenous shock (Bach and Trabold 2000; Frijters et al. 2004; Fuchs-Schündeln and Schündeln 2005; Fuchs-Schündeln 2008). After World War II, Germany was divided into two parts. At this time, both parts of the country were mostly indistinguishable with respect to economic measures (Alesina and Fuchs-Schündeln 2007), but diverged diametrically afterwards. By 1952, the inner German border was so fiercely guarded

[5] Eventually, this lack of individual incentives also contributed to the low level of productivity in Eastern Bloc countries (VanArk 1996), particularly to the lack of productivity of the GDR as compared to the FRG (VanArk 1995).

that it was extraordinary difficult for East Germans to enter West Germany. The city of Berlin was the only place where it was still feasible to cross the border. However, the building of the Berlin Wall in 1961 closed even this gate. During the following decades, the two German states developed very differently. Under the influence of the Western Allies, West Germany became a democratic state with a free market economy, whereas East Germany turned into a socialist, centrally planned economy under the influence of the Soviet Union. Naturally, these diametrical contexts influenced their inhabitants' worldviews and attitudes toward the state and society. And as a consequence, the East and West Germans were not only physically separated by walls and barbed-wire fences but became also separated from each other by their implicit institutions. It was only when the Berlin Wall fell in 1989 and Germany was subsequently reunified that this sharp separation came to an end.

German reunification in 1990 resulted in the present situation where all Germans, regardless of whether they were raised in the GDR or the FRG, now share a common democratic constitution that guarantees the rule of law, property rights, and (economic) freedom. In other words, all market actors in Germany today operate within a very similar institutional framework. However, implicit institutions (i.e., mindsets and value systems) cannot be changed by edict and we therefore expect persisting differences in the social norms and values prevalent in these two formerly separated parts of Germany. Considering that the socialist ideology systematically oppressed entrepreneurship and entrepreneurial virtues, the suspected differences in implicit institutions between East Germans and West Germans should result in systematic differences in the desire to become an entrepreneur that continue to exist years after the German reunification.

4 Empirics on the Macro Level

4.1 Analytical Framework

We hypothesize that implicit institutions have an influence on the decision to start a business and become an entrepreneur. Since East Germans were brought up in a socialist country, we suppose they may be more critical toward entrepreneurship than their fellow citizens who grew up in the Federal Republic of Germany. All else equal, this should be expressed by a smaller number of entrepreneurs in the eastern part of Germany as compared to the western part. However, this ceteris paribus condition is difficult to fulfill due to substantial structural differences between East and West Germany, as we detail next.

As "agents of change and growth" (OECD 1998, p. 11), entrepreneurs are expected to play a crucial role in the transition from a centrally planned economy, such as that of the former GDR, to a free market economy, such as it now exists in reunified Germany. "Entrepreneurs not only seek out potentially profitable economic opportunities but are also willing to take risks to see if their hunches are right" (OECD 1998, p. 11). The immense structural change that occurred in East Germany following the "jump start" of reunification (Sinn and Sinn 1992) certainly created a great set of opportunities to start up new businesses and firms. Hence, the very first years after reunification are characterized by intense entrepreneurial activity in the East German regions. Implementation of a market economy in the former socialist East resulted in the privatization of state-owned firms as well as in new firm startups in all sectors. It was, in short, an extraordinary promising time to become an entrepreneur. The absence of an established market structure, not to mention the lack of competitors, was fertile ground for new ventures.

We use data provided by the German Social Insurance Statistics to examine differences in the level of entrepreneurial activity as measured by firm foundations. The Social Insurance Statistics requires each employer to report certain information, e.g., qualifications, about every employee subject to obligatory social insurance. The information collected can be transformed into an *establishment file* that provides longitudinal information about the establishments and their employees.[6]

Since the decision to become an entrepreneur is, amongst others, subject to regional heterogeneity, we identify regions in West Germany and in Each Germany that are similar with regard to the regional factors driving entrepreneurship. Accordingly, we concentrate on regions in West Germany that are close to the former border with East Germany. Before reunification, these FRG regions were classified as peripheral and they received reduced public infrastructure investment. The private sector also invested rather modestly in the boarder region and consequently, these regions were economically underdeveloped and still are to a great extent. Despite all these problems, a few northern regions near the former boarder always have done and continue to do quite well. We omit these particular border regions from our analysis. Instead, we concentrate on the border regions that are classified as development regions under Objective 2 from 2000 to 2006 (respectively, Objective 5b from 1994 to 1999) of the European Structural Funds. To compensate for the gap resulting from the exclusion of the non-funded regions, we additionally include the Objective 2 regions in eastern Bavaria, which did not lie on the inner German border but border on the Czech Republic, former CSFR. Thus, these regions also faced the problems inherent to being on the outskirts of the free world, hemmed in by the Iron curtain.

The whole area of the former GDR falls into Objective 1 of the Structural Funds, thus making it eligible for public funding of economic development. We concentrate on the East German districts that adjoin the former border plus those districts that adjoin these border districts. These regions along the border are similar to their western counterparts with respect to geography and natural resources. Furthermore, their situation in the GDR was comparable to the conditions to the western districts in one important aspect: all of these districts were located at the far end of the respective country. Thus, the border region is most suitable for comparing East and

[6] For a detailed description of this data, see Fritsch and Brixy (2004). A detailed description of how the number of start-ups is computed is provided by the authors upon request.

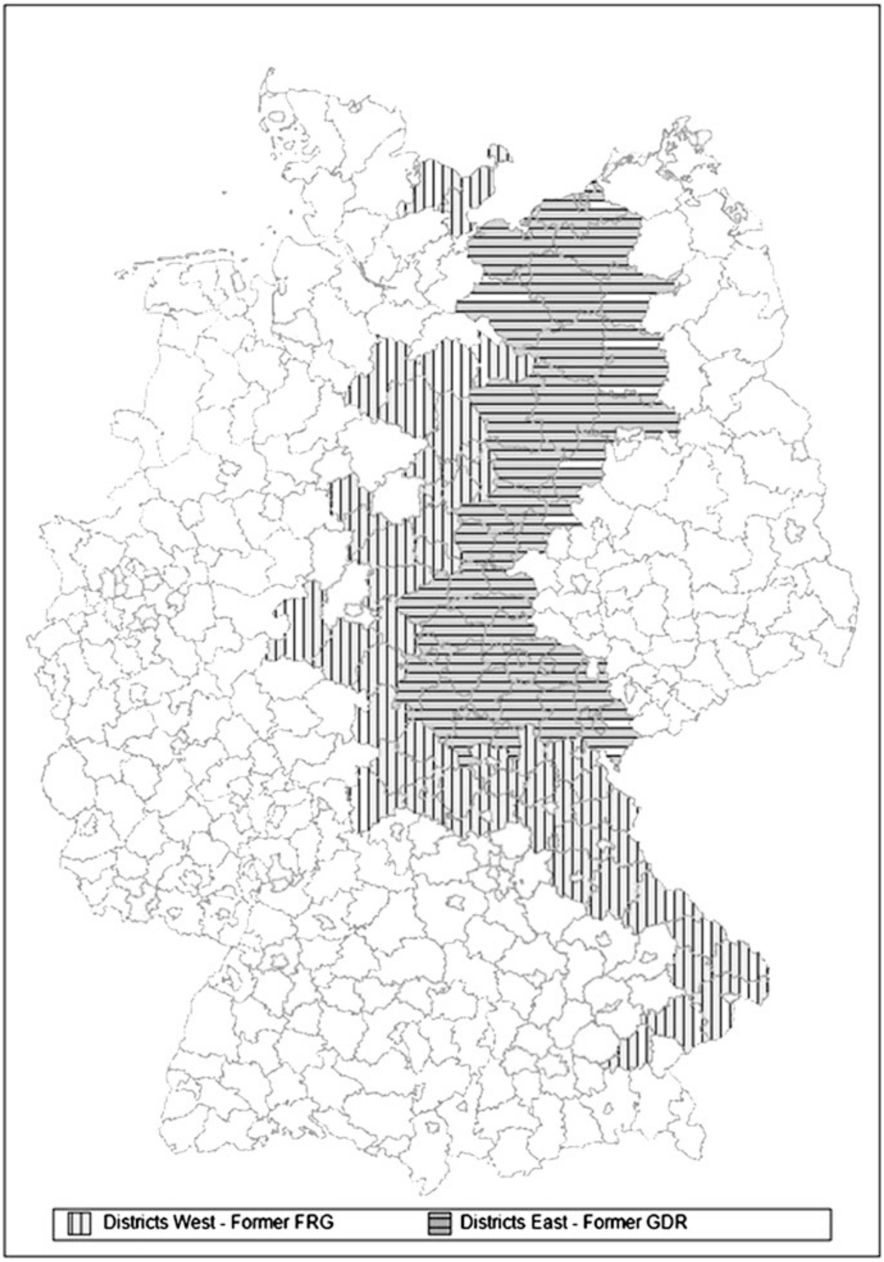


Fig. 1 Area under investigation

West Germany, since the districts within these regions are as similar to each other as it is possible to be. Some descriptive statistics for the observed districts are presented in Appendix 1, whereas Fig. 1 presents a graphical description.

Table 1 Start-up rates in manufacturing: comparison between East and West German districts

		1999	2000	2001	2002	2003	2004
Number of start-ups in manufacturing	West	20.54	21.38	17.8	18.88	17.82	16.76
	East	33.32	23.36	20.72	21.96	19.70	19.19
	Diff.	12.78***	1.98	2.92**	3.08*	1.88	2.43*
Start-up rate I (start-ups/ employees in manuf.)	West	0.0021	0.0021	0.0018	0.0019	0.0019	0.0019
	East	0.0054	0.0038	0.0035	0.0036	0.0033	0.0033
	Diff.	0.0034***	0.0018***	0.0017***	0.0017***	0.0014***	0.0014***
Start-up rate II (start-ups/firms in manuf.)	West	0.0563	0.0591	0.0515	0.0546	0.0529	0.0516
	East	0.0963	0.0713	0.0669	0.0712	0.0659	0.0668
	Diff.	0.0400***	0.0122***	0.0154***	0.0167***	0.0129***	0.0152***

Notes: * denotes 10 % level of significance; ** denotes 5 % level of significance; *** denotes 1 % level of significance

4.2 Differences in Start-Up Rates

To assess differences in the entrepreneurial activity between East and West Germany, we compare start up activities between 47 East German and 50 West German border districts for each year from 1999 to 2004. To do so, we calculate two different start-up rates by dividing the number of start ups by number of employees and number of firms in manufacturing respectively. According to our hypothesis, implicit institutions should have a negative effect on start-up rates in the eastern regions if all confounding factors were the same. However, as Table 1 reveals, this effect cannot be found using macro data. In fact, the start-up rates in the eastern regions exceed the western start-up rates in every year of analysis, regardless of how the rate is calculated.

Overall, the number of startups in manufacturing decreases in both West and East Germany, albeit not uniformly. Over time, the difference between the eastern and the western districts diminishes. If we follow a labor market approach and calculate *start-up rate I* as the number of startups in manufacturing divided by the number of employees in manufacturing, we see a decline in both the West and the East over time. The difference between East and West decreases but remains significant. Under this start-up rate, the probability that an employee in the East German regions founds a business in manufacturing is continuously higher, even 14 years after reunification. If we take an ecological approach[7] and calculate *start-up rate II* as the number of startups divided by the number of manufacturing firms in the respective regions, we obtain a slightly different result with more variance. This start-up rate is relatively stable in the western districts, but experiences a remarkable decline in the eastern districts from 1999 to 2000. After 2000, the rate seemingly decreases in the eastern area even though it continues to be significantly higher than the start-up rates in the western area.

[7] For a comparison of the labor market approach and the ecological approach, see Audretsch and Fritsch (1994).

All together, these regional data do *not* show that implicit institutions inherited from the socialist era in eastern Germany have a negative influence on the start-up activity in the eastern parts of Germany. There are three possible explanations for this counterintuitive result. It might be that (1) implicit institutions do *not* vary between the two regions at all. Or, perhaps it is that (2) implicit institutions *do* differ between the two regions, but do not affect entrepreneurship. Alternatively, it could also be that (3) implicit institutions are significantly different in East as compared to West Germany and do hinder entrepreneurship in eastern Germany, but that this effect is *overcompensated* by the prevailing transition process.

Which of these is true, cannot be satisfyingly examined using district-level data. This is because the transition process is still in progress and the turbulence induced by reunification still blurs macro-data (Barro and Sala-i-Martin 1991; Fritsch 2004; Schindele 2010). We see significant differences in both opportunity entrepreneurship (cf. Reynolds et al. 2005) as well as necessity entrepreneurship (cf. Kirzner 1973) in the eastern part of our area under investigation. The figures presented in Appendix 2 provide even more evidence that East Germany is experiencing an ongoing development process that distorts the empirical analysis of aggregated data. In the remainder of the chapter, we hence face up to the problems of macro-level data and turn to micro-level data instead to further investigate the role of implicit institutions in entrepreneurship.

5 Empirics at the Micro Level

In order to examine our hypothesis that implicit institutions developed under the socialist regime of the GDR negatively influence entrepreneurship, we now adjust our empirical strategy and turn to micro-level data. First, we show that the socialist regime in the GDR really changed societal norms and values related to entrepreneurship. Second, we collect evidence that these brought about implicit institutions indeed have a negative impact on an individual's propensity to become an entrepreneur.

To these ends, the German General Social Survey (ALLBUS) is a valuable data source.[8] The dataset is based on biennial, representative surveys of the German population conducted through personal interviews. ALLBUS covers a wide range of topics pivotal to empirical research in social sciences. A core set of questions is asked in every wave of the survey, with various sets of additional questions added in different years.[9]

---

[8] The ALLBUS program was financially supported by the German Research Foundation (DFG) from 1980 to 1986 and in 1991. Further surveys were financed on a national and federal state (Laender) level via the GESIS network (Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen).

[9] Terwey et al. (2007) provide detailed information on the ALLBUS surveys in general and present all variables available in the cumulated dataset from 1980 until 2006.

5.1 Entrepreneurial Attitudes of East and West Germans

We use the 1991 wave, which a.o. contains information on individual risk attitude with regard to job security, and the 1994, 1998, 2000 and 2004 waves, which contain information on individual's norms and values expectedly related to entrepreneurship. A first look at variables concerning individual attitudes towards the state's responsibility (norms), the economy and society (values), as well as risk-aversion reveal significant differences in implicit institutions between East and West Germans.

Table 2 shows simple mean comparisons of a whole set of value and norm variables between West Germans and East Germans. Briefly summing up the results, the descriptive analysis clearly and consistently shows that East Germans tend to be less self-reliant and more reliant on the state; reveal a rather skeptical attitude towards the market economy; and show higher levels of risk-aversion than their West German counterparts. As can be seen in the last column of Table 2, t-tests confirm that these differences are all significant at a 1 % level. Thus, implicit institutions as shaped by the socialist regime appear to exist in the eastern parts of Germany even years after reunification.

5.2 Occupational Choice and Entrepreneurial Attitude

In order to assess whether these implicit institutions are indeed associated with the decision to become an entrepreneur, we include the attitude variables depicted in Table 2 into an occupational choice equation. Although we believe that they are all related with the decision to become an entrepreneur, we deliberately choose to include only eight of the ten attitude variables in a multivariate setting. The reason for this is that two of the variables might severely suffer from reverse causality, namely the statement on the "Role of enterprise" and the "Intervention of the state". It might very well be that being an entrepreneur positively affects the agreement to the statement that employers' profits foster the economy. The same is true for the state intervention variable: Being an entrepreneur might drive the statement that the state has to care for employment and price stability *even if this cuts the rights of employers*. Since we do not know how this potential reverse causality problem might bias the coefficients of the other attitude variables of interest, we do not include them in our estimations.

The estimated occupational choice equation is described by Eq. 1:

$$\Pr(y_i = 1|\cdot) = \alpha + \beta_1 att_i + X_i\beta_2 + \varepsilon_i, \qquad (1)$$

where $\Pr(y_i = 1|\cdot)$ is the conditional probability of being an entrepreneur. y is an indicator variable that takes the value of unity if person i is an entrepreneur and takes the value 0 if person i is dependently employed. att is the level of agreement with the statements revealing attitudes. X is a set of control variables that might influence a

Table 2 Differences in attitudes between East and West Germans

	East Germans	West Germans	
	Level of agreement on a scale from 1 to 4 (1 = fully agree; 4 = don't agree at all)		Difference
Role of enterprise*: "Employers' profits foster the economy."* (obs.: 900 and 1,704)	2.206 *0.031*	1.972 *0.020*	0.234 ***
Welfare state: *"The current social security system reduces work incentives."* (obs.: 900 and 1,700)	2.992 *0.031*	2.485 *0.023*	0.507 ***
Fairness*: "Economic profits are distributed fairly in Germany."* (obs.: 887 and 1,674)	3.363 *0.021*	2.942 *0.018*	0.421 ***
Income differences*: "Income differences give incentives to work hard."* (obs.: 2,304 and 3,210)	2.564 *0.019*	2.293 *0.015*	0.271 ***
Rank differences*: "Rank differences are performance based and therefore acceptable."* (obs.: 2,324 and 3,196)	2.727 *0.019*	2.464 *0.015*	0.262 ***
Status differences: *"Social status differences are just – by and large"* (obs.: 2,341 and 3,208)	3.221 *0.016*	2.629 *0.015*	0.592 ***
State intervention*: "The state has to care for employment and price stability even if this cuts the rights of employers."* (obs.: 894 and 1,688)	1.951 *0.030*	2.287 *0.022*	−0.336 ***
National assistance: *"The state has to care for the sick, poor, old and unemployed."* (obs.: 2,380 and 3,267)	1.437 *0.013*	1.802 *0.013*	−0.366 ***
Performance: *"Everybody should get the money he needs – regardless of any performance."* (obs.: 2,327 and 3,230)	2.639 *0.020*	2.754 *0.016*	−0.115 ***
	East Germans	West Germans	Difference
	Level of importance on a scale from 1 to 7 (1 = not important; 7 = very important)		
Risk attitude: *"How important is a secure job to you?"* (obs.: 631 and 859)	6.661 0.029	6.087 0.047	0.574 ***

Notes: In the cells below the respective questions we report the number of observations in *parentheses*. The *first number* refers to East Germans, i.e., individuals who currently live in East Germany. The *second number* refers to West Germans. Column 1 and Column 2 depict group means for East Germans and West Germans respectively, while Column 3 shows mean differences; standard errors are given in *italics*. * denotes 10 % level of significance, ** denotes 5 % level of significance, *** denotes 1 % level of significance

person's occupational choice, i.e., we include control variables for gender, home ownership, marital status, nationality, education, working experience, and individual unemployment history. All control variables except working experience differ significantly between East and West Germany. Finally, we include a dummy for

East Germany and year dummies if appropriate. We estimate occupational choice equations by simple probit models.

In order to make use of all information on the attitude variables available over the years, we run four estimations, where each single estimation includes, in addition to the controls, only those attitude variables having a perfect overlap with respect to the years they were collected in. In particular, we run a probit occupational choice model for the year 1991 using the “Risk attitude” variable. Another probit estimation is executed for the years 1994 and 2004 where we have got information on the “Welfare state” and “Fairness” attitudes. The third estimation uses information on the “Performance” and “National assistance” attitudes for the years 1991, 1994, 2000, and 2004, whereas the final estimation exploits information on “Income differences”, “Rank differences”, and “Status differences” attitudes provided in the years 1991, 1994, 1998, 2000, and 2004. The results of the models are reported in Table 3.

As can be seen from Table 3, the coefficients of all the attitude variables under investigation point to the direction we would suspect and are precisely estimated. Almost all of them are found statistically significant at a 1 % level. Summing it up, the results suggest that a lack of self-reliance, a sceptical attitude towards the market economy, as well as risk-aversion is associated with a lower probability of becoming an entrepreneur. Accordingly, implicit institutions are indeed associated with an individual’s propensity to become an entrepreneur. The differences in implicit institutions between East and West Germans shown in Table 2 translate into differences in the individual probability to become an entrepreneur

Of course, it is important to note that we cannot claim to identify causal effects of implicit institutions on entrepreneurship here, because, despite of controlling for well-known and important determinants of occupational choice, unobserved characteristics might be correlated with the risk attitude, norm and value variables, and at the same time have an impact on an individual’s occupational choice, which would cause the error term to be correlated with the variable of interest. However, what we can show so far is that the socialist regime led to an environment where we find implicit institutions to be negatively associated with entrepreneurship, without knowing the details about the accurate channels, though.

5.3 What’s the Channel? More than “Just” (Un)Observed Regional Characteristics?

So far, we can argue that the socialist regime in the former GDR left an environment where societal norms and values prevail that are negatively associated with entrepreneurship. One might argue that the substantial norm and value differences are mainly due to observable and unobservable labor market heterogeneity, since e.g. the level of unemployment in East Germany is much higher than it is in West Germany. Note in this context that our dataset includes only the dependently employed and the entrepreneurs the unemployed are excluded. Nevertheless, even for those who are

Table 3 Occupational choice of becoming an entrepreneur

	Cross section	Repeated cross section	Repeated cross section	Repeated cross section
	1991	1994, 2004	1991, 1994, 2000, 2004	1991, 1994, 1998, 2000, 2004
Risk attitude	−.020 ***			
	.004			
Welfare state		−.034 ***		
		.006		
Fairness		−.017 **		
		.008		
Performance			.016 ***	
			.004	
National assistance			.029 ***	
			.005	
Income differences				−.004
				.004
Rank differences				−.014 ***
				.005
Status differences				−.015 ***
				.005
Controls	Yes	Yes	Yes	Yes
No. of observations	1,591	2,911	6,276	7,285
Wald test	93.21 ***	150.73***	333.45***	372.38***
Pseudo R^2	0.112	0.078	0.079	0.078

Notes: The table reports probit marginal effects where the dependent variable "occupational status" is unity for entrepreneurs and zero for employees. The variable "Job security" refers to the question "How important is a secure job for you?", where the level of importance ranges from 1 (not important) to 7 (very important). The variable "Welfare state" refers to the statement "The current social security system reduces work incentives." The variable "Fairness" refers to the statement "Economic profits are distributed fairly in Germany.", "Performance" refers to the statement "Everybody should get the money he needs – regardless of any performance.", "National Assistance" refers to the statement "The state has to care for the sick, poor, old and unemployed.", "Income Differences" refers to the statement "Income differences give incentives to work hard.", "Rank Differences" refers to the statement "Rank differences are performance based and therefore acceptable.", and "Status Differences" refers to the statement "Social status differences are just – by and large." All these statements could be evaluated on a scale from 1 (fully agree) to 4 (disagree). Robust standard errors are given in *italics*. * denotes 10 % level of significance, ** denotes 5 % level of significance, *** denotes 1 % level of significance

employed, economic conditions may differ a lot between East and West Germany, with the consequence that e.g. people living in East Germany might on average be more afraid of getting unemployed than those living in West Germany, which could influence their norms and values. In a next step, we would like to show that there is more behind the norm and value differences than this, i.e., we set out to demonstrate that the legacy of the socialist regime is not "just" unobserved regional characteristics which affect implicit institutions.

Therefore, we refine our analysis to concentrate on those East Germans who moved to West Germany after 1989 and compare them to born and raised West Germans. These movers had been socialized in the GDR and thus had been "treated" with the communist ideology, but were now confronted with the same economic conditions as their West German counterparts after they moved.

This strategy could be reasonably criticized on the grounds that the group of movers might be subject to a selection bias, i.e. that movers are in general energetic and self-reliant and thus different in attitude as compared to the average East German. However, note that the selectivity of the group of movers should consequently work against us by reducing the differences in norms and values between East and West Germans. This means that any remaining differences for this subgroup mark a lower bound and are likely to be more pronounced in the overall East German sample.

The mean values of the societal norms and value variables presented in Table 4 are in line with our expectation that the subgroup of East German movers should show smaller differences to West Germans than other East Germans. But even if the selection problem, which works against us, is present to some degree, we still find differences even between this selective East German subgroup and West Germans. All differences continue to run in the same direction as before and most of them are still statistically significant. Thus, these differences support the view that we indeed observe an effect of the communist ideological treatment on implicit institutions, which does not merely go through the channel of observable and unobservable characteristics of a poor economic environment.

The number of observations for East Germans who moved to West Germany after 1989 is not large, but still provides insightful results. To obtain these results, we collapsed information on norms and values, if it was available, for different points in time. Note the great difference in the attitude towards risk between East German movers and West Germans. This information on risk attitude is only gathered in 1991. Consequently, we only observe those East German movers who moved to West Germany between 1989 and 1991, i.e., right after the fall of the Berlin Wall. This explains both the low number of observations as well as the great gap in the means. When this information was collected, these movers did not have much time to adapt to West German norms and values. Rather, this is the group for which the socialist past was still very recent indeed.

6 Conclusions

The goal of this chapter is to disentangle the effects of explicit and implicit institutions on individuals' entrepreneurial intentions. Explicit institutions can change rather quickly; implicit institutions, however, here defined as societal values and norms, develop and change much more slowly. To identify the effect of implicit institutions on an individual's entrepreneurial intentions requires a situation where people now living under the same explicit institutions were raised

Table 4 Differences in attitudes between West Germans and East German movers

	East German Movers	West Germans	
	Level of agreement on a scale from 1 to 4 (1 = fully agree; 4 = don't agree at all)		Difference
Role of enterprise: *"Employers' profits foster the economy."* (obs.: 58 and 1,704)	2.190 *0.114*	1.972 *0.020*	0.218 **
Welfare state: *"The current social security system reduces work incentives."* (obs.:59 and 1,700)	2.746 *0.117*	2.485 *0.023*	0.261 **
Fairness: *"Economic profits are distributed fairly in Germany."* (obs.: 59 and 1,674)	3.169 *0.088*	2.942 *0.018*	0.227 ***
Income differences: *"Income differences give incentives to work hard."* (obs.: 97 and 3,210)	2.485 *0.087*	2.293 *0.015*	0.192 **
Rank differences*: "Rank differences are performance based and therefore acceptable."* (obs.:100 and 3,196)	2.490 *0.086*	2.464 *0.015*	0.026
Status differences*: "Social status differences are just – by and large"* (obs.: 100 and 3,208)	2.880 *0.079*	2.629 *0.015*	0.251 ***
State intervention: *"The state has to care for employment and price stability even if this cuts the rights of employers."* (obs.: 59 and 1,688)	2.102 *0.110*	2.287 *0.022*	−0.185 *
National assistance*: "The state has to care for the sick, poor, old and unemployed."* (obs.:99 and 3,267)	1.737 *0.073*	1.802 *0.013*	−0.065
Performance*: "Everybody should get the money he needs – regardless of any performance."* (obs.: 100 and 3,230)	2.650 *0.094*	2.754 *0.016*	−0.104
	Level of importance on a scale from 1 to 7 (1 = not important; 7 = very important)		
Risk attitude*: "How important is a secure job to you?"* (obs.: 14 and 631)	6.857 *0.097*	6.087 *0.047*	0.770 ***

Notes: In the cells below the respective questions we report the number of observations in *parentheses*. The *first number* refers to late East German movers, i.e., East Germans that were born in the former GDR but moved to West Germany not before 1989. The *second number* refers to West Germans. Column 1 and Column 2 depict group means for late East German movers and West Germans respectively, while Column 3 shows mean differences; standard errors are given in *italics*. * denotes 10 % level of significance, ** denotes 5 % level of significance, *** denotes 1 % level of significance

and socialized under different regimes and thus—assumedly—developed different observable values and norms. In this regard, Germany provides a suitable natural experiment due to its unique history of separation into two distinct systems and states, the socialist GDR and the non-socialist FRG, following World War II. Because, according to Alesina and Fuchs-Schündeln (2007), both parts of the country were quite comparable before this split, observable differences after the separation are in all probability driven by the prevailing ideologies in each part, which gradually produced different values.

We suppose that the socialist regime in the GDR influenced implicit institutions in a way which is negative for entrepreneurship. However, comparing the start-up rates of East Germany with those of West Germany, we find that entrepreneurial activity is higher in the former GDR than it is in the regions that always belonged to the FRG. This result holds if we focus on the corridor along the former inner-German border, thus encompassing only regions having similar natural conditions and equal experience with being located so close to the Iron Curtain. This higher level of entrepreneurial activity appears to be the result of significant differences in both opportunity and necessity entrepreneurship in East Germany, thus obscuring the hypothesized negative effect of the socialist regime on implicit institutions negatively related to entrepreneurship. Indeed, we see some evidence of economic development convergence with West Germany in the former GDR, but the eastern part of the country is still "catching up." In the face of this transition, comparatively many people found their own business, even if their job orientation is not exceedingly entrepreneurial. Whether opportunity or necessity entrepreneurship is dominant within this process might be an interesting question for further research.

Using micro-level data, we show that the socialist regime in the GDR had indeed a causal effect on implicit institutions, i.e., a whole set of norm and value variables. Further analyses clearly reveal that these implicit institutions were shaped in a way which is negatively associated with entrepreneurship. To analyze whether the influence of the socialist past runs merely via the channel of observable and unobservable labor market characteristics, i.e., a generally bad economic environment in East Germany, we switch the focus to those individuals who were born in East Germany and then migrated to West Germany after the fall of the Berlin Wall.[10] We find that the socialist legacy can even be found in this very selective subgroup of East Germans. Hence, we suggest that the differences in implicit institutions are not merely driven by the specific economic conditions in East Germany.

Our analyses suggest that implicit institutions in the form of values and norms affecting individuals' mindsets prevail over and above explicit institutions. Individuals in a presently similar environment but who were socialized under different ideologies do differ in their underlying value systems. These differences can affect economic decisions, perhaps most especially the decision about whether to start a business. This finding should be a particularly important consideration in the design of policies geared toward stimulating entrepreneurial activities. Our findings strongly advise against too-general, catch-all policies. Some facets of entrepreneurship are doubtless universal, such as the necessary financing. However, our study shows that the incentives to become an entrepreneur are also affected by social respectively cultural factors which might vary between regions. Growing up

[10] See for migration patterns of East Germans Burda (1993), Burda et al. (1998), and Hunt (2006).

under a socialist regime appears to be one such factor. According to Alesina and Fuchs-Schündeln (2007), underlying values, or in our terms, implicit institutions, can take several generations to change. Supporting this change might be another policy issue.

Appendix

Appendix 1: Descriptive Statistics for Area Under Investigation

	East 47			West 50		
Number of regions (NUTS3)	Min	Mean	Max	Min	Mean	Max
Total number of firms	1,170	3,047	7,337	1,480	3,009	6,390
Number of firms in manufacturing	91	322	977	115	363	687
Total number of employees	16,424	18,969	115,063	11,681	36,704	109,056
Number of employees in manufacturing	2,119	6,448	17,138	2,677	12,376	61,546
Total number of startups	103	328.67	1,272	105	275.49	745
Number of start ups in manufacturing	4	23.04	79	3	18.86	53
Total number of firm shutdowns	0	279.06	1,160	0	197.56	739
Number of shutdowns in manufacturing	0	23.33	97	0	19	57
Inhabitants	44,076	108,160	237,833	49,462	124,085.50	266,070
Population density	40	251.57	1,170	42	203.72	1,534
GDP	834,195	1,904,423	5,811,596	858,014	2,814,694	9,005,517
GDP per capita	12.06	17.51	30.14	12.78	23.05	73.89

Appendix 2: Persisting Differences Between East and West Germany (Districts' Averages)

		1999	2000	2001	2002	2003	2004
GDP	West	2,713,335	2,755,099	2,824,650	2,837,641	2,851,413	2,906,028
	East	1,810,424	1,845,522	1,893,191	1,926,538	1,947,363	2,003,500
	Diff.	902,911***	909,576***	931,459***	911,103***	904,051***	902,528***
GDP per capita	West	22.17	22.52	23.13	23.25	23.34	23.89
	East	16.28	16.70	17.32	17.84	18.17	18.76
	Diff.	5.88***	5.82***	5.81***	5.42***	5.17***	5.13***
Firms in manufacturing	West	377.36	374.62	366.4	358.8	353.24	345.18
	East	346.81	335.15	324.91	315.26	307.00	300.53
	Diff.	30.55	39.47	41.49*	43.54*	46.24*	44.6*
Employees in manufacturing	West	12,649	12,675	12,673	12,350	12,081	11,829
	East	6,324	6,488	6,549	6,503	6,419	6,404
	Diff.	6,324***	6,186***	6,124***	5,847***	5,662***	5,425***
Shutdowns in manufacturing	West	28.24	28.16	29.38	28.22	-	-
	East	38.38	34.45	35.89	31.26	-	-
	Diff.	10.14***	6.29**	6.51**	3.04	-	-
Inhabitants	West	123,999	124,078	124,149	124,230	124,146	123,911
	East	110,009	109,375	108,579	107,783	106,996	106,218
	Difference	13,990*	14,704*	15,570**	16,447**	17,150**	17,692**

Notes: * denotes 10 % level of significance, ** denotes 5 % level of significance, *** denotes 1 % level of significance

References

Alesina A, Fuchs-Schündeln N (2007) Good bye Lenin (or not?)—the effect of communism on people's preferences. Am Econ Rev 97(4):1507–1528

Audretsch DB (2007) The entrepreneurial society. Oxford University Press, Oxford

Audretsch DB, Fritsch M (1994) On the measurement of entry rates. Empirica 21(1):105–113

Bach S, Trabold H (2000) Ten years after German monetary, economic and social union: an introduction. Quart J Econ Res 69(2):149–151

Barro RJ, Sala-i-Martin X (1991) Convergence across states and regions. Brooking Pap Eco Ac 22 (1):107–182

Bauernschuster, S, Falck, O, Gold, R, Heblich, S (2012) The shadows of the socialist past: lack of self-reliance hinders entrepreneurship. Eur J Pol Econ 28(4):485–497

Bauernschuster S, Falck O, Heblich S (2010) Social capital access and entrepreneurship. J Econ Behav Organ 76(3):821–833

Bernhard H, Fehr E, Fischbacher U (2006) Group affiliation and altruistic norm enforcement. AEA Pap Proc 96(2):217–221

Blanchflower DG, Oswald AJ (1998) What makes an entrepreneur? J Labor Econ 16(1):26–60

Burda MC (1993) The determinants of east–west German migration – some first results. European Econ Rev 37(2–3):452–461

Burda MC, Härdle W, Müller M, Werwatz A (1998) Semiparametric analysis of German east–west migration intentions. Facts and theory. J Appl Econom 13(5):525–541

Camerer C, Lovallo D (1999) Overconfidence and excess entry: an empirical approach. Am Econ Rev 89(1):306–318
Falck O, Heblich S, Luedemann E (2009) Identity and entrepreneurship. CESifo working paper (2661), Munich
Fölster S (2002) Do lower taxes stimulate self-employment? Small Bus Econ 19(2):135–145
Frijters P, Haisken-DeNew J, Shields MA (2004) Money does matter! Evidence from increasing real income and life satisfaction in east Germany following reunification. Am Econ Rev 94(3):730–740
Fritsch M (2004) Entrepreneurship, entry and performance of new businesses compared in two growth regimes: East and West Germany. J Evol Econ 14(5):525–542
Fritsch M, Brixy U (2004) The establishment file of the German social insurance statistics. Schmollers Jahrbuch/J Appl Soc Sci Stud 124(1):183–190
Fuchs-Schündeln N (2008) The response of household saving to the large shock of German reunification. Am Econ Rev 98(5):1798–1828
Fuchs-Schündeln N, Schündeln M (2005) Precautionary savings and self-selection. Evidence from the German reunification 'experiment'. Q J Econ 120(3):1085–1120
Gilson RJ (1999) The legal infrastructure of high technology industrial districts: Silicon Valley, route 128, and covenants cot to compete. New York Univ Law Rev 74(3):575–628
Guiso L, Sapienza P, Zingales L (2004) The role of social capital in financial development. Am Econ Rev 94(3):526–556
Halaby CN (2003) Where job values come from: family and schooling background, cognitive ability, and gender. Am Sociol Rev 68(2):251–278
Hout M (1984) Status, autonomy and training in occupational mobility. Am J Sociol 89(6):1379–1409
Hunt J (2006) Staunching emigration from east Germany. Age and the determinants of migration. J Eur Econ Assoc 4(5):1014–1037
Kihlstrom RE, Laffont J-J (1979) A general equilibrium entrepreneurial theory of firm formation based on risk aversion. J Polit Econ 87(4):719–748
Kirzner IM (1973) Competition and entrepreneurship. University of Chicago Press, Chicago
Klepper S (2009) Silicon Valley–a chip off the old Detroit bloc. In: Audretsch DB, Strom R (eds) Entrepreneurship, growth, and public policy. Cambridge University Press, Cambridge, UK, pp 79–115
Lazear EP (2005) Entrepreneurship. J Labor Econ 23(4):649–680
Michelacci C, Silva O (2007) Why so many local entrepreneurs? Rev Econ Stat 89(4):615–633
Miller DR, Swanson GE (1958) The changing American parent. A study in the Detroit area. Wiley, New York
Moore JH (1992) Measuring soviet economic growth: old problems and new complication. J Inst Theor Econ 148(1):72–92
Mortimer JT, Lorence J (1979) Work experience and occupational value socialization. A longitudinal study. Am J Sociol 84(6):1361–1385
North D (1990) Institutions, institutional change, and economic performance. Harvard University Press, Cambridge, MA
North D (1991) Institutions. J Econ Perspect 5(1):97–112
OECD (Organisation for Economic Co-Operation and Development) (1998) Fostering entrepreneurship. OECD, Paris
Parker SC (2009) The economics of entrepreneurship. Cambridge University Press, Cambridge
Parker SC, Robson MT (2004) Explaining international variations in self-employment. Evidence from a panel of OECD countries. Southern Econ J 71(2):287–301
Rainer H, Siedler T (2009) Does democracy foster trust? J Comp Econ 37(2):251–269
Reynolds PD, Bosma N, Autio E, Hunt S, de Bono N, Servais I, Lopez-Garcia P, Chin N (2005) Global entrepreneurship monitor. Data collection design and implementation 1998–2003. Small Bus Econ 24(3):205–231

Sanders JM, Nee V (1996) Immigrant self-employment: the family as social capital and the value of human capital. Am Sociol Rev 61(2):231–249

Saxenian A (1994) Regional advantage: culture and competition in Silicon Valley and route 128. Harvard University Press, Cambridge, MA

Schindele Y (2010) How long does it take to become an entrepreneurial society? The case of German convergence in self-employment. Jena Economic Research Papers, #2010–2015

Schumpeter JA (1912) The theory of economic development. Oxford University Press, New York

Sinn G, Sinn HW (1992) Jumpstart. The economic reunification of Germany. MIT Press, Cambridge

Stuart TE, Sorenson O (2005) Social networks and entrepreneurship. In: Alvarez S, Agarwal R, Sorenson O (eds) The handbook of entrepreneurship: disciplinary perspectives. Springer, Berlin, pp 211–228

Tabellini G (2010) Culture and institutions: economic development in the regions of Europe. J Eur Econ Assoc 8(4):677–716

Terwey M, Bens A, Baltzer S (2007) Datenhandbuch ALLBUS 1980–2006, ZA-No. 4241. GESIS, Cologne

VanArk B (1995) The manufacturing sector in East Germany. A reassessment of comparative productivity performance, 1950–1988. Jahrbuch für Wirtschaftsgeschichte 2:75–100

VanArk B (1996) Convergence and divergence in the European periphery. Productivity in eastern and southern Europe in retrospect. In: VanArk B, Crafts N (eds) Quantitative aspects of post war European economic growth. Cambridge University Press, Cambridge, pp 271–326

Academic Entrepreneurship and the Geography of University Knowledge Flows in the UK

Maria Abreu and Vadim Grinevich

Abstract Knowledge flows between universities and industry have long been recognised as important determinants of regional economic development. However, a major unresolved issue is the exact nature of the mechanisms through which knowledge flows take place. We analyse the spatial patterns of knowledge exchange activities by considering a wide range of mechanisms including joint research, consultancy services, personnel exchange and informal advice. The analysis is based on a recently completed survey of UK academics, providing micro-data on over 22,000 academics across all subject areas. Our results show that the geography of academic entrepreneurship varies widely by type of activity.

Keywords Geography of knowledge • University-business links • Technology transfer • Regional innovation

1 Introduction

Knowledge flows between universities and industry have long been recognised as important determinants of regional economic development (Jaffe 1989; Feldman 1994; Anselin et al. 1997; Henderson et al. 1998; Cooke 2001, 2002; Arundel and Geuna 2004). However, a major unresolved issue is the exact nature of the mechanisms through which knowledge flows take place (Breschi and Lissoni 2001). A key figure is the academic entrepreneur, an individual who makes knowledge

JEL codes: O31, O32, R11

M. Abreu (✉)
Pembroke College and Department of Land Economy, University of Cambridge, UK
e-mail: ma405@cam.ac.uk

V. Grinevich
Department of Architecture, University of Cambridge and School of Business, Leadership and Enterprise, University Campus Suffolk, UK

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_10,

transfer possible by generating value for his/her research outside of academia (Audretsch and Lehmann 2005). The focus of the literature has generally been on a narrow number of knowledge transfer mechanisms including patenting, the creation of spin-out companies, and the licensing of research outputs. This is partly because these are easy to measure outputs and the link to economic impact is fairly straightforward. The literature has also tended to abstract from the characteristics of the individual academic entrepreneur and has instead focused on institutional determinants such as university policy and incentive systems (Friedman and Silberman 2003; Powers and McDougall 2005), the role of intermediaries such as technology transfer offices (TTOs) and incubators (Collins and Wakoh 2000; Del Campo et al. 1999; Markman et al. 2004), and cultural norms (Christman et al. 1995; Louis et al. 2001).

Our aim is to fill some of the gaps in the literature by exploring the geography of specific mechanisms of academic entrepreneurship, while accounting for both individual and institutional determinants. We analyse the spatial patterns of knowledge exchange activities by considering a wide range of mechanisms including joint research, consultancy services, personnel exchange and informal advice. Rather than restrict the analysis to science and technology, we also consider other disciplines including the health sciences, social sciences, humanities and the creative arts. The analysis is based on a recently completed survey of UK academics, providing micro-data on over 22,000 academics from across the entire higher education sector in the UK.

Our analysis aims to show how individual and institutional characteristics affect the geography at which knowledge flows take place. Implicit and explicit cultural factors are important determinants of entrepreneurship more generally (Bauernschuster et al. 2011; Bellini et al. 2011), but their relevance has not previously considered in the context of the geography of academic entrepreneurship. We show that the geography of academic entrepreneurship varies by type of activity, and that individual and institutional characteristics affect the type of entrepreneurial activities that an individual academic is involved in, and therefore the geography at which knowledge flows occur. In doing so we are also aiming to show how institutions shape the geography of innovation, and therefore ultimately the geography of economic performance.

The remainder of the chapter is organised as follows. Section 2 provides an overview of the relevant literature and sets out our conceptual framework. Section 3 describes the data and methodology used in the analysis, and the variables included in the model. Section 4 presents descriptive statistics on the geography of academic entrepreneurship activities, while Sect. 5 discusses the regression results. Section 6 summarises the findings and concludes.

2 Knowledge Flows and Space

A large body of empirical evidence has shown that knowledge flows from university research tend to be geographically localised. In a hugely influential study, Jaffe (1989) shows that patenting activity increases with university research expenditures within the same US state, although the geographical localisation effect varies substantially by technological area. Jaffe's conclusions are reinforced by Acs et al. (1992), who use a similar methodology but a different measure of innovative activity: the number of innovations recorded in leading technology, engineering, and trade journals in each manufacturing industry. Interestingly, the impact of university research turns out to be greater for innovation counts than for patenting, showing that localised spillovers are of a greater magnitude than is suggested by patenting alone.

Using the same knowledge production function approach, Anselin et al. (1997) provide further evidence on the positive relationship between university research and innovative activity. In contrast to previous studies, the authors use data on innovation counts and R&D employment at both the state level and the level of Metropolitan Statistical Area (MSA), a more disaggregated unit of analysis. The results show that the effects of university research on innovative activities by firms occur at distances of up to 50 miles. Using an alternative measure of productivity in the form of patent citations, Jaffe et al. (1993) also find strong evidence of localisation effects, with patent citations more likely to originate from the same state and MSA as the cited patents, after controlling for the pre-existing concentration of related research activity using a set of "control patents". Most of these studies show that the extent of localisation is strongly contingent on academic discipline, technological area or industrial sector (Jaffe 1989; Mansfield 1991, 1995, 1998; Thompson and Fox-Kean 2005; Abramovsky et al. 2007).

A major unresolved issue in the literature is the exact nature of the mechanisms through which knowledge flows take place. In his pioneering study of patenting activity, Jaffe (1989) acknowledges that the spillover mechanisms have not been modelled explicitly, and that the results cannot therefore be interpreted structurally. In a similar vein, Breschi and Lissoni (2001) argue for a re-examination of the evidence on localised knowledge spillovers, and in particular, of the many different knowledge transmission mechanisms that "link together universities, public laboratories, private companies and individuals, each of them serving different purposes and, as a consequence being affected in different ways by physical distances", in order to better understand the spatial bounds of knowledge flows. In this context, the key to understanding the nature of the knowledge spillover mechanisms, and hence their spatial dimension, is the individual academic entrepreneur, who is engaged with external organisations within the context of his/her institutional setting. The extent to which knowledge flows occur is dependent on individual characteristics (such as gender, age, academic position, subject area, previous experience) and institutional characteristics (such as support facilities, rules governing commercialisation activities and culture).

The academic entrepreneur has generally been defined narrowly as an individual whose activities extend beyond research and teaching roles and into commercialisation (Audretsch et al. 2005). However, recent attempts have been made to broaden this definition, to account for the fact that economic value (resulting from knowledge exchange activities) can occur through less formal activities such as joint research, consultancy, external teaching and training, and prototyping and testing (Klofsten and Jones-Evans 2000; Link et al. 2007; D'Este and Iammarino 2010). Greater incidence of these less formal activities may be partly a reaction to the perceived bureaucratic inflexibility of intermediaries such as Technology Transfer Offices (TTOs), which has had the perverse effect of promoting the use of less formal channels in order to circumvent the restrictions of the formal commercialisation process (Link et al. 2007). Formal activities are also not good indicators of the extent of academic entrepreneurship in the social sciences, humanities and the creative arts, where patenting of research outputs is much less common. In our analysis we therefore extend the definition to also include informal mechanisms of knowledge exchange such as networking, organising student placements and providing informal advice which, although less structured, still provide the means for university knowledge to reach external organisations and in the process generate economic value.

3 Data and Methods

Our analysis is based on a large-scale survey of UK academics, conducted over 2008–09 as part of an ESRC-funded project based at the Centre for Business Research, University of Cambridge (Abreu et al. 2009).[1] The sampling frame for the survey included all academics based at UK higher education institutions who at the time of asking were involved in teaching and/or research. Because there is no unified listing of academic staff active in the UK, the sampling frame was constructed using e-mail addresses available on university websites, and the survey was administered through an online web-survey tool. The total number of survey recipients was 126,120, and the achieved sample was 22,556, which also includes a number of paper-based questionnaires, resulting in a response rate of 17.8 %.[2] As far as we are aware, this is the first survey of its kind to cover all disciplines, institutions and job categories within a country's higher education sector. The survey includes questions on activities with private, public and third sector organisations, individual characteristics, views on the benefits and difficulties of

[1] The survey was part of a 2-year project on "University-Industry Knowledge Exchange: Demand Pull, Supply Push and the Public Space Role of Higher Education Institutions in the UK Regions", with grant number ESRC- RES-171-25-0018, which was carried out over 2007–2009. The project involved Maria Abreu, Vadim Grinevich, Alan Hughes and Michael Kitson, and was sponsored by Economic and Social Research Council (ESRC) in partnership with the Scottish Funding Council (SFC), Department for Employment and Learning (DEL) in Northern Ireland, the Higher Education Funding Council for England (HEFCE) and the Higher Education Funding Council for Wales (HEFCW).

[2] See Abreu et al. (2009) for further details. The data are available through the UK Data Archive (http://www.data-archive.ac.uk/) and are listed under "Cambridge Centre for Business Research Survey of Knowledge Exchange Activity by United Kingdom Academics, 2005–2009", archive no. SN 6462.

academic entrepreneurship and the geography of interactions. The questions refer to the 3 year period prior to the survey.

In addition to the survey, we use institutional data provided by the "Higher Education – Business and Community Interaction Survey 2007–08", which includes questions on funding and university resources used to support third stream activities over the period 2007–08.[3] We use a classification of universities based on a frequently made distinction between the Russell Group and 1994 Group of research-intensive universities, other old universities, post 1992 universities that have only recently achieved university status, and specialised colleges such as arts, medical and agricultural colleges. This classification, although somewhat artificial (not all Russell Group universities have a comparable research intensity, for instance), is helpful since it is often used in the higher education policy debate, and provides a useful typology for UK higher education institutions. We also allow for more subtle differences between institutions by considering a measure of the weight given to research activities (in terms of promotion), as reported by the respondents to the survey, in addition to the above classification.

3.1 Methodology

Our analysis proceeds in two steps. First, we provide an overview of the types of activities that academics are engaged in outside of the traditional academic activities of teaching and research. These activities range from participation in research consortia and joint research projects, to consultancy, informal advice and community-based activities. We also analyse the extent to which participation in these activities varies by geography.

We explore these issues in more detail using a set of ordered probit regressions. Our dependent variable is an ordinal measure of the greatest geographical distance at which an individual academic interacts with an external partner, out of local, regional, national or overseas. We run a series of models, including one for "all activities", and separate regressions for a few activities of interest, ranging from the formal (joint research) to the more informal (informal advice). We define local as occurring within 10 miles of the university; regional as occurring within the same NUTS 1 region (but beyond 10 miles); national as occurring within the UK level (but outside of the region); and overseas as occurring at a location outside of the UK. If an academic is engaged in activities at multiple geographies, we record the highest geography at which a specific activity takes place. Our definition is based on administrative geography, except for the lowest level, which is distance-based. The reason for this choice is to avoid a situation where an activity occurs at a location

[3] The "Higher Education – Business and Community Interaction Survey 2007–08" data are available through the Higher Education Funding Council for England (http://www.hefce.ac.uk/pubs/hefce/2009/09_23).

just over a regional border, but which is in fact very close to the academic entrepreneur (within 10 miles).

Our ordered probit model is based on ordinal responses to questions about the geography of specific activities in the survey. These are coded 1 for local, 2 for regional, 3 for national and 4 for overseas. We denote y_i^* the unobserved distance at which the interaction takes place, with y_i its observable ordinal counterpart. The geography of the activities can be expressed as a function of a vector of explanatory variables (X_i), containing individual and institutional determinants. We assume the model follows a linear relationship:

$$y_i^* = X_i{}'\beta + u_i \text{ where } u_i \sim N(0, \sigma^2) \tag{1}$$

The latent measure of distance y_i^* is related to the observable ordinal variable y_i as follows:

$$\begin{aligned} y_i &= 1 \text{ if } y_i^* \le \lambda_1 \\ y_i &= 2 \text{ if } \lambda_1 < y_i^* \le \lambda_2 \\ y_i &= 3 \text{ if } \lambda_2 < y_i^* \le \lambda_3 \\ y_i &= 4 \text{ if } y_i^* > \lambda_3 \end{aligned} \tag{2}$$

where $\lambda_1 < \lambda_2 < \lambda_3 < \lambda_4$ are unknown cut-off points (or threshold parameters). The parameter vectors β and λ can be estimated using maximum likelihood, where the likelihood function:

$$\begin{aligned} l_i(\lambda, \beta) = {} & 1[y_i = 1] \log[\Phi(\lambda_1 - X_i\beta] \\ & + 1[y_i = 2] \log[\Phi(\lambda_2 - X_i\beta) - \Phi(\lambda_1 - X_i\beta \\ & + 1[y_i = 3] \log[\Phi(\lambda_3 - X_i\beta) - \Phi(\lambda_2 - X_i\beta] \\ & + 1[y_i = 4] \log[1 - \Phi(\lambda_3 - X_i\beta] \end{aligned} \tag{3}$$

and Φ denotes the cumulative distribution function of the standard normal. In order to ensure identification, we omit a constant term from the vector of explanatory variables X_i. We also correct the variance-covariance matrix for heteroscedasticity of an unknown form using the robust (or sandwich) estimator of variance.

3.2 Variables Included in the Analysis

Our analysis seeks to understand the spatial dimension of academic activities with external individuals and organisations, and in particular, whether the distance at which these activities take place is affected by individual and institutional factors. Our analysis is therefore structured around the findings of the literature on the determinants of academic entrepreneurship. Although the latter has mostly focused on the institutional level, micro-level have become more widespread as the quality and availability of micro-data sources has improved.

3.2.1 Individual Determinants

Previous research has shown that individual characteristics can have a significant impact on the extent of academic entrepreneurship (Rothaermel et al. 2007). These findings can be grouped into three categories; those related to demographic variables, those that are specific to the discipline or type of research undertaken, and those related to current and past experience.

Of the demographic variables, the most widely researched factor is the academic life cycle. The literature has identified two forces at work. On the one hand, academics are more focused on publishing when they are younger or less established, as they seek to achieve tenure and become well known in their fields, while older academics are more likely to capitalise on their reputation in order to commercialise (Carayol 2007; Levin and Stephan 1991; Stephan et al. 2007). On the other hand, younger cohorts of academics are more familiar with commercialisation, and therefore are more comfortable with the compromises involved (Azoulay et al. 2007). Empirical findings on this topic have been inconclusive, with studies finding positive, negative, or insignificant effects of life cycle on academic entrepreneurship (Azoulay et al. 2007; Morgan et al. 2001; Stephan et al. 2007; Ambos et al. 2008; Link et al. 2007) while others identify an inverted-u shaped relationship (Levin and Stephan 1991; Thursby and Thursby 2005). Studies that include both age and academic position have found that age has a negative effect, while position has a positive effect, with professors more likely to engage with external organisations, regardless of their age (Bercovitz and Feldman 2003; D'Este and Perkmann 2011).

A second demographic variable of interest is gender, with most studies finding that female academics are less likely to engage with external organisations, particularly with respect to formal activities such as patenting and licensing (Ding et al. 2006; Thursby and Thursby 2005; Whittington and Smith-Doerr 2005). This may be due to a number of reasons, including female academics being less likely to have commercial sector experience, being less likely to belong to networks that also include industrial partners, and being are more likely to feel ambivalent about the ethics and benefits of commercialisation (Murray and Graham 2007).

The second set of factors is related to the type of research an individual academic is engaged in. Some academic disciplines such as the life sciences naturally combine fundamental and applied work in a way that allows applications to directly follow from university research (Murray 2002; Stephan et al. 2007), while in other disciplines such as theoretical physics a substantial amount of additional work is required before the results can be commercialised. These findings are closely related to a categorisation developed by Stokes (1997), who makes a distinction between pure basic research, user-inspired basic research and applied research. We follow the Stokes (1997) classification in this chapter, and also control for other unobserved differences related to cultural factors by including subject dummies in the regression analysis.

A final category is previous and current entrepreneurial experience. The literature on entrepreneurship has found that prior experience, such as having started or owned a business, or having a close relative who owns a business, have a positive effect on an individual's entrepreneurial behaviour (Klofsten and Jones-Evans 2000). The evidence with respect to current academic activities is less clear cut. A more research intensive position might entail greater pressures to publish rather than commercialise, but this relationship may be changing over time, with the proliferation of research-only roles in university-affiliated research centres that are also partly-funded by external organisations (Carayol 2007). We analyse the impact of these factors by including variables to capture previous experience and current activities such as teaching only, research only, or both research and teaching.

There are several reasons to expect these findings to differ across activities and geographies. We would expect more established academics to engage in informal activities, such as sitting on advisory boards or providing informal advice, at higher frequency and at greater distances than younger or less established academics, since engaging in these activities is partly dependent on being well known outside their immediate communities. Likewise, we would expect individuals with more prior experience of business or industry to be able to more easily engage in formal activities that involve complex negotiations or access to finance, such as research consortia or joint research. The role of belonging to a particular discipline is also interesting; some subject areas that are closely linked to public sector work, such as education and the health sciences, would be more likely to involve collaborations with local or regional partners, while others based on large-scale industry, such as engineering, would be more likely to lead to national or overseas collaborations.

3.2.2 Institutional Determinants

With respect to institutional determinants, the literature on academic entrepreneurship has focused mainly on the role of the Technology Transfer Office (TTO), which is both in charge of protecting the institution's intellectual property, and helping academic staff to capitalise on their research. The dual role of the TTO creates a complex set of incentives, whereby academic staff must decide whether to disclose their findings to the TTO, and the TTO must subsequently decide whether to commercialise them, and how to negotiate with potential users (Jensen and Thursby 2003; Siegel et al. 2007). The high levels of bureaucracy and inflexible administrative rules of a large number of TTOs can lead to academic staff opting to disclose their findings to potential users through informal channels, in return for access to equipment, data or student placements (Siegel et al. 2004; Thursby et al. 2001). A related issue is the prevalent culture towards commercialisation and entrepreneurship within a department or higher education institution, regardless of the official line on research commercialisation. A divide may also exist between research-intensive universities and institutions with a greater focus on teaching, although this divide may manifest itself through engagement in different types of activities, with the more research-intensive universities focusing joint research with leading firms and non-commercial concerns at greater distances, while more teaching-intensive universities may be more likely to work with local business in developing teaching curricula or arranging student placements.

A second area of interest is in the composition of the TTO, in terms of its staff and the facilities provided to academic entrepreneurs. The literature has found that academics are generally dissatisfied with the level of bureaucracy and skills of TTO staff (Link et al. 2007; Siegel et al. 2004). The ideal situation is one where there is flexibility, allowing academics to capitalise on their research in different ways, and where the TTO has the mix of skills necessary to help them in this endeavour. Having a compulsory contracting system may therefore be counter-productive, while providing access to commercialisation facilities without a compulsion to use them is likely to be beneficial (Link et al. 2007). While these factors have been shown to affect the incidence of academic entrepreneurship, they may also affect the variety of activities undertaken and the distance at which these take place, since greater flexibility allows a departure from the established norm.

4 Geography of Academic Entrepreneurship Activities

In order to analyse the geography of academic activities, we have re-classified the 22 activities covered in the survey into a smaller number of mechanisms (see Table 1 for a classification). The activities are classified into *joint R&D*, capturing formalised interactions where both partners are involved in the research; *personnel exchange*, denoting activities that involve the physical relocation of staff (and thus of human capital); *networks*, which covers conferences, lectures and other activities that involve meeting individuals from external organisations; *contract R&D*, where research occurs on the academic side only; *advisory roles*, covering informal knowledge exchange activities; and *community engagement*, which includes any activities that involve showing or involving the community in the findings of academic research.

As Fig. 1 shows, the extent of academic activities varies in both scale and geography with the type of activity.[4] The most frequent activities are the two most informal ones, networks (90 %) and advisory roles (77 %), followed by the most formal, joint R&D (61 %). The graph also shows that, contrary to public perception, a significant percentage of academics participate in community

[4] Because the survey allows respondents to choose more than one geography for each type of activity, Fig. 1 has been drawn so that the total percentage for each activity is the percentage of academics engaged in that activity at any geographical level, while the breakdown by geography shows the percentage of those engaged in that activity who indicated a specific geography.

Table 1 List of academic activities with external organisations

Type	Activities
Joint R&D	Research consortia, joint publications, joint research
Personnel exchange	Hosting of personnel, secondment, student placements
Networking	Attending conferences, invited lectures/seminars, networks
Contract R&D	Prototyping and testing, contract research
Advisory roles	Joint curriculum design, consultancy services, sitting on advisory boards, standard setting forums, informal advice, personnel training
Community engagement	Funding for physical facilities, public lectures, public exhibitions, enterprise education, school-based projects

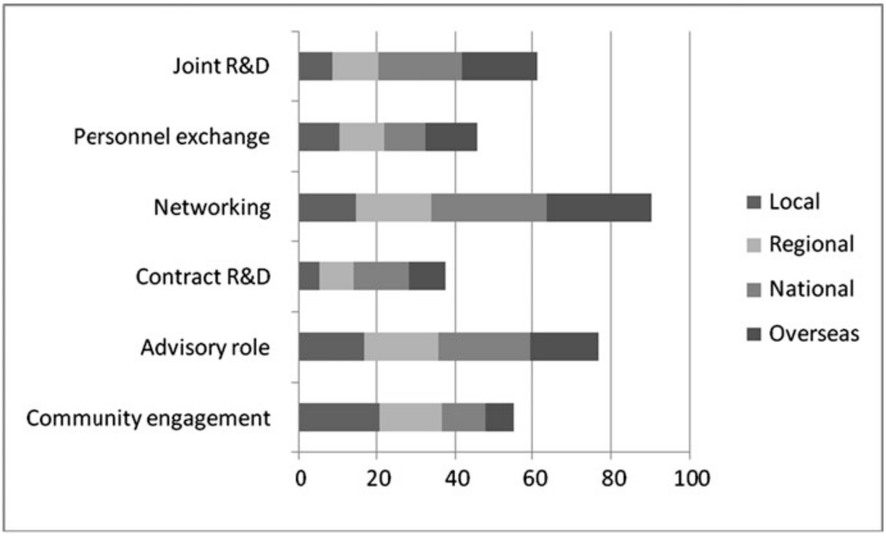


Fig. 1 Geography of academic activities, by broad activity

engagement activities (55 %). The geography at which the activities occur can be seen more easily as percentages of the total for each activity (Fig. 2). Activities involving formal research collaborations occur more frequently at larger geographies, perhaps because the stakes are higher, so those involved will look for the ideal partner, regardless of location. Second most distant are those involving movements of people, such as personnel exchange and networks. Most local are those involving informal advisory roles and community engagement.

We hypothesised in Sect. 3.3 that academic entrepreneurship is likely vary by discipline. Figure 3 shows the extent and geography of academic activities by the individual's discipline. As expected, the more distant interactions occurred for academics in the biological sciences, engineering and the physical sciences, where more than half of all activities where at the national or overseas level. In contrast, most interactions for academics in the health sciences, creative arts and

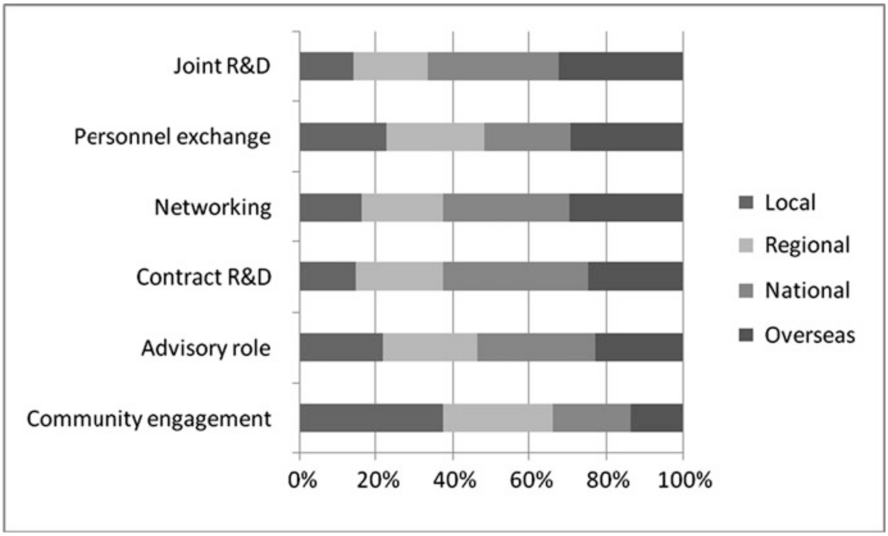


Fig. 2 Geography of academic activities, by broad activity (% of total)

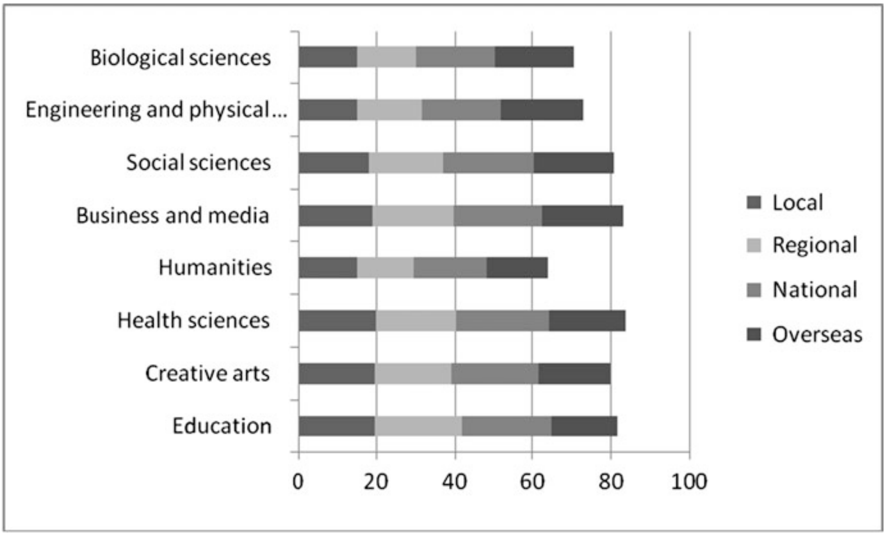


Fig. 3 Geography of academic activities, by discipline

education occurred at the local or regional level. In terms of overall levels of academic entrepreneurship, not surprisingly, the lowest levels were for the humanities (64 %), while the highest were for the business and media (83 %), the health sciences (83 %) and the social sciences (80 %).

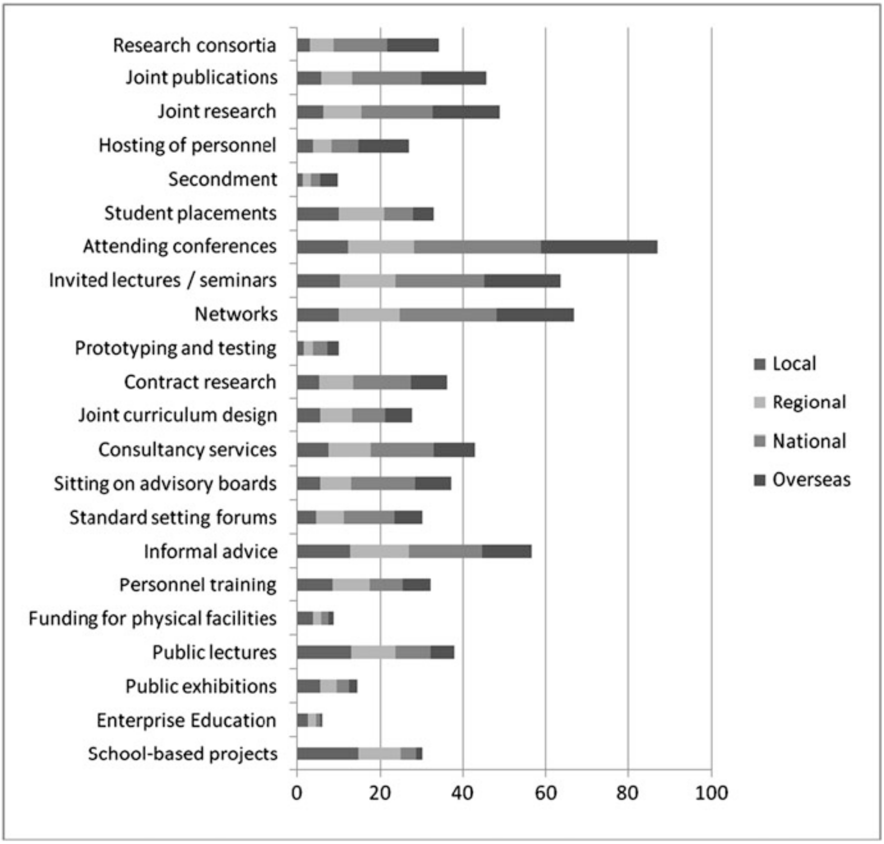


Fig. 4 Academic activities by type of disaggregated activity

The geography of activity, while fairly consistent at the broader activity level, also varies at a more disaggregated level (Fig. 4). Among personnel exchange activities, the movement of more experienced staff (in the form of *hosting of personnel* or *secondment*) occurs at greater geographies than *student placements*. Similarly, involvement in *school-based projects* is more likely to occur on a local level than other forms of community engagement activities, such as *public lectures*.

5 Geography of Formal and Informal Activities

The descriptive analysis discussed in Sect. 4 strongly suggests that the geography of academic activities with external partners varies by type of activity. We now take this finding a step further, and analyse whether the effect of variables traditionally

Table 2 Ordered probit regressions for the geography of academic activities

	All activities	Joint research	Consultancy services	Informal advice
Individual characteristics				
Age: under 30†				
Age: 30–39	0.080*	0.045	0.103	0.195***
	(0.045)	(0.059)	(0.078)	(0.063)
Age: 40–49	−0.016	−0.050	−0.023	0.122*
	(0.047)	(0.060)	(0.078)	(0.064)
Age: 50 and over	−0.167***	−0.143**	−0.053	0.050
	(0.048)	(0.062)	(0.078)	(0.064)
Position: assistant staff†				
Position: researcher	0.271***	0.172**	0.683***	0.472***
	(0.054)	(0.070)	(0.090)	(0.075)
Position: lecturer	0.185***	0.028	0.307***	0.150**
	(0.055)	(0.074)	(0.088)	(0.073)
Position: reader, senior staff	0.395***	0.208***	0.182**	−0.026
	(0.055)	(0.073)	(0.089)	(0.074)
Position: professor	0.878***	0.575***	0.333***	0.168**
	(0.059)	(0.076)	(0.089)	(0.073)
Female	−0.153***	−0.187***	−0.189***	−0.157***
	(0.020)	(0.026)	(0.027)	(0.023)
Subject: health sciences†				
Subject: biological sciences	0.422***	0.318***	0.115**	0.331***
	(0.041)	(0.045)	(0.054)	(0.045)
Subject: Eng. and physical sciences	0.397***	0.330***	−0.052	0.097***
	(0.035)	(0.037)	(0.041)	(0.037)
Subject: social sciences	−0.018	0.067*	0.042	−0.023
	(0.030)	(0.037)	(0.039)	(0.035)
Subject: business and media	0.091**	0.046	−0.019	−0.095**
	(0.044)	(0.051)	(0.051)	(0.047)
Subject: humanities	−0.106***	0.222***	0.088	0.089**
	(0.037)	(0.060)	(0.053)	(0.045)
Subject: creative arts	0.126**	−0.004	−0.116*	0.037
	(0.050)	(0.067)	(0.060)	(0.053)
Subject: education	−0.117***	0.033	−0.076	−0.066
	(0.039)	(0.050)	(0.049)	(0.044)
Basic research†				
User-inspired research	0.288***	−0.085**	0.104***	0.118***
	(0.028)	(0.038)	(0.039)	(0.033)
Applied research	0.328***	−0.136***	0.080**	0.089***
	(0.026)	(0.036)	(0.037)	(0.032)
Other type of research	0.032	−0.170*	−0.024	−0.027
	(0.052)	(0.090)	(0.076)	(0.064)
Employed in small company	0.078***	0.026	−0.011	0.026
	(0.026)	(0.030)	(0.030)	(0.027)
Owned small company	0.170***	0.025	0.121***	0.106***
	(0.031)	(0.034)	(0.032)	(0.030)

(continued)

Table 2 (continued)

	All activities	Joint research	Consultancy services	Informal advice
Employed in large company	0.003	0.007	0.052*	−0.026
	(0.026)	(0.030)	(0.030)	(0.028)
Employed in public sector	−0.028	−0.012	−0.042	−0.033
	(0.025)	(0.029)	(0.030)	(0.027)
Employed in third sector	0.062**	0.056*	0.122***	0.095***
	(0.027)	(0.033)	(0.033)	(0.029)
Not previously employed	−0.072**	0.095**	0.005	0.054
	(0.031)	(0.037)	(0.039)	(0.034)
Research only	−0.039	0.061*	0.134***	0.071*
	(0.033)	(0.036)	(0.044)	(0.037)
Teaching only	−0.393***	−0.094	−0.062	−0.279***
	(0.060)	(0.124)	(0.088)	(0.083)
Both research and teaching†				
Weight of research	0.247***	0.243***	0.268***	0.279***
	(0.036)	(0.046)	(0.047)	(0.040)
Weight of commercialisation	−0.090**	−0.077*	−0.081*	−0.129***
	(0.037)	(0.043)	(0.045)	(0.041)
HE type: Russell group	0.135***	0.068*	0.115***	0.094***
	(0.032)	(0.037)	(0.039)	(0.034)
HE type: 1994 group	0.008	−0.020	0.094**	0.076*
	(0.037)	(0.045)	(0.046)	(0.041)
HE type: other old†				
HE type: post 1992	−0.006	−0.068	−0.025	0.022
	(0.042)	(0.051)	(0.051)	(0.045)
HE type: specialised college	0.113	0.189**	0.300***	0.274***
	(0.070)	(0.087)	(0.094)	(0.079)
Licensing capability	0.075	0.038	0.082	0.100*
	(0.047)	(0.065)	(0.062)	(0.056)
Commercialisation facilities	0.040	0.184*	−0.056*	0.083
	(0.091)	(0.104)	(0.034)	(0.099)
Contracting system	−0.036	−0.056*	−0.152**	−0.058*
	(0.027)	(0.033)	(0.070)	(0.030)
Staff indemnity insurance	−0.067	−0.030	0.135	−0.046
	(0.053)	(0.067)	(0.111)	(0.062)
λ_1	−0.530**	−0.243	0.165	0.221
	(0.235)	(0.297)	(0.313)	(0.268)
λ_2	−0.005	0.411	0.939***	0.918***
	(0.234)	(0.297)	(0.313)	(0.268)
λ_3	0.855***	1.236***	1.886***	1.755***
	(0.234)	(0.297)	(0.313)	(0.269)
Chi-square	1883.80***	985.86***	944.23***	1013.76***
Observations	19,464	9,850	8,358	10,905

Note: Standard errors in *parentheses*. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. †Reference category. Standard errors are estimated using the robust or sandwich estimator of variance. The dependent variable is an ordinal measure of the greatest distance at which an activity takes place: 1 = local, 2 = regional, 3 = national and 4 = overseas found to affect academic entrepreneurship also affect the distance at which interactions take place.

We run four ordered probit models, shown on Table 2. The first column shows the results for the geography of academic entrepreneurship activities, where the dependent variable captures the greatest geography at which an individual academic interacts with external organisations, for *all* academic entrepreneurship activities. The results for the demographic variables are generally in line with those of the literature on entrepreneurship; older academics are less likely to be involved in external activities at greater distances (possibly due to a cohort effect), as are female academics. Position, on the other hand, has a positive association with distance, and professors are much more likely than junior staff to be involved in activities overseas.

The academic discipline and type of research also have a large effect on the distance of academic activities. Academics in the biological sciences, engineering and physical sciences, business and media and, somewhat surprisingly, the creative arts are all more likely to interact at greater distances, while those in the humanities and education are more likely to interact locally. The type of research is also important, with those involved in user-inspired or applied research more likely to be involved at greater distances, relative to those working on basic research. The findings on previous experience show that having been employed or having owned a small company in the past has a positive association with distance, while not having previously employed has a negative association, in keeping with the entrepreneurship literature that suggest that networks, financial acumen and contacts in the business world resulting from past experience lead to greater academic activities with external organisations. In terms of academic roles, being involved in teaching only (as opposed to both teaching and research) is associated with local rather than longer-distance activities, possibly because academics involved in teaching only are more likely to be involved in student placements and community engagement activities such as giving public lectures.

In keeping with previous work, we find that institutional variables are less significant than individual variables in explaining the extent of academic entrepreneurship (Abreu and Grinevich 2010). Working in an institution that gives more weight to research activities has a positive effect on distance, while working in an institution with a greater focus on commercialisation has a negative effect, possibly because the latter institutions are likely to focus on community engagement and impact on the regional economy. Other institutional variables that generally tend to affect academic entrepreneurship are not significant in explaining distance of interactions, with the exception of belonging to a Russell Group university, which is associated with fewer local and more overseas activities.

The three remaining ordered probit models show how the coefficients change when the benchmark model in the first column is restricted to specific activities. All three of the activities chosen, joint research, consultancy services and informal advice, are research related, but range from the formal to the more informal. This allows us to identify whether and how the coefficients vary as the type of activity varies. The first interesting result is related to career life cycle; across all activities older academics are less likely to interact at greater distances (possibly a cohort effect), while more established academics are more likely to interact at greater distances, after controlling for age. The effect of being a professor, in particular, is positive on distance for all types of activity, but the effect is larger for more informal activities. Status is therefore an important determinant of the distance at which academic activities take place, particularly for more informal activities.

A second interesting finding concerns the type of research. Our benchmark regression shows that being involved in more applied work is associated with interactions at greater distances, but the activity-specific models show that this is only the case if the activity is informal. The more formal the activity, the less likely it is that interactions occur at a greater distance, or in other words, joint research projects are more likely to occur on a local level if the research is applied, and on a national or overseas level if the research is basic, while the opposite is true for informal advice. This finding is consistent with other results in the literature, in particular Mansfield (1991, 1995, 1998).

The research intensity of the institution also follows an interesting pattern, with academics located at more research-intensive universities being more likely to interact at greater distances, with the effect at its strongest for more informal activities. A similar finding also holds for academics belonging to specific types of universities, such as Russell Group universities or specialised colleges. The availability of commercialisation facilities is associated with greater distances for more formal activities, but lower distances for informal activities, while the existence of a compulsory contracting system restricts the interactions to local activities, for all types of activity.

6 Conclusions

In this chapter we explore the individual and institutional determinants of academic entrepreneurship, and analyse how these factors affect the geography at which academic activities with external organisations and individuals take place. The literature on knowledge spillovers has generally found that knowledge flows are bounded in space, but these findings are driven by significant differences across sectors, technological areas and academic disciplines. The literature has also tended to overlook the specific mechanisms through which knowledge transfer occurs, abstracting from the significant complexity of university-business activities. Our aim is to fill this gap by presenting a more comprehensive analysis of the geography of academic entrepreneurship and its determinants.

We use data from a unique large-scale survey of academics, covering over 22,000 academics from all disciplines and higher education institutions in the UK, to analyse academic entrepreneurship, broadly defined to include any activity that allows an academic to generate value for his/her research outside academia. Our results show that the geography of academic entrepreneurship varies widely by type of activity. In particular, formal research collaborations occur at greater geographies, possibly because the stakes are higher, so that those involved will look for the ideal partner, regardless of distance. Second most distant are personnel exchange activities, although the geography for these also varies by degree of formality (type of arrangement). Activities involving senior or more established staff tend to occur at greater distances than those involving student placements. The most local activities are those related to providing informal advice and community engagement.

Our results also show that demographic and other individual factors are most relevant to explaining the extent of academic entrepreneurship activities. Older academics are less likely to be involved at greater distances, perhaps because, in keeping with the entrepreneurship literature, younger academics are more familiar with commercialisation and other activities. The academic discipline of the individual is also very relevant, with academics in the biological sciences, engineering and the physical sciences being more likely to interact with partners located overseas, while those in the humanities and education being more likely to be involved with local partners. In keeping with previous work, we also find that institutional factors are less important than individual factors in explaining academic entrepreneurship activities, with the exception of the research intensity of the institution, which has a positive association with the average distance at which activities take place.

While it is not possible to say how great the impact of the different activities is in terms of the economic value of the knowledge that is being generated, we can draw conclusions on the effects of institutional characteristics on the geography of knowledge flows. Academics based at universities that place a greater focus on research activities are more likely to engage in activities at greater distances, while a focus on commercialisation activities the focus to local interactions. This holds for all activities, but is most pronounced in the case of informal advice. Similarly, having a compulsory university contracting system is associated with more local activities, perhaps because the facilities provided are more geared towards the needs of local businesses.

The picture that emerges from individual characteristics is also interesting. After controlling for academic subject (across which there are important differences), individuals who are younger, and those who occupy more senior positions (after controlling for age) are more likely to engage in activities at greater distances, as are those who undertake applied research and who have previous entrepreneurial experience. This suggests that academics who are more successful, and whose work is in greater demand, are more likely to engage at greater distances. This in turn would seem to imply that knowledge flows from universities are not confined to the local area, but are spread more widely, although there are important differences by type of activity (less formal activities occur at smaller distances).

In order to capture the impact of these activities, more research is needed on the interaction between the availability of knowledge, the willingness of academic entrepreneurs to engage with businesses and other organisations, and crucially, the capacity of businesses and other external organisations to make use of the knowledge and adapt it for the purposes of their own innovation activities

(Iammarino et al. 2011). The use by academic entrepreneurs of multiple activities to engage with external partners also highlights the need to better understand how networks of academic and external researchers are built and maintained, and whether, for instance, they are closely related to teaching and student placements (see Crescenzi et al. 2011, on the role of social networks in innovation processes). Finally, while universities are important components of the regional innovation system, not least through their provision of human capital, the geography of knowledge flows is complex, and our findings show that university knowledge can spread at multiple geographies. The relative importance of university knowledge flows and the impact of other university activities such as teaching and training on regional innovation and total factor productivity growth processes is an important area for future research (Diliberto and Usai 2011).

References

Abramovsky L, Harrison R, Simpson H (2007) University research and the location of business R&D. Econ J 117:C114–C141

Abreu M, Grinevich (2010) Personal, institutional and spatial determinants of academic entrepreneurship in the UK. Paper presented at the 50th anniversary European Congress of the Regional Science Association International, Jönköping, Sweden, 19–23 Aug 2010

Abreu M, Grinevich V, Hughes A, Kitson M (2009) Knowledge exchange between academics and the business, public and third sectors. UK Innovation Research Centre, University of Cambridge and Imperial College, London

Acs ZJ, Audretsch D, Feldman M (1992) Real effects of academic research: comment. Am Econ Rev 82(1):363–367

Ambos TC, Mäkelä K, Birkinshaw J, D'Este P (2008) When does university research get commercialized? Creating ambidexterity in research institutions. J Manage Stud 45:1424–1447

Anselin L, Varga A, Acs ZJ (1997) Local geographic spillovers between university research and high technology innovations. J Urban Econ 42:422–448

Arundel A, Geuna A (2004) Proximity and the use of public science by innovative European firms. Econ Innov New Technol 13(6):559–580

Audretsch D, Lehmann E (2005) Does the knowledge spillover theory of entrepreneurship hold for regions? Res Policy 34:1191–1202

Audretsch D, Aldridge T, Oetti A et al (2005) The knowledge filter and economic growth: the role of scientist entrepreneurship. Discussion papers on entrepreneurship, growth and public policy N.1106. Max Planck Institute of Economics, Jena

Azoulay P, Ding W, Stuart T (2007) The determinants of faculty patenting behavior: demographics or opportunities? J Econ Behav Organ 63:599–623

Bauernschuster S, Falck O, Gold R, Heblich S et al (2011) Explicitly implicit: how institutional differences influence entrepreneurship. Ch 9 in this volume

Bellini E, Ottaviano G, Pinelli D, Prarolo G et al (2011) Cultural diversity and economic performance: evidence from European regions. Ch 8 in this volume

Bercovitz J, Feldman M (2003) Technology transfer and the academic department: who participates and why? Paper presented at the DRUID summer conference 2003, Copenhagen, 12–14 June 2003

Breschi S, Lissoni F (2001) Knowledge spillovers and local innovation systems: a critical survey. Ind Corp Change 10(4):975–1005

Carayol N (2007) Academic incentives, research organization and patenting at a large French university. Econ Innov New Technol 16:119–138
Christman JJ, Hynes T, Fraser S (1995) Faculty entrepreneurship and economic development: the case of the University of Calgary. J Buss Venturing 10:267–281
Collins S, Wakoh H (2000) Universities and technology transfer in Japan: recent reforms in historical perspective. J Technol Transf 25:213–222
Cooke P (2001) Regional innovation systems, clusters, and the knowledge economy. Ind Corp Change 10:945–974
Cooke P (2002) Knowledge economies. Clusters, learning and cooperative advantage. Routledge, London
Crescenzi R, Gagliardi L Percoco M et al (2011) Culture and innovation: evidence from Italy. Ch 10 in this volume
D'Este P, Perkmann M (2011) Why do academics engage with industry? The entrepreneurial university and individual motivations. J Technol Trans 36:316–339
D'Este P, Iammarino S (2010) The spatial profile of university-business research partnerships. Pap Reg Sci 89:335–350
Del Campo AA, Sparks A, Hill RC, Keller RT (1999) The transfer and commercialisation of university-developed medical imaging technology: opportunities and problems. IEEE Trans Eng Manage 46(3):289–299
Di Liberto A, Usai S (2011) TFP convergence across European regions: a comparative spatial dynamics analysis. Ch 5 in this volume
Ding WW, Murray F, Stuart TE (2006) Gender differences in patenting in the academic life sciences. Science 313:665–667
Feldman M (1994) The geography of innovation. Kluwer, Amsterdam
Friedman J, Silberman J (2003) University technology transfer: do incentives, management and location matter? J Technol Trans 28:17–30
Henderson R, Jaffe A, Trajtenberg M (1998) Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. Rev Econ Stat 80:119–127
Iammarino S, Piva M, Vivarelli M, von Tunzelmann N et al (2011) Firm capabilities and cooperation for innovation: Evidence from the UK regions. Ch 14 in this volume
Jaffe A (1989) Real effects of academic research. Am Econ Rev 79(5):957–970
Jaffe A, Trajtenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patents citations. Q J Econ 108(3):577–598
Jensen R, Thursby M (2003) Proofs and prototypes for sale: the licensing of university inventions. Am Econ Rev 91:240–259
Klofsten M, Jones-Evans D (2000) Comparing academic entrepreneurship in Europe – the case of Sweden and Ireland. Small Bus Econ 14:299–309
Levin SG, Stephan PE (1991) Research productivity over the life cycle: evidence for academic scientists. Am Econ Rev 81:114–132
Link AN, Siegel DS, Bozeman B (2007) An empirical analysis of the propensity of academics to engage in informal university technology transfer. Ind Corp Change 16:641–655
Louis KS, Jones LM, Anderson MS, Blumenthal D, Campbell EG (2001) Entrepreneurship, secrecy, and productivity: a comparison of clinical and non-clinical life sciences faculty. J Technol Trans 26:233–245
Mansfield E (1991) Academic research and industrial innovation. Res Policy 20:1–12
Mansfield E (1995) Academic research underlying industrial innovations: sources, characteristics and financing. Rev Econ Stat 77:55–65
Mansfield E (1998) Academic research and industrial innovation: an update of empirical findings. Res Policy 26:773–776
Markman GD, Gianiodis PT, Phan PH, Balkin DB (2004) Entrepreneurship from the ivory tower: do incentive systems matter? J Technol Transfer 29:353–364

Morgan RP, Kruytbosch C, Kannankutty N (2001) Patenting and invention activity of U.S. scientists and engineers in the academic sector: comparisons with industry. J Technol Trans 26:173–183

Murray F (2002) Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. Res Policy 31:1389–1403

Murray F, Graham L (2007) Buying science and selling science: gender differences in the market for commercial science. Ind Corp Change 16:657–689

Powers JB, McDougall PP (2005) Policy orientation effects on performance with licensing to start-ups and small companies. Res Policy 34:1028–1042

Rothaermel FT, Agung SD, Jiang L (2007) University entrepreneurship: a taxonomy of the literature. Ind Corp Change 16:691–791

Siegel D, Waldman DA, Atwater LE, Link AN (2004) Toward a model of the effective transfer of scientific knowledge from academicians to practitioners: qualitative evidence from the commercialization of university technologies. J Eng Technol Manage 21:115–142

Siegel D, Veugelers R, Wright M (2007) Technology transfer offices and commercialization of university intellectual property: performance and policy implications. Oxford Rev Econ Pol 23(4):640–660

Stephan PE, Gurmu S, Sumell A, Black G (2007) Who's patenting in the university? Evidence from a survey of doctorate recipients. Econ Innov New Technol 16:71–99

Stokes DE (1997) Pasteurs quadrant: basic science and technological innovation. Brookings Institution Press, Washington, DC

Thompson P, Fox-Kean M (2005) Patent citations and the geography of knowledge spillovers: a reassessment. Am Econ Rev 95(1):450–460

Thursby JG, Thursby MC (2005) Gender patterns of research and licensing activity of science and engineering faculty. J Technol Trans 30:343–353

Thursby JG, Jensen RA, Thursby MC (2001) Objectives, characteristics and outcomes of university licensing: a survey of major U.S. universities. J Technol Trans 26:59–70

Whittington KB, Smith-Doerr L (2005) Gender and commercial science: women's patenting in the life sciences. J Technol Trans 30:355–370

Part III
Agglomeration Economies, the Location of Economic Activities and Innovation

Evaluating the Role of Clusters for Innovation and Growth in Europe

Andrés Rodríguez-Pose and Fabrice Comptour

Abstract The analysis of clusters has attracted considerable interest over the last few decades. The articulation of clusters into complex networks and systems of innovation – generally known as regional innovation systems – has, in particular, been associated with the delivery of greater innovation and growth. However, despite the growing economic and policy relevance of clusters, little systematic research has been conducted into their association with other factors promoting innovation and economic growth. This chapter addresses this issue by looking at the relationship between innovation and economic growth in 152 regions of Europe during the period between 1995 and 2006. Using an econometric model with a static and a dynamic dimension, the results of the analysis highlight that: (a) regional growth through innovation in Europe is fundamentally connected to the presence of an adequate socioeconomic environment and, in particular, to the existence of a well-trained and educated pool of workers; (b) the presence of clusters matters for regional growth, but only in combination with a good 'social filter', and this association wanes in time; (c) more traditional R&D variables have a weak initial connection to economic development, but this connection increases over time and, is, once again, contingent on the existence of adequate socioeconomic conditions.

Keywords Clusters • Economic growth • European Union • Innovation • Regional innovation systems

---

JEL Codes: O32, O52, R11

A. Rodríguez-Pose (✉)
London School of Economics, London, UK
e-mail: a.rodriguez-pose@lse.ac.uk

F. Comptour
College of Europe, Bruges, Belgium

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_11,

1 Introduction

One of the traditional advantages associated with clusters of firms has been their capacity to engender greater innovation and to transform this innovation into economic growth (Porter 2000). Groups of firms working in the same or in closely related sectors are deemed to generate agglomeration economies and knowledge spillovers. These spillovers, in turn, are at the root of self-reinforcing processes of innovation and growth (Capello 1999). Physical proximity among firms is considered to facilitate the emergence of interaction and the formation of interpersonal and firm networks leading to the genesis of complex collective learning mechanisms (Melachroinos and Spence 2001; Storper and Venables 2004). Knowledge spillovers and collective learning mechanisms thus help transform mere clusters of firms into ‘neo-Marshallian industrial districts’ (Becattini 1987), ‘new industrial spaces’ (Scott 1988), ‘innovative milieux’ (Aydalot 1986), ‘learning regions’ (Morgan 1997), or ‘regional innovation systems’ (Cooke et al. 1997; Cooke and Morgan 1998), where firms and the territories they are located in – together with their intrinsic social and structural characteristics and interactions – are put at the centre of the innovation process and of the generation of economic growth. Hence, local social structures, interaction, and collective learning processes within clusters are viewed as making firms located in close physical proximity more innovative and more dynamic than isolated firms (Baptista and Swann 1998).

The link between clusters of firms, innovation, and economic growth has generally been based on a large number of case studies where the learning processes of firms in dense institutional environments are documented. However, as Martin and Sunley (2003, 22) acknowledge – possibly because of the constant resort to what can be considered as favourable cases – the positive connection between the presence of clusters and innovation and economic growth is far from well documented. There are relatively few studies that address the link between clusters, innovation and growth from a comparative perspective and even fewer that try to venture into quantitative analyses of a large number of territories, in order to assess whether the positive relationship between clusters, innovation, and growth found in specific cases stands the scrutiny of including not only successful clusters, but also areas a priori less prone to the emergence of collective learning process.

This chapter tries to address this gap in the literature by studying the interaction of the presence of clusters with other factors deemed to promote innovation – such as investment in research and development (R&D), patent applications, or the presence of ‘innovation prone’ socioeconomic environments – and economic growth across 152 regions located in fifteen European Union (EU) countries over the period 1995–2006. Using pooled cross-section regressions, the model intends to capture both the static and the dynamic connection between a series of innovation promoting factors grouped into three different composite variables or ‘innovation filters’ – the ‘R&D filter’, the ‘social filter’ and the ‘clusterisation index’ – specially designed in order to proxy the complex interaction among growth enhancing innovation variables.

In order to achieve this aim, the chapter is structured into five main sections. After this introduction, the analytical framework of the study is framed in the theoretical literature, paying special attention to the analysis of clusters and regional innovation systems. The third section is devoted to the question of how to operationalise the key factors emerging from the theoretical section. The fourth section presents the model and the results of both the static and dynamic analyses of the connection between different groups of innovation generating factors and economic growth in Europe. The main conclusions of the analysis are presented in the final section.

2 From Clusters to Innovation and Growth

Clusters or “the geographic concentrations of interconnected companies, specialized suppliers, service providers, firms in related industries, and associated institutions” (Porter 2008, 213) have been at the centre of much of the literature aiming to understand and describe the link between innovation and economic growth. This literature has tended to highlight the importance of the presence of agglomerations of firms, organizations, and institutional actors, located in close geographical proximity and interlocked in intricate systems of cooperation, competition, and knowledge diffusion, for the genesis and the spreading out of innovation and, subsequently, economic growth. This literature – often grouped under the label of ‘regional systems of innovation’ literature – traditionally combines two main fields of theory: the innovation systems strand, on the one hand, and insights from regional science, on the other (Doloreux and Parto 2005, 134–5). The ‘regional systems of innovation’ framework is fundamentally grounded on the innovation systems theory developed in the 1980s by evolutionary theorists (cf. Iammarino 2005). This theory is based on the idea that economic performance is not only the result of individual firms’s efforts, but also of a series of other factors, external to the firm, that create an environment that is more or less prone to innovation and economic growth (Dosi 1988). Whether any particular territory is capable of becoming more innovative and, as a result, more dynamic, depends on the presence of a complex system of “inter-organisation networks, financial and legal institutions, technical agencies and research infrastructures, education and training systems, governance structures, innovation policies, etc.” (Iammarino 2005, 499). This implies not only the co-location of firms and related industries, but also a degree of specialization combined with a certain level of scope or breadth across a range of industries included in the cluster and a minimum scale or critical mass of firms (Spencer et al. 2010, 702). The capacity of any territory to innovate and grow is considered to be closely dependent on the presence of these regional or local systems of innovation.

Regional science complements the regional innovation systems approach by bringing the specificities of the regional scale to the fore. By taking into account “the internal and dynamics regularities of territorially embedded socio-economic structures” (Iammarino 2005, 501), regional science allows to determine to what extent specific regions are genuine ‘loci of innovation’ (Doloreux and Parto 2005, 135). This is achieved by focusing on two aspects: first, physical proximity among economic actors as a driver of innovation and, second, the idiosyncratic and innovation enhancing characteristics of a region.

Physical proximity is often regarded as the key aspect making some regions genuine ‘loci of innovation’. The basic reasoning is that innovation travels with difficulty and suffers from strong distance decay effects. Indeed, most analyses looking at the geographical diffusion of knowledge spillovers have highlighted that these knowledge effects are neither felt beyond the boundaries of the functional metropolitan region, in the case of the US (Anselin et al. 1997; Varga 2000; Sonn and Storper 2008), nor do they surpass, in the case of Europe, the distance that can be reasonably covered by a person by car or public transport in a day – circa 200 km (Moreno et al. 2005; Crescenzi et al. 2007; Rodríguez-Pose and Crescenzi 2008). Hence, innovation benefits from the proximity of the different actors involved in the generation, diffusion and absorption of knowledge and contributes, in turn, to the emergence of clusters. Economic actors clustered in close geographical proximity tend to innovate more and to benefit more from knowledge spillovers than those working in remote locations. Clusterisation also enables firms to exchange knowledge and information fast and increases the chance for an innovative firm to find partners and early-adopters of a new technology (Moore and McKenna 1999). From this perspective, the ‘clusterisation’ of firms working in the same sector or even “competing in the same industry or collaborating across related industries tends to trigger processes that create not only general dynamism and flexibility but also learning and innovation” (Doloreux and Parto 2005, 137). ‘Clusterisation’ is more effective, however, when it involves other actors in the innovation process beyond firms. That is, when universities, R&D research centres, and other public and private institutions create ‘dense’ environments of socioeconomic actors, weaving complex networks of interaction that become the channels through which knowledge is disseminated and transformed into economically viable activity. Once again, clusters work best for innovation and economic growth when they are not just mere collocations of firms in similar or related sectors, but when they become regional systems of innovation.

Some research strands have also stressed that the best way to generate and absorb innovation is through a mixture of local ‘buzz’ and ‘global pipelines’ (Bathelt et al. 2004; Wolfe and Gertler 2004). While local ‘buzz’ represents the quintessential elements of physical proximity, encompassing face-to-face contacts and other forms of human interaction in dense environments (Storper and Venables 2004), ‘global pipelines’ channel knowledge through cognitive, social, and institutional mechanisms, overcoming physical distance (Bathelt et al. 2004).

Physical proximity alone, however, does not suffice to generate innovation and growth. Other characteristics are at play in order to transform regions into truly functioning innovation systems. It is commonly accepted that regions with a similar institutional framework and organisation “may show different abilities to accommodate innovation” (Iammarino 2005, 503). Factors such as ‘social capability’ and ‘technological congruence’ (Abramovitz 1986; Fagerberg 1987 and 1994) contribute to determine to what extent any given region or territory is ‘innovation prone’ or ‘innovation averse’ (Rodríguez-Pose 1999). ‘Social capability’ refers to the capacity of a region to shape its institutional framework in order to support the emergence of what is known as the ‘socio-institutional environment’ or the ‘innovation-supportive culture’ (Doloreux and Parto 2005, 135) required for the generation of innovation. Local socio-institutional environments that favour entrepreneurship are, for example, more likely to generate systems which will be innovation enhancing than those environments that do not. ‘Technological congruence’ refers to the idea of technological frontier (Abramovitz 1990), i.e. the proximity of a region to develop cutting-edge knowledge and thus to make the most from new investment in the promotion of innovation. A region’s technological congruence depends, in turn, on characteristics, such as the presence of a specialized labour market or of developed ‘local learning processes’ integrating company networks (Doloreux and Parto 2005, 135).

The presence of a good ‘social capability’ and strong ‘technological congruence’ contributes to bridge the gap between the supply side of innovation – mainly the institutional sources of knowledge creation – and the demand side, featured by the productive systems that develop and apply such knowledge (Braczyk et al. 1998). Innovation thus becomes a territorially-embedded process (Rodríguez-Pose and Crescenzi 2008, 54). In one way or another, this notion of ‘territorial embeddedness’ has been present in all approaches highlighting the importance of clusters for innovation and growth, articulating concepts such as ‘innovative milieux’ (Camagni 1995), ‘learning regions’ (Morgan 1997), ‘industrial districts’ (Becattini 1987), and, not least, that of ‘regional innovation systems’. True territorial embeddedness is, however, considered to be “feasible only at regional level” (Cooke 2006, 6). Indeed, the regional dimension allows the different actors involved in the process of knowledge-sharing and exchange to get to know each other, to work together, and to trust one another. All these aspects make the region “the best geographical scale for an innovation-based learning economy” (Doloreux and Parto 2005, 136).

There has certainly been no shortage of high quality research dealing with the implications of clusters for innovation and economic growth (Cheshire and Malecki 2004). Among this research, qualitative case-study analyses abound. Most of these studies have focused on a handful of cases, including a limited number of well-known technology clusters, such as Cambridge (e.g. Keeble et al. 1999), or of industrial clusters in the Third Italy or Baden-Württemberg. Other research has stepped away from these traditional cases and wandered into apparently less fertile ground. Cumbers. Mackinnon, and Chapman’s (2003) analysis of SMEs in the Aberdeen oil complex represents one such example. However, while many of these analyses provide deep insights into the internal and external relationships that may – or may not – make clusters hotbeds of innovation and growth, there is always the uncertainty of whether we have been simply observing the lushest trees, while, at the same time, overlooking the overall condition of the forest. More systematic analyses, trying to map out clusters across Europe have been few and far between. Crouch et al. (2001), in perhaps the most ambitious attempt to date, have mapped local production systems across France, Germany, Italy, and the United Kingdom, however the analysis has been confined to national borders, generally avoiding dynamic quantitative analysis. This noticeable absence of robust quantitative evidence is without doubt the result of problems with measuring the intricate interactions, the institutional linkages and the complexity of the collective learning processes happening within clusters, learning regions, or regional innovation systems. But this absence of more systematic analysis flies in the face of recent improvements in databases measuring clusters and of the importance clusters have acquired in policy circles. The belief that clusters, in general, and regional systems of innovation, in particular, are key drivers of innovation and growth has become widespread among academics and policy-makers alike. The diffusion of the cluster concept by leading management academics such as Michael Porter and the impetus that research on regional innovation systems has acquired in recent years have lead to the extensive implementation of cluster policies as a means to achieve economic dynamism.

Yet the perception of clusters as the fundamental drivers of innovation and growth is challenged by more traditional theoretical strands dealing with the genesis and diffusion of innovation. One of these strands is the linear model of innovation, which is based on the basic premise that innovation and growth are driven by greater investment in research and development (R&D) (MacLaurin 1953). The greater the investment in R&D, the greater the output, and the greater the economic growth. Linear models of innovation and growth have thus fundamentally focused on the role of two parameters: the level of expenditure in R&D of a country (or a region), as the key input, and the number of patent applications, as the main output. In particular, "R&D investment becomes [even] more essential when industries move closer to their technological frontier" (Aghion 2006, 2). Other factors, such as the protection of intellectual property rights, also matter for innovation. However, beyond these basic factors, most other parameters are considered either not to count for the genesis of innovation, or to play a mere supporting role.

The linear model neglects, however, another key aspect of the innovation process: 'the context' in which it occurs or, as mentioned earlier, its territorial-embeddedness (Rodríguez-Pose and Crescenzi 2008, 54). From a linear model perspective, innovation is seen as a static process, not influenced by the dynamics and the quality of the different interactions between the actors at play. Yet, from a different perspective, the context in which the interaction among economic actors takes place is fundamental in determining whether innovation will occur or not, or whether it will be assimilated by economic actors or not. This is what Rodríguez-Pose (1999) has called the 'social filters', or the unique combination of "innovative and conservative (...) elements that favour or deter the development of successful regional innovation systems" (Rodríguez-Pose 1999, 82) in any given territory. These elements are neither the networks, nor the institutions which permit the formation of regional innovation systems, but the substrata which encourage the creation and success of these local networks and institutions. They include, among others, the level of education and skills in the population, the level of use of human resources, the demographic dynamism, risk-taking, and the sectoral specialization. The unique combination of these factors in any particular space makes any territory either 'innovation prone' or 'innovation averse' (Rodríguez-Pose 1999).

However, despite these contrasting, and not always complementary, approaches, relatively little effort has been made in order to discriminate between them and to identify which approach has a greater sway over the generation and diffusion of innovation and economic growth. Do clusters have a greater influence over innovation than investment in R&D? Is the role of education greater than that of R&D and the presence of regional systems of innovation in generating economic growth? The interaction among these factors has also been underexplored. Does the presence of a favourable social filter reinforce the potentially positive effects of the presence of clusters on innovation and growth? And how does it interact with R&D? These are questions which have been overlooked or, at most, addressed tangentially by the literature studying innovation and economic growth and which have been mainly examined in case studies. This chapter aims to cover this gap in the literature by looking at the interaction between R&D, social conditions, and the presence of clusters and regional innovation systems across the regions of the enlarged EU for the period between 1995 and 2006, from both a static and dynamic perspective.

3 From Theory to Practice

3.1 Operationalising the Model

That the questions presented in the previous section have been somewhat neglected can be largely put down to the difficulties in defining – and, consequently, operationalising – most of the concepts involved in this type of analysis. In particular, the concept of what a regional system of innovation is far from straightforward. The most commonly accepted definition is that by Cooke, Gómez Uranga and Etxeberría (1998, 1581), who consider that when "firms and other organisations are systematically engaged in interactive learning through an institutional milieu characterised by embeddedness" they make up a regional innovation system. The interaction between the production and the institutional structure generates territorially-embedded networks which determine the genesis, import capacity, diffusion, and assimilation of knowledge within any given cluster (Howells 1999; Evangelista et al. 2002). These networks generate, in turn, a governance and a business structure within the cluster (Braczyk et al. 1998). The governance dimension involves the "soft infrastructure of enterprise innovation support" (Cooke 2006, 6), such as "public policy, institutions, and knowledge infrastructure" (*ibid*). The business dimension includes the "industrial base: [..] the type of firms, the level of R&D investment, the level of linkages" (*ibid*, p. 7).

Most of these networks, institutions and dimensions are idiosyncratic and dependent on the context on which every cluster is placed. As the characteristics of each region and locality are unique, operationalising clusters in a quantitative manner is virtually impossible (Iammarino 2005). It is often the case that regions with, on paper, very similar socio-institutional structures diverge (often wildly) in terms of their innovative capacity. These differences underline that "there is no single model that is able to generalize the dynamics of successful regional innovation systems" (Doloreux and Parto 2005, 138) and question whether the regional innovation framework can be really applied beyond the identification of 'stylized regional innovation systems' (Iammarino 2005), that is, purely theoretical concepts with no clear equivalent on the ground.

Despite these gargantuan difficulties, some authors have embarked on the heroic task of trying to identify clusters and/or design cluster policies in Europe on a large scale. This is, for example, the case of the pioneering work of Jacobsson et al. (2006), who, using functional analysis, aim to identify and measure the different functions of a cluster and the different steps in its creation. In a more systematic way, the European Commission has used the INNOVA initiative to gather best practices from European clusters and to promote them (EC 2006). But it is possibly the European Cluster Observatory (ECO) the organisation, which has made the greatest effort in order to systematically identify, measure, and map clusters in Europe. Their measures – not exempt, as any such measure, of controversy – are used in this chapter in order to assess clusterisation across the regions of Europe.

Operationalising other constituents of innovation and growth, such as R&D and, in particular, 'social filters' is also problematic. But the indicators behind the construction of this type of variables tend to generate, by and large, greater consensus.

3.2 *Identifying the Variables*

Bearing in mind the caveats presented above, in this section we now define the variables included in the analysis. In order to do this we follow previous empirical work and, in particular, the work of Rodríguez-Pose and Crescenzi (2008), who resorted to a series of parameters to measure 'social filters' across the regions of Europe.

The dependent variable is perhaps the most straightforward and widely accepted of all the variables included in the analysis: the growth of the logarithm of the regional GDP per capita. The explanatory variables deserve, by contrast, much greater attention.

Following the three key strands presented in the theoretical section (linear model, 'context', and clusters and regional systems of innovation), in order to analyse the link between (regional) economic growth and the factors that generate innovation in Europe's regions, we resort to three basic explanatory variables, which we call the three filters. These are the 'R&D' filter, the 'social' filter and the 'clusterisation' filter.

R&D Filter – The 'R&D filter' is directly derived from the basic principle of the linear model of innovation. We create a composite index using the two basic input and output variables of this approach. The former is represented by the regional expenditure in R&D as a percentage of GDP, whereas the latter is depicted by the number of patent applications per million inhabitants in any given region. Despite the controversy surrounding patent applications as a measure of innovation outputs – not all sectors patent in the same way, not all patents lead to true innovation, and not all patents lead to short term economic returns – the inclusion in the analysis of the number of patent applications responds to its value as a proxy for the capacity of a region to absorb and generate knowledge and its correlation with regional economic growth. *R&D expenditure* and *patent application* are given equal weight in the resulting 'R&D filter' index. As could be expected, the higher the R&D expenditure and the higher the patent applications per capita, the higher the value of the R&D filter.

Social Filter – The concept of 'social filter' aims at building a composite index reflecting the socio-economic conditions that make a region innovation prone or innovation averse. This filter reflects the 'territorially-embedded' character of innovation as often presented in regional innovation systems approaches. Multiple aspects can play a role in the emergence of innovation. Among these we highlight: (a) local market rigidities, (b) demographic aspects, (c) education, skills, and human capital, and (d) the scientific base of the region.[1]

The 'social filter' variable used in this chapter is based on that of Rodríguez-Pose and Crescenzi (2008), including some additional variables, in order to reproduce better the socioeconomic setting in which innovation and growth take place. The first aspect covered – that of market rigidities – refers to the local use of resources. The variables covered by this domain include long term unemployment (*long term unemployment*) as a means of measuring the degree of rigidity in the local labour market and, at the same time, as a potential indication of the share of the active population with inadequate or insufficient skills. The second variable is agricultural employment (*agricultural employment*), used as a proxy to partially measure levels of 'hidden unemployment', especially prevalent in some of the new members of the EU. These two parameters are also indirectly linked with the productivity level of the labour force. The last variable in this domain is the level of corporate tax rate (*corporate tax rate*). The rationale for the inclusion of this variable is based on the complaints often raised by entrepreneurs and other economic actors. A high level of corporate taxation is said to diminish the investment capacity of firms (especially in R&D) and to be a disincentive for location in certain regions.

The second aspect covered by the 'social filter' relates to the demographic characteristics of a region. It is assumed that the total population of a region (*total population*) may have an impact on its innovative capacity and thus on its growth potential. Indeed, in regions with large populations, the presence of a large market pool will make it easier for a company to find workers with the right skills and knowledge. Moreover, a larger population may be at the source of both greater diversification (Jacobs type) and specialization (Marshall-Arrows-Romer type) externalities. The influence of the number of people living in a region on innovation and growth is complemented by the average age of the population (*percentage of young*). The impact of this variable on economic growth is difficult to predict theoretically. On the one hand, a young population is often associated with less risk aversion and greater openness to innovation. On the other, if a large percentage of the young is still studying or in full-time training, their immediate impact on economic growth is bound to be limited.

[1] Another aspect is local institutions, which are, however, hard to measure at the regional level for the whole of Europe.

The third domain refers to the education and skills level of the population. Education is widely regarded as a key source for innovation and economic growth. Two variables are included in this domain: the share of the population with a higher education degree (*education population*) and the percentage of adults participating in lifelong learning activities (*lifelong learning*).

The final domain in our 'social filter' index reflects the importance of the presence of scientists in the innovation process. The variable included is the share of employed in science and technology (*hr in science & techno*), as a proxy for the human resources devoted directly to the generation of new knowledge. A strong scientific community in a region can be considered as a competitive advantage for innovation and growth. This aspect was not included in Rodríguez-Pose and Crescenzi's (2008) operationalisation of 'social filter'.

Principal Components Analysis (PCA) is used in order to create the resulting composite variable 'social filter'. The advantage of resorting to PCA is that it can be used as a means for identifying patterns in data and of merging "a set of variables [...] into an individual indicator able to preserve as much as possible of the variability of the initial information" (Rodríguez-Pose and Crescenzi 2008, 57). The results of amalgamating all the variables included in the 'social filter' into one composite variable by means of PCA are presented in Table A.5 in Annex 2. The first principal component – used as our 'social filter' variable – accounts for 38 % of the total variance. The contribution of individual variables to the composite variable 'social filter' has the expected sign: high long-term unemployment, agricultural employment, corporate tax rates or young populations lower the social filter index; big populations, high educational achievement and life-long learning levels in the population and a good endowment of researchers in science and technology increase the social filter index.

Clusterisation Index – The third and final filter represents an approximation – given the complexity of the task – at capturing the 'clusterisation' effects which are, according to the literature on clusters and regional innovation systems, believed to be directly behind the economic dynamism of a region. The logic for including this index is based on the importance of proximity in the generation of innovation as explained in the theoretical section of the chapter. The variables included try to measure the propensity of firms to cluster – or concentrate geographically – in similar or related industries. As mentioned earlier, such dynamics are expected to create important internal flows of knowledge and a strong potential to innovate.

The three variables used in this 'clusterisation index' stem from data collected by the European Cluster Observatory (ECO). The ECO has identified clusters in the 27 members states of the EU, sorting them by region and assessing them through a detailed methodology.[2] Using different criteria, the ECO develops a series of

[2] The detailed methodology is available directly on the website: www.clusterobservatory.eu. Only clusters with at least 1,000 workers are taken into consideration in order to "prevent the appearance of very small insignificant clusters" (Cluster Observatory website in Methodology: Evaluation of regional cluster strength).

regional indices. The index of cluster specialisation (*specialisation*) exploits employment data in order to create a specialisation quotient representing the employment intensity of a given regional cluster sector compared to the employment intensity in general for this region (please see Annex 1 for the exact formula). The *Focus* index captures the share of a region's total employment represented by a specific cluster. If this share is large, this means that the 'clusterisation' effects for that sector in this region are strong (see Annex 1 for the exact formula). The last variable included in this index intends to control for the diversification of clusters in a region (*diversification*), i.e. the presence of economic cluster activities in different industries. If a region is characterized by the existence of several clusters in various sectors (even if these clusters are relatively small in comparison to those in regions with only one large cluster), it can expect to benefit from diversification or Jacobs-type externalities, likely to foster greater innovation and growth.

As in the previous filter, the three variables are combined into a composite one using PCA (Table A.7 in Annex 2). The first principal component, used as the 'clusterisation filter variable', accounts for 49 % of the total variance. Greater specialisation, focus, and diversification of clusters in a region result in a higher clusterisation index.

3.3 Data and Geographical Coverage

The analysis covers 152 regions in 15 EU Member States for the period 1995–2006.[3] The economic analysis is conducted at NUTS 2[4] regional level for most of the countries – Austria, Czech Republic, Finland, France, Italy, Hungary, Poland, Portugal, Slovakia and Spain. NUTS 1 regions have been used for Belgium, Germany, Greece, the Netherlands and the United Kingdom, both for reasons of data constraints and as a need to reflect – at least in the case of decentralised countries – similar tiers of government and levels of decision making capacity.[5]

The data used in the chapter stem from two main sources: the European Statistical Office (Eurostat) and the European Cluster Observatory. Eurostat data

[3] Bulgaria, Cyprus, Denmark, Estonia, Ireland, Latvia, Lithuania, Luxembourg, Malta, Romania, Slovenia and Sweden were excluded because of lack of sufficient and/or reliable regional data on R&D expenditure.

[4] Nomenclature of Territorial Unit for Statistics as defined by the European Commission on http://ec.europa.eu/comm/eurostat/ramon/nuts/home_regions_en.html

[5] In addition, some specific regions have been excluded because of lack of data. This is the case of all the French Overseas Departments and Territories, and of the regions of the Åland islands (Finland), Açores and Madeira (Portugal) and the African enclaves of Ceuta and Melilla (Spain)

Table 1 Definition of the variables and data sources

Variable	Definition	Sources
Dependent variable		
Growth of GDPpc	GDP PPS per inhabitant	Eurostat
R&D filter		
R&D expenditure	Percentage of GDP	Eurostat
Patents	Applications per million inhabitants	Eurostat
Social filter		
Long term unemployment	Percentage of total unemployment	Eurostat
Agriculture employment	Percentage of total employment	Eurostat
Corporate tax rate	Percentage of corporate benefits (national proxy)	Eurostat
Percentage of young	People aged 15–24 as percentage of total population	Author's calculations based on Eurostat data
Total population	Percentage of national population	Eurostat
Education	Percentage total population with tertiary education (levels 5–6 ISCED 1997)	Eurostat
Life long learning	Percentage of Adults (25–64) participating in education and training	Eurostat
Human resources in science and technology	Percentage of active population	Eurostat
ClusterIndex		
Specialisation	cf Annex I	European cluster observatory
Focus	cf Annex I	European cluster observatory
Diversification	Number of clustered industries in the region per 100,000 employees	Author's calculation based on European cluster observatory data

are mainly used for the variables included in the 'R&D' and 'Social filter' and for the dependent variable. European Cluster Observatory data are used in the construction of the 'Clusterisation filter'. Missing data were estimated using trends. All data are gathered at the regional level, with the exception of the corporate tax rate, which is national. It is also worth noting that the European Cluster Observatory bases its data on what it calls a 'reference year' (corresponding to the year of the most recent available data). This 'reference year' differs for each country. This implies taking the assumption that the Clusterisation Index of a region is homogenous over the period of analysis.

The names, definitions, and sources of the 14 variables included in the analysis are presented in Table 1.

4 The Model and Empirical Analysis

4.1 The Model

The econometric model used in the empirical analysis adopts the following form:

$$\Delta \ln GDPpc_{i,(t+1)-t} = \alpha + \beta_1 \ln GDPpc_{i,t-1} + \beta_2\, RDFilter_{i,t} + \beta_3\, SocFilter_{i,t} + \beta_4\, ClusterIndex_{i,t} + \beta_5 ND + \varepsilon. \quad (1)$$

Where:

$\Delta \ln GDPpc_{i,(t+1)-t}$	Is the growth of GDP per capita in region *i* during the period of analysis
α	Is a constant
$\ln GDPpc_{i,t-1}$	Represents the natural logarithm of GDP per capita in region *i* at the beginning of the period of analysis
$RDFilter_{i,t}$	Denotes the R&D filter conditions in region *i* and time *t*
$SocFilter_{i,t}$	Represents the social filter conditions in region *i* and time *t*
$ClusterIndex_{i,t}$	Denotes the degree of clusterisation in region *i* and time *t*
ND	Are a series of national dummies
ε	Is the error term

The specific characteristics of the data included in the 'Clusterisation index' – the use of a 'reference year' by the European Cluster Observatory – constrain us to estimating the model by means of heteroskedasticity-consistent pooled OLS (Ordinary Least Square) regressions. This method has the advantage of allowing us to present both a static and – by resorting to annual lags – a dynamic image of the association between the different indices included as independent variables and regional economic growth. All the estimates carried out are based on a robust variance matrix estimator which is valid in the presence of heteroskedasticity or serial correlation (Wooldridge 2006). VIF tests have been conducted for all the variables in the model, with no multicollinearity having been detected. In order to account for unobserved national fixed effects, a set of national dummies (variable *ND*) is included in the model.

The main aim of the analysis is to examine the association between the composite variables representing the competing explanations of the factors behind innovation presented in the theoretical section – clusters and regional innovation systems, the local socioeconomic conditions or social filter, and traditional investment in R&D and patent applications – and economic growth. The analysis is conducted in three steps. First, a static picture is presented in Table 2. In this table the results of running 14 different regressions are reported. These regressions aim to capture both the aggregate connection between each composite variable representing the different approaches to the analysis of innovation or filters (regression 1), as well as the individual correlation of between each individual variable included in each of the three filter variables (regressions 2–9 for the *Social Filter,*

Table 2 Pooled (HC-OLS) regressions of regional log GDP/capita

	1	2	3	4(*)	5	6	7	8	9	10	11	12	13	14
Constant	2.157***	0.955***	1.288***	0.691***	2.048***	0.561***	1.459***	1.086	2.165***	2.327***	2.032***	2.012***	2.008***	2.067***
	0.350	0.256	0.2674	0.132	0.261	0.195	0.292	0.154	0.281	0.346	0.367	0.329	0.340	0.356
Log GDPpc	0.757***	0.947***	0.901***	0.988***	0.896***	0.979***	0.854***	0.9101***	0.722***	0.738***	0.770***	0.770***	0.769***	0.766***
	0.038	0.027	0.0297	0.014	0.029	0.021	0.035	0.006	0.030	0.0373	0.040	0.036	0.037	0.039
R&D filter	0.009	0.034***	0.036***	0.006	0.027***	0.038***	0.019**	0.024***	0.003			0.004	0.008	0.008
	0.008	0.006	0.006	0.006	0.007	0.006	0.008	0.006	0.008			0.009	0.008	0.008
Social filter	0.049***									0.0494***	0.051***	0.049***	0.048***	0.049***
	0.005									0.0046	0.005	0.005	0.005	0.005
Clusterisation index	0.013**	−0.004	0.002	−0.003	−0.007	−0.003	0.010*	0.000	0.016***	0.014**	0.011**			
	0.005	0.004	0.003	0.003	0.004	0.003	0.006	0.003	0.004	0.005	0.005			
National dummies	X	X	X		X	X	X	X	X	X	X	X	X	X
Social filter														
Long term unemployment		−0.002***												
		0.001												
Agriculture employment			−0.007***											
			0.001											
Corporate tax rate				−0.010***										
				0.001										
Percentage of young					−0.045***									
					0.005									
Total population						−0.001								
						0.001								
Education							0.025***							
							0.004							
Lifelong learning								0.047***						
								0.005						
HR in science and techno									0.019***					
									0.001					
R&D filter														
Patent application										0.000**				
										0.000				

R&D expenditure											−0.009			
											0.007			
Clusterisation index														
Specialisation												0.0126*		
												0.0065		
Focus													0.011*	
													0.006	
Diversification														0.011
														0.007
R^2	0.925	0.883	0.885	0.898	0.895	0.880	0.914	0.905	0.916	0.926	0.925	0.924	0.924	0.924
F	614.21	1012.33	546.29	1737.74	493.00	1313.86	533.36	1104.96	418.44	538.64	773.24	561.30	705.68	517.31
Number of observations	1,756	1,760	1,760	1,760	1,760	1,760	1,760	1,760	1,756	1,756	1,756	1,756	1,756	1,756

*, **, *** indicates significances at 10 %, 5 % and 1 % respectively (*) this regression has been run without national dummies since a national proxy has been used for the Corporate tax rate

Table 3 Dynamic analysis

	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
Constant	2.157***	1.956***	1.689***	1.683***	1.283***	1.159***	1.392***
	0.350	0.314	0.282	0.243	0.253	0.256	0.248
Log GDPpc	0.757***	0.785***	0.820***	0.853***	0.874***	0.891***	0.899***
	0.038	0.034	0.031	0.028	0.027	0.027	0.028
R&D filter	0.009	0.015**	0.017***	0.017***	0.016***	0.014**	0.015**
	0.008	0.007	0.006	0.006	0.006	0.006	0.007
Social filter	0.049***	0.043***	0.037***	0.031***	0.030***	0.025***	0.024***
	0.005	0.005	0.004	0.004	0.004	0.004	0.005
Clusterisation index	0.013**	0.011**	0.009*	0.006	0.004	0.002	0.002
	0.005	0.005	0.005	0.005	0.005	0.005	0.005
R^2	0.925	0.932	0.940	0.947	0.956	0.964	0.968
F	614.21	582.94	630.35	705.77	763.65	927.23	1281.37
Number observations	1,756	1,596	1,436	1,276	1,116	956	796

*, **, ***indicates significances at 10 %, 5 % and 1 % respectively

regressions 10 and 11 for the *R&D Filter* and regressions 12–14 for the *Clusterisation Index*), on the one hand, and economic growth in the regions of Europe, on the other. The three composite filter variables – R&D filter, social filter and clusterisation index – have been standardised in order to make it possible to compare their effects on regional growth. The standard errors are presented (in italic) under the value of the coefficient. In each regression tests have been conducted in order to account for the good specification and goodness of fit of the model.

Second, the dynamic dimension of the relationship is reported in Table 3. This table includes seven pooled HC-OLS regressions, with the dynamic effect achieved by regressing regional per capita growth on the initial GDP per capita and lagged filters, where the number of lags is $n \in [1; 6]$. While this econometric approach enables us to give a global picture of the dynamics of the model, it has the drawback of reducing the number of observations after each lag. In any case, even after six annual lags, the number of observations (n = 796) remains relatively large. As in the case of the static analysis, the potential presence of spatial serial correlation is controlled for. No multicollinearity is detected in the model.

Third, and in order to better assess how the association between the different filters may affect innovation and growth, the dynamic analysis is rerun substituting each of the dependent variables of interest in model (1) – R&D filter, Social filter, and Clusterisation index – by their pairwise interaction.

4.2 Static Analysis

The first fact that can be underlined in the static analysis is the goodness of fit of the model (Table 2). A very high proportion of the variance in regional growth is explained, implying that the combination of the more traditional variables of innovation, with the social filter and its components, and the different indicators aimed at identifying the presence of clusters have a powerful association with regional economic growth.

Most variables are significant and tend to remain so despite the introduction of different controls. This is the case of the initial GDP per capita of a region, which is positively and robustly associated with regional economic growth in all 14 regressions (Table 2). When the three composite filter variables are considered together, the social filter and the clusterisation index have a positive and significant relationship to economic growth, but the R&D filter variable is not significant (Table 2, Regression 1). This, in principle, represents a confirmation of the views of those strands of research which have highlighted importance of both the presence of clusters and complex regional innovation systems, on the one hand, and the basic socioeconomic conditions on which these networks and systems can be constructed, on the other, for economic growth. Indeed successive regressions (Regressions 2 to 9) reveal the close interaction between the presence of clusters and of favourable socioeconomic conditions. When the composite social filter variable is excluded from the analysis, the coefficient of the clusterisation index becomes generally insignificant (Regressions 2, 3, 4, 5, 6 and 8, Table 2). This also points to the fact that the association of clusters with regional economic growth is closely related to the presence of a good level of education in the population (Regression 7), with an emphasis on life-long learning (Regression 8) and, essentially, with the existence of a good pool of researchers (Regression 9). The presence of a relatively ‘high-tech’ labour force thus seems to play a major role in the settlement of innovation-enhancing socio-economic conditions and, more globally, in the economic growth of a region. This may be a confirmation of some of the basic characteristics associated with regional innovation systems. In these complex systems the existence of a pool of researchers surrounded by a highly educated workforce will naturally tend to form a community where innovation is generated, diffused, and absorbed in the workplace. This is, in essence, the ‘local learning process’, as defined by Doloreux and Parto (2005). If companies in a region are, in addition, geographically clustered, this is likely to increase intra-regional knowledge flows between high-tech workers and educated people. Therefore, clusterisation, on the one hand, and the presence of a high density of researchers and of a well educated labour force, on the other, will reinforce each other in the generation of innovation and growth. The greater the density of clusters in any given region, the easier the knowledge flow between innovative firms and the rest of the production fabric, facilitating the diffusion and absorption of knowledge. This renders the impact of clusters significant to economic growth. The absence of these conditions, in contrast, makes clusters almost irrelevant for growth.

Factors such as the presence or absence of long-term unemployed, of greater or lower levels of agricultural employment, of a younger or older population, or the overall dimension of the region neither enhance, nor reduce the potential relationship between clusters and economic growth. In fact, they contribute to make them irrelevant (Table 2).

Extracting the social filter from the analysis renders the more traditional R&D variables of R&D expenditure and patent applications positive and significant (Regressions 2 through 9, Table 2), with the exception of when the R&D filter is considered in combination with corporate tax rate (Regression 4) and the regional human resources devoted to science and technology (Regression 9).

Turning to the individual variables included in each filter – while controlling for other filters – exposes other interesting associations. First, the decomposition of the R&D filter variable into its two components brings to the fore a significant and positive correlation between the number of patent applications and growth (regression 10) whereas, investment in R&D turns out as non significant (regression 11). The weak association between R&D expenditure and economic growth, at least in the short term, comes in support of the views which highlight the relative irrelevance of policies dominated by public investment in R&D in environments associated with inadequate or weak socioeconomic conditions and in the presence anaemic networks and systems to absorb it (Cooke 2001). However, the results regarding patent applications are in line with the linear approach to innovation.

The different regressions including individual social filter variables (Regressions 2–9) give a more detailed information about the socioeconomic conditions which may matter for innovation and growth in the regions of Europe. Among the socioeconomic variables that have a positive and significant association with economic growth, the educational parameters clearly stand out. Both coefficients of the level of education of the population and of a life-long learning dimension in the workforce are strongly positive and significant and of great importance for growth (Regressions 7 and 8). The human resources devoted to science and technology go in the same direction (Regression 9). By contrast, the level of long term unemployment, that of agricultural employment, the corporate tax rate, and the percentage of young (Regressions 2 to 5) are negatively and significantly associated with regional economic growth. The demographic size of a region is completely dissociated from growth, once the R&D and clusterisation indices are included in the analysis (Regression 6, Table 2).

Finally, of the variables making up the clusterisation index, specialisation and focus are positively and significantly – albeit at the 10 % level – correlated with regional economic growth (Regressions 12 and 13, Table 2), The coefficient of the variable representing the diversification of clusters is, however, not significant (Regression 14).

In brief, the static analysis exposes the very strong, positive, and robust association between the social filter of a region and its economic growth. The strength of this relationship is significantly stronger than that of the other two filters with regional growth. The link between R&D and patents and growth, on the one hand, and the presence of clusters, combining both specialisation and diversity externalities, and growth, on the other, is contingent on their interplay with the presence or absence of adequate social filters. The R&D variable only becomes significant when the social filter is not taken into account, while the relevance of the existence of clusters in a region for economic growth only comes to the fore in areas with adequate social filters (Table 2). The capacity by economic actors to absorb innovation across European regions depends on the overall combination of social conditions and, more specifically, on the educational endowment of the population and on the existence of a 'high-tech literate' labour force. Clusters also matter, but their importance for growth is contingent on the existence of adequate social filters. Weak or rigid social filters – characterised by factors such as the prevalence of long term unemployment, low productivity employment and high levels of corporate taxation – may damage significantly the innovation potential of a region and render the association between clusters and economic growth irrelevant. Adequate social filters (i.e. those featured by well-educated populations, a high-tech labour force and limited market rigidities) combined with the capacity to transform R&D into patents quickly, and to develop clusters both specialised and focused – relative to those in other regions – are at the base of the formation of innovative and economically dynamic regions.

4.3 Dynamic Analysis

The dynamic analysis in an up to 7 year horizon is presented in Table 3. It adds a series of interesting nuances to the relationship between the key factors behind innovation and economic growth, outlined in the static approach. The most relevant finding is the enduring importance of an adequate social filter for regional economic growth in Europe. The social filter is the only composite variable to remain significant throughout the whole period of analysis, despite the fact that the strength of its relationship with regional economic growth wanes in time. The association of the social filter with the variation of regional economic growth in Europe is only half as strong when considering a 6 year time lag as when no time lags are considered (Table 3).

Another important finding is the contrasting trajectories of the relationship between the R&D filter, on the one hand, and the clusterisation index, on the other, and regional economic growth. As highlighted in the static analysis, the presence of a greater specialisation and focus in clusters in favourable socioeconomic environments is connected to higher growth in the short term. This positive relationship is, however, short-lived. The strength and the significance of the coefficient starts to wane quickly and becomes non significant beyond 3 years (Table 3). The R&D filter, by contrast, is insignificant in the first year considered, but becomes significant after 1 year. The strength of this association remains more or less intact during the remaining years. The importance of this association also increases over time, especially as the intensity of the connection between the social filter and regional economic growth starts to decline (Table 3). This may be a signal that, at least in the European case, the importance of clusters and innovation systems for regional economic growth may have been somewhat overstated. Conversely, hard R&D indicators may have a greater sway over short and medium-term economic performance than admitted by some recent strands of literature.

4.4 Dynamic Analysis with Interaction Terms

But how does the interplay between the three different filters affect economic growth? In order to get a more accurate picture of how the interaction between the factors behind innovation promote regional growth in Europe, the dynamic analysis is rerun substituting the independent variables representing each of the filters by their interactions – interaction between the R&D filter and the Social filter, between the R&D filter and the Clusterisation index, and between the Social filter and the Clusterisation index. All regressions are run including the GDP per capita of the region and national dummies.

The results of this analysis, presented in Table 4, underline once again the importance of social conditions for the genesis of innovation and growth. The interaction between the Social filter and the R&D filter yields a positive and significant coefficient, which remains so over the period of analysis. Adequate social conditions – and, in particular, a good human capital endowment (Rodríguez-Pose and Crescenzi 2008) – facilitate the transformation of R&D investment and patent applications into economic growth (Table 4). The interaction between the presence of clusters and a good R&D environment is, by contrast, not associated with higher levels of growth. Regions which benefit from high levels of investment in R&D and from a relative good endowment of clusters do not necessarily grow faster that regions lacking these characteristics, in the absence of adequate social filters which would help transform these factors into greater economic dynamism. Similarly, the interaction between the social filter and the presence of clusters is completely dissociated from the economic performance of the region (Table 4).

5 Conclusion

The objective of the chapter has been to assess through the use of an econometric model with a static and a dynamic dimension the association between the different factors that promote innovation and economic growth across the regions of Europe. In particular, we have analysed the role that the presence of clusters within regions play in this relationship. The intention was to overcome the tendency by most of the literature on clusters to concentrate on the most favourable cases (Martin and Sunley 2003), which was ultimately raising important questions about the role of clusters in the generation of innovation and economic growth. Are all clusters a source of innovation and growth? Or is it just those that happen to be located in the right environments, in the right sectors, and/or in places where adequate management is available and adequate support policies have been implemented? The

Table 4 Dynamic analysis with interaction terms between the different filters

		Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
Interaction between R&D filter and social filter	Coefficient	0.0023***	0.0028***	0.0025***	0.0025***	0.0021***	0.0018**	0.0016*
	Standard error	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	R^2	0.879	0.898	0.916	0.931	0.943	0.9542	0.9590
	F	2244.47	1778.56	1405.31	1208.18	1049.95	860.57	732.72
Interaction between R&D filter and clusterisation index	Coefficient	−0.0007	−0.0013	−0.0007	−0.0001	0.0010	0.0030	0.0042
	Standard error	0.004	0.004	0.004	0.004	0.005	0.005	0.005
	R^2	0.8775	0.8965	0.9144	0.9291	0.9420	0.9537	0.9588
	F	2214.27	1827.81	1362.87	1143.91	947.89	751.51	638.85
Interaction between social filter and clusterisation index	Coefficient	−0.0008	−0.0010	−0.0010	−0.0011	−0.0011	−0.0012	−0.0012
	Standard error	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	R^2	0.8787	0.8976	0.9153	0.9299	0.9428	0.9540	0.9589
	F	2428.88	1933.90	1555.65	1384.80	1141.24	957.06	818.87
Number observations		1,756	1,596	1,436	1,276	1,116	956	796

chapter has thus examined the role of clusters across regions in Europe, looking not just at the brightest trees in the forest – the Cambridges, Venetos, Jutlands or Württembergs of the cluster world – but also at the average and even the moribund trees – i.e. the clusters which happen to be located perhaps in the wrong environments, the wrong sectors, and with inadequate management and policies. The size of employment in clusters relative to overall employment, the dominance of specific clusters, and cluster diversification were the three criteria used in order to measure the presence of clusters across regions in Europe. Two other composite indices or filters, covering 'hard' innovation indicators – the R&D filter – and the socioeconomic conditions on which innovation takes place – the Social filter – were included in the analysis in order to represent the other factors which can promote regional innovation and growth.

Three primary conclusions can be extracted from the analysis. First and foremost is the importance of having a favourable socioeconomic setting in order to foster innovation and growth. Much more than the presence or absence of clusters, having a good level of education, a strong endowment of skills in the population or a workforce with sufficient high tech skills is not just crucial in order to generate and absorb innovation, but also as a way of ultimately promoting greater economic growth. Having a good employment/unemployment balance is also equally important for innovation and economic growth. Fiscal incentives can also become useful in fostering innovation, if they help attract companies with a high innovative potential. These socioeconomic conditions weave a complex substratum that allows certain territories to become more innovation prone than others.

Second, regional clusters have a strong association with economic growth in the static model, especially when they help increase the knowledge flow in already highly integrated communities, among well endowed with firms, skilled workers, researchers and scientists. However they appear only as 'second-best factors' in relation to the social filter. This may be partly a result of the way the clusterisation effect is measured in the analysis. The method used may have introduced, as the European Cluster Observatory explains, "a bias towards employment-intensive clusters" (ECO, website). Therefore, these data will need to be completed by other information – not yet available at the European level – such as "wage bill, productivity or value added [in order] to shift the balance in favour of capital–or knowledge-intensive cluster categories" (*ibid*). In any case, the results may also highlight that the association between the presence of clusters, innovation, and economic development in the regions of Europe is (a) contingent on the presence of adequate social filters that would help make the transition from a mere cluster of firms into a real regional system of innovation, and (b) less relevant in time than the socioeconomic substrata on which the clusters are based. Clusters seem to matter when they become the hub for regional systems of innovation, but this tends to happen only when they are located in innovation prone environments with adequate social filters and even in these cases, their influence seems to be weaker than, for example, investment in R&D. Hence, the influence of clusters for economic growth may be lower than what many think. What really matters for economic growth is setting up in every territory the adequate conditions for innovation, including greater education and life-long learning opportunities, a better and more efficient use of human resources, a better matching of investment in training and innovation to local production fabric and more emphasis in science and technology.

The third conclusion is the limited short-term association between R&D investment and patent applications and economic development across the regions of Europe. However, the presence of adequate social conditions helps improve the returns on R&D investment and patents over time.

The research presented here probably sends a message of warning against the adoption of one-size-fits-all and even 'mesmeric' types of cluster policies for local economic development (Taylor 2010). Policies aimed at fostering or encouraging the agglomeration of firms may, without paying attention to local conditions and potential, end up yielding lower results – if at all – than expected. Indeed the analysis points towards the need of addressing local social filter bottlenecks as a precondition for achieving greater returns in R&D and in cluster policies. However, neither all clusters have the same transactions costs and internal relations characteristics, nor the same technological regimes and knowledge features (Iammarino and McCann 2006). This implies a need to make greater distinctions in policy-making among different types of clusters, as different clusters in different contexts may require different types of intervention (Gordon and McCann 2000). In any case, while the analysis presented here provides a springboard for some potential practical policy implementations and recommendations, it also calls for further research, and in particular of research trying to better reproduce and capture the effects of different types of clusters.

Appendix A. Annex 1: The Exact Formula of *Specialisation* and *Focus*

These formulas are directly extracted from the European Cluster Observatory's website: www.clusterobservatory.eu

A.1. Specialisation Quotient

$$SQ_{r,s} = \frac{e_{r,s}/E_s}{E_r/E}$$

SQr,s = the specialisation quotient for region r and cluster sector s
er,s = the number of employees for region r and cluster sector s
Es = the total number of employees in all regions for sector s
Er = the total number of employees in all cluster sectors for region r
E = the total number of employees in all regions and all cluster sectors

Put in a simpler way the Specialisation Quotient is given by

$$\frac{(\text{Employment in a region in a category}) \,/\, (\text{Total employment in a region})}{(\text{Employment in a category in Europe}) \,/\, (\text{Total employment in Europe})}$$

A.2. Focus

$$F_{r,s} = \frac{e_{r,s}}{E_r}$$

er,s = the number of employees for region r and cluster sector s
Er = the total number of employees in all cluster sectors for region r

Appendix B. Annex 2: PCA Analysis

In this annex, the results of the three Principal Components Analyses are given

B.1. Principal Component Analysis for Social Filter

Table A.5 Eigenanalysis of the correlation matrix: social filter

Component	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Comp8
Eigenvalue	3.09646	1.56254	1.11289	0.80251	0.518857	0.439485	0.34847	0.118794
Proportion	0.3871	0.1953	0.1391	0.1003	0.0649	0.0549	0.0436	0.0148
Cumulative	0.3871	0.5824	0.7215	0.8218	0.8867	0.9416	0.9852	1

Table A.6 Coefficients of the PCA: social filter

Variable	Comp1	Comp2
LT unemployment	−0.2879	−0.1887
Agriculture employment	−0.4096	0.2502
Corporate tax rate	−0.0119	−0.7014
Young people	−0.3441	0.4336
Total population	0.1056	0.3473
Education	0.4435	0.1291
Life long learning	0.4175	0.2867
HR in science and technology	0.4986	−0.0495

B.2. Principal Component Analysis for Clusterisation Index

Table A.7 Eigenanalysis of the correlation matrix: clusterisation index

Component	Comp1	Comp2	Comp3
Eigenvalue	1.48376	0.807631	0.708606
Proportion	0.4946	0.2692	0.2362
Cumulative	0.4946	0.7638	1

Table A.8 Coefficients of the PCA: clusterisation index

Variable	Comp1	Comp2
Specialisation	0.6046	−0.2846
Focus	0.5899	−0.4676
Diversification index	0.5352	0.8369

References

Abramovitz M (1986) Catching up, forging ahead, and falling behind. J Econ Hist 46:385–406

Abramovitz M (1990) The catch-up factor in postwar economic growth. Econ Enquiry 28:1–18

Aghion P (2006) A primer on innovation and growth. Bruegel policy brief, Issue 2006/06, Oct 2006. Available at http://www.bruegel.org/Files/media/PDF/Home/Primer_Innovation_Growth_13Oct.pdf

Anselin L, Varga A, Acs Z (1997) Local geographic spillovers between university research and high technology innovations. J Urban Econ 42:422–448

Aydalot P (1986) Milieux innovateurs en Europe. GREMI, Paris

Baptista R, Swann P (1998) Do firms in clusters innovate more? Res Policy 27:525–540

Bathelt H, Malmberg A, Maskell P (2004) Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. Prog Hum Geog 28:31–56

Becattini G (1987) Mercato e forze locali. Il distretto industriale. Il Mulino, Bologna

Braczyk HJ, Cooke P, Heidenreich M (1998) Regional innovation systems. UCL Press, London

Camagni R (1995) The concept of innovative milieu and its relevance for public policies in European lagging regions. Pap Reg Sci 74:317–340

Capello R (1999) Spatial transfer of knowledge in high technology milieux: learning versus collective learning processes. Reg Stud 33:353–365

Cheshire P, Malecki EJ (2004) Growth, development, and innovation: a look backward and forward. Pap Reg Sci 83:249–267

Cooke P (2001) Regional innovation systems, clusters, and the knowledge economy. Ind Corp Ch 10:945–974

Cooke P (2006) Regional innovation systems as public goods, publication of UNIDO Strategic Research and Economics Branch

Cooke P, Gómez Uranga M, Etxeberria G (1997) Regional innovation systems: institutional and organizational dimensions. Res Policy 26:475–491

Cooke P, Gómez Uranga M, Etxeberria G (1998) Regional systems of innovation: an evolutionary perspective. Environ Plann A 30:1563–1584

Cooke P, Morgan K (1998) The associational economy: firms, regions, and innovation. Oxford University Press, Oxford

Crescenzi R, Rodríguez-Pose A, Storper M (2007) The territorial dynamics of innovation: a Europe-United States comparative analysis. J Econ Geog 7:673–709

Crouch C, Le Galès P, Trigilia C, Voelzkow H (eds) (2001) Local production systems in Europe. Rise or demise? Oxford University Press, Oxford, pp 25–45

Cumbers A, Mackinnon D, Chapman K (2003) Innovation, collaboration, and learning in regional clusters: a study of SMEs in the Aberdeen oil complex. Environ Plann A 35:1689–1706

Doloreux D, Parto S (2005) Regional innovation systems: current discourse and unresolved issues. Technol Soc 27:133–153

Dosi G (1988) Sources, procedures and microeconomic effects of innovation. J Econ Lit 26:1120–1171

EC (2006) Praxis manual for innovation policy-makers and practitioners, EC Publication: DG Enterprise & Industry

Evangelista R, Iammarino S, Mastrostefano V, Silviani A (2002) Looking for regional systems of innovation evidence from the Italian innovation survey. Reg Stud 36:173–186

Fagerberg J (1987) A technology gap approach to why growth rates differ. Res Policy 16:87–99

Fagerberg J (1994) Technology and international differences in growth rates. J Econ Lit 32:1147–1175

Gordon I, McCann P (2000) Industrial clusters: complexes, agglomeration and/or social networks. Urban Stud 37:513–532

Howells J (1999) Regional systems of innovation? In: Archibugi D, Howells J, Michie J (eds) Innovation policy in a global economy. Cambridge University Press, Cambridge, pp 67–93

Iammarino S (2005) An evolutionary integrated view of regional systems of innovation: concepts, measures and historical perspectives. Eur Plan Stud 13:497–519

Iammarino S, McCann P (2006) The structure and evolution of industrial clusters: transactions, technology and knowledge spillovers. Res Policy 35:1018–1036

Jacobsson S, Bergek A, Carlsson B, Lindmark S, Rickne A (2006) Analyzing the dynamics and functionality of technological innovation systems: a scheme of analysis. Report delivered to VINNOVA

Keeble D, Lawson C, Moore B, Wilkinson F (1999) Collective learning processes, networking and 'institutional thickness' in the Cambridge region. Reg Stud 33:319–332

Maclaurin WR (1953) The sequence from invention to innovation and its relation to economic growth. Q J Econ 67:97–111

Martin R, Sunley P (2003) Deconstructing clusters: chaotic concept or policy panacea? J Econ Geog 3:5–35

Melachroinos KA, Spence N (2001) Manufacturing productivity growth across European Union states: 1978–94. Environ Plann A 33:1681–1703

Moore G, McKenna R (1999) Crossing the chasm: marketing and selling high tech products to mainstream customers. Harper, New York

Moreno R, Paci R, Usai S (2005) Spatial spillovers and innovation activity in European regions. Environ Plann A 37:1793–1812

Morgan K (1997) The learning region: institutions, innovation and regional renewal. Reg Stud 31:491–503

Porter ME (2000) Location, competition, and economic development: local clusters in a global economy. Econ Dev Q 14:15–34

Porter ME (2008) Clusters and competition: new agendas for companies, governments, and institutions. In Porter ME (ed) On competition. Harvard Business Review Books, Cambridge, MA, pp 213–304

Rodríguez-Pose A (1999) Innovation prone and innovation averse societies: economic performance in Europe. Growth Change 30:75–105

Rodríguez-Pose A, Crescenzi R (2008) R&D, spillovers, innovation systems and the genesis of regional growth in Europe. Reg Stud 41:51–67

Scott AJ (1988) New industrial spaces: flexible production organization and regional development in North America and Western Europe. Pion, London

Sonn JW, Storper M (2008) The increasing importance of geographical proximity in technological innovation: an analysis of U.S. patent citations, 1975–1997. Environ Plann A 40:1020–1039

Spencer GM, Vinodrai T, Gertler MS, Wolfe DA (2010) Do clusters make a difference? Defining and assessing their economic performance. Reg Stud 44:697–715

Storper M, Venables AJ (2004) Buzz: face-to-face contact and the urban economy. J Econ Geog 4:351–370

Taylor M (2010) Clusters: a mesmerising mantra. Tijdschr Econ Soc Ge 101:276–286

Varga A (2000) Local academic knowledge spillovers and the concentration of economic activity. J Reg Sci 40:289–309

Wolfe DA, Gertler MS (2004) Clusters from the inside and out: local dynamics and global linkages. Urban Stud 41:1071–1093

Wooldridge J (2006) Introductory econometrics: a modern approach. South-Western Publishing company, Mason, Third edition international student edition

Input–Output Linkages, Proximity to Final Demand and the Location of Manufacturing Industries

Giordano Mion

Abstract In this chapter I develop an empirical framework to estimate the role of agglomeration externalities, especially those stemming from input–output linkages, in the location process of US manufacturing plants. Furthermore, drawing on the model of Holmes and Stevens (J Econ Geogr 4: 227–250, 2004b), I propose a way to reconcile some previous puzzling results about proximity to consumers' demand and the scope of agglomeration forces. Results suggest that flows of intermediate goods have a positive impact, especially for big plants, on local specialization. By contrast, consumers' demand has a negative effect and this result is consistent with theory. However, the majority of both effects comes from very local interactions with spatial spill-overs being quite weak but with a very large geographical scope. This result suggests some kind of strong non-linearity in the underlying spatial process. Very close interactions are extremely important but, when considering what is beyond the limit of local markets, then distance does not matter so much.

Keywords Manufacturing concentration • Input–output linkages • Agglomeration externalities • Market proximity

---

JEL Classification: L60, R12, R15, R31, R34.

I'm very grateful to Vernon Henderson and Eric Strobl for fruitful discussions and comments. The usual disclaimer applies.

G. Mion (✉)
Department of Geography and Environment, London School of Economics (lse), UK cep, London, UK and cepr, London, UK
e-mail: g.mion@lse.ac.uk

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_12,

1 Introduction

Market proximity is a key feature in location theory and in particular in the so-called New Economic Geography (NEG) literature. The interaction of transport costs and increasing returns to scale creates incentives for non price-taking firms to cluster together in order to be close to large markets. As shown in Fujita et al. (1999), there are at least two channels trough which this agglomeration process operates: input–output linkages and final consumers' demand.

The first mechanism is emphasized in Krugman and Venables (1995) and Puga (1999), and the basic idea behind it is that industries characterized by strong input–output relations can both save trade costs and have a richer availability of the different varieties of each product by choosing the same location. A recent empirical paper by Amiti and Cameron (2007) on Indonesian data provides evidence that good access to production and demand of manufacturing intermediates has a positive effect on manufacturing nominal wages. Furthermore, they found that the geographical scope of such an effect is not negligible, with proximity to input suppliers being significant up to 230 km. However, in a related paper, Rosenthal and Strange (2003) found a much smaller scope for agglomeration economies. Studying US new plants' location, they actually found that new arrivals are more likely to be attracted as employment in the own industry increases up to only 15 miles.

A second force pushing towards the clustering of activities is the proximity to consumers' demand. As shown in Krugman (1991) and Ottaviano et al. (2002), firms want to be close to large consumer markets in order to exploit their increasing returns to scale technology. Both Hanson (2005) and Mion (2004) found evidence that proximity to households' disposable income has a positive effect on aggregate factor remunerations. Furthermore, when looking at local sectoral wages, Amiti and Cameron (2007) found that proximity to demand increases wages, this effect being significant up to 85 km. By contrast, in their study of French manufacturing location, Gaigné et al. (2003) found a negative effect of demand in surrounding regions on local industry specialization.

Instead of being simply a result of country differences, I will provide later on some theoretical arguments that can reconciliate these (apparently) conflicting results. In particular, building on Holmes and Stevens (2004b) model, I will show how it is reasonable that agglomeration externalities turn into different patterns for plants' location compared to local wages analysis. Once I draw a link between theory and these recent empirical findings, I will proceed through the main part of the chapter. In particular, using manufacturing data on US counties, I will regress a measure of industrial concentration on the major determinants of location in NEG models: access to input–output flows, access to consumers' demand, and local wages. This will allow me to provide further evidence of the importance of agglomeration externalities for industrial location as well as to test the a priori on the signs of parameters.

Particular attention will be devoted to measure the degree of spatial interaction by estimating a specific distance decay parameter coming from a polynomial function. Results suggest that all variables are extremely relevant at very localized level. Both input–output goods demand and supply in a county positively affect the concentration of a an industry in that location. By contrast, consumers' demand has a negative effect and this result is consistent with the model of Holmes and Stevens (2004b). As for distance effects, demand and supply coming from other locations have a very weak but significant effect with an elasticity of around 0.5. The relatively stronger importance of local variables is in line with the findings of Rosenthal and Strange (2003), while the surprisingly slowly declining distance effect is a brand new result. This suggests some kind of strong non-linearity in the underlying spatial process. Very close interactions are crucial. However, when considering what is beyond the limit of local markets, then distance does not matter so much. The breakdown of the agglomeration analysis by establishments size is also a novelty of the chapter. In particular, data suggests that input–output flows (especially at very local level) play a much more important role in the location process of big plants, while small ones can be found more easily close to big consumers' markets.

This chapter certainly belongs to the strand of literature that is concerned with the determinants of location and regional growth that includes, among others, D'Costa et al. (2013), Ellison and Glaeser (1999), Rodriguez-Pose and Comptour (2013), and Rosenthal and Strange (2001).[1] The importance I give to input–output linkages and market access within a NEG framework is common to other papers like Davis and Weinstein (1999), Combes and Lafourcade (2001) and Redding and Venables (2004). This second strand of literature is characterized by the desire to make the empirical analysis more directly related to structure coming from the underlying theory. In particular, the question of the scope of agglomeration externalities (that I also deal with) is a specific concern in both Mion (2004) and Amiti and Cameron (2007). Finally, my comparison of local vs spatially weighted variables is common to both Henderson (2003) and Rosenthal and Strange (2003).

The rest of the chapter is structured as follows. Section 2 introduces the reader to the problems associated when trying to link multi-industry models to data evidence, and contains a brief exposition of Holmes and Stevens (2004b) model. Section 3 deals with the construction of the variables used in the regression analysis as well as with data sources. Drawing on previous studies on spatial externalities, I derive in Sect. 4 the basic econometric specification. In Sect. 5 I present estimations obtained with the baseline specification, as well as a number of robustness checks and more disaggregated results. Finally, Sect. 6 concludes and proposes directions for further research.

2 Insights from Theory

This section presents some stylized facts about the location of economic activities in the US and raises the issue of the apparent contrast of some of these features with NEG predictions. The subsequent analysis of Holmes and Stevens (2004b) model will allow to reconciliate these inconsistencies with theory.

[1] See Rosenthal and Strange (2004) for a review of the literature.

2.1 Some Stylized Facts

As convincingly pointed out in Holmes and Stevens (2004b) manufacturing in the US is a "rural" activity compared to services. Data on US cities from the 1997 Economic Census are organized in such a way that the 328 primary metropolitan statistical areas (PMSAs) are divided into three city sizes: small (PMSA population under half a million), medium (PMSA of half a million to two million), and large (PMSA over two million). Let us consider *LQ* as the share of industry sales of establishments located in a given city size class divided by the share of population in that size class. Holmes and Stevens (2004b) show that the service industries – wholesale trade, finance and insurance, and professional, scientific, and technical services – are heavily concentrated in the large cities group (the *LQ* increases from approximately 0.5 in the smallest cities to approximately 1.3 in the largest cities for the three service sectors). By contrast, manufacturing activity is concentrated in small cities (the *LQ* declines from 1.15 to 0.86). Further evidence of this pattern is provided in Holmes and Stevens (2004a). They show that in the U.S. economy rural areas are actually net exporters of manufactured goods while large cities are net importers.

One easy explanation of such a phenomenon is that manufacturing is relatively more land consuming than services while land rents are increasing in population size. There are certainly many other stories one can tell, but the key point is that, at first glance, it would be difficult to find some consistent story *within* the NEG framework. In fact, in all NEG models the increasing returns sectors producing differentiated products is thought to be manufacturing. Therefore manufacturing is expected to display, more than other activities, a tendency to concentrate close to big markets (the so called market access effect). This is clearly in sharp contrast with the evidence I presented so far, casting some doubts on the relevance of the theory. The first empirical framework that faced this inconsistency is the analysis of French manufacturing activities by Gaigné et al. (2003). When regressing an index of local industry specialization on consumers' income coming from surrounding locations they actually found a negative and significant coefficient (thus implying a negative market access effect).[2]

The model of Holmes and Stevens (2004b) can accommodate these results by coming back to one fundamental element of the NEG literature: differences in transportation costs. Transportation costs are particularly high in the service sector. Face-to-face communication is often crucial with services, and the transportation costs of moving people is surely higher than the cost of moving goods. Therefore, service activities concentrates in large cities because a large home market makes it possible to both economize on the cost of moving people and to achieve economies of scale. By contrast, the manufacturing sector concentrates in small cities. With the relatively low cost of shipping goods rather than moving people, manufacturing plants in small cities can in fact obtain scale economies by shipping to a national market. To put it differently, both sectors want to be in the large cities, but service

[2] Gaigné et al. (2003) considered this result as misleading and did not try to interpret it.

activities have, due to their relatively higher transportation costs, more incentives to "be in the center" and can manage to pay the higher rents and wages one has to afford in big cities.

Another puzzling issue in applied analysis is the considerable variation in the geographical scope of agglomeration externalities found in different studies. When looking at the spatial pattern of industry wages, Amiti and Cameron (2007) found that proximity to local demand and inputs matters up to 230 km. By considering aggregate factor earnings in Italy, Mion (2004) found a similar result in that access to consumers' demand was significant up to 200 km. By contrast, in their analysis of new plants' location, Rosenthal and Strange (2003) found a much smaller scope for agglomeration economies with new arrivals being more likely to be attracted as employment in the own industry increases up to only 15 miles. Coherently with the latter result, Gaigné et al. (2003) found that both input–output linkages and consumers' demand coming from neighbor locations have a weak impact on local industrial specialization.

Holmes and Stevens (2004b) offer a novel way to solve this puzzle. The basic insight of the reasoning contained in their model is that labor is, to a great extent, an homogenous good for which many sectors compete at the same time. For each sector, the benefit of being in a certain location decreases with the distance from big central markets. This creates a kind of wage gradient, that represent the maximum price that firms of a given sector are willing to pay in order to produce and hire people in a certain place. Sectors like services have an higher wage gradient close to the center because high transportation costs make centrality crucial. They can thus manage to over-bid other sectors and locate close to big cities while other activities like manufacturing stay in peripheral areas. Nevertheless, the wage gradient coming out from this competition for location/workers is the upper bound of the wage gradients of the different sector which cannot be really separated from each other. Consequently, wages will decline slowly with distance from the center, resulting in the observable large spatial scope for agglomeration externalities that has been found in Amiti and Cameron (2007) and Mion (2004).

By contrast, when looking at local specialization, the link one can trace between concentration of production and access to demand and inputs is more sector specific. This should reasonably lead to a smaller spatial scope for agglomeration externalities, as it is the case in Gaigné et al. (2003), and Rosenthal and Strange (2003). Nevertheless, the true extent of spatial externalities should be, parallel to the case of wages, underestimated. This reasoning can also help explain why proximity to consumers' demand is usually found to be positive in manufacturing wage analysis, while the converse is true for specialization regressions.

Contrary to input–output relations, which are always found to have a positive impact, competition for consumers' proximity is highly rival. The access to supply and demand of input goods is in fact industry specific and so there is no clear consensus on an ideal location for all sectors. Each industry can thus find its own place where it can benefit from the proximity of strategic partners. By contrast, everybody would like to be close to final demand, and there is thus competition for the ideal location. As I mentioned before, manufacturing has less incentives than others to be in the center, and equilibrium forces push it toward rural areas, i.e. places where there is a bad access to consumers' demand. This justifies why one can observe a negative correlation between the specialization of a region in manufacturing and its access to final demand. On the other hand, wage analyses do not allow to isolate the equilibrium relation between these two variables and pick up the upper bound willingness to pay of the different sectors to be close to the center. This happens because the pressure of all such sectors on local labor markets contributes to determine the location-specific wage.

2.2 The Model

I will now present some basic features of Holmes and Stevens (2004b) model to which the reader may refer for a detailed exposition. Imagine a continuum of goods indexed by $x \in [0,1]$. For each good at each location, the authors further assume that a certain component of the local demand for the good must be satisfied by local production and that a second component may be satisfied either by local production or by imports.[3] Each good is thus further indexed by $k \in 1,2$, where k indicate whether the good must be locally produced (k = 1) or can be imported (k = 2). The representative consumer has a Cobb-Douglas utility function:

$$U = \int_0^1 \lambda \ln q(x,1) + (1-\lambda) \ln q(x,2) dx \tag{1}$$

where $q(x,k)$ denotes the consumption of a commodity (x,k). In equilibrium, the preference parameter λ will denote the share of income spent on goods that must be locally produced.

The technology for production is the same across all goods and labor is the only factor of production. However, producers can chose among two alternatives. A backstop constant returns to scale technology, that is supposed to be the more efficient when production is limited to local consumption, and an increasing to scale one that is more efficient when goods are also exported. In particular let $c(q)$ denote the labor input required to produce q units of a particular good and define the average cost (in labor units) as $a(q) = c(q)/q$. For the increasing returns to scale technology the average cost is supposed to be minimized at a point $q^* \in (0,\infty)$, where q^* is defined as the lowest point at which the minimum is attained (i.e., q^* is the minimum efficient scale) and $a^* = a(q^*)$. By contrast, the other technology requires γ units of labor to obtain one unit of output independently of the scale of production and $\gamma > a^*$.

[3] Consider, for example, a service good called "grocery wholesaling". One component of the demand for this good is for the wholesaling of milk, an important service that must be provided locally because of milk's perishability. A second component, the wholesaling of crackers, could be satisfied by imports because perishability is not an issue.

There is a continuum of locations indexed by $i \in [0, 1]$. Let $l(i)$ be the population of workers (each endowed with one unit of labor) at location $i \in [0, 1]$, and assume that population is strictly increasing in i. Now, let $\underline{l} = l(0)$ be the minimum population and $\bar{l} = l(1)$ be the maximum population. Population of each location is fixed, i.e. there is no labor mobility.

While labor is immobile, goods can be traded. The transportation cost of shipping a good from one location to another is τ that is independent from distance. The transportation cost takes the standard "iceberg" form: in order to move one unit of a good to another location, $1 + \tau$ units must be shipped, so τ units get lost in transit. Goods are allowed to differ in transportation cost. There are three different levels of transportation cost and goods with the same transportation cost are regarded as being in the same sectors. The three sectors are manufacturing, services, and retail. Let μ, σ, and ρ be the fraction of industries in the three sectors, where $\mu + \sigma + \rho = 1$. Assume the goods are ordered such that $x \in [0,\mu)$ are manufactured goods, $x \in [\mu,\mu + \sigma)$ are services, and $x \in [\mu + \sigma,1]$ are retail goods. Make the extreme assumption that manufactured goods have zero transportation cost, $\tau_M = 0$, retail transportation costs are infinite, $\tau_R = \infty$, and service transportation costs are intermediate, $\tau_s \in (0,\infty)$. Note that goods differ across sectors only in transportation costs; the production technology and the way the goods enter the utility function is otherwise symmetric.

Finally, the authors assume a local strategic structure and zero profits in the long run. Firms take prices in other locations as given and behave as perfect competitors in the export market (due to the continuous number of producers in the entire economy). By contrast, in the local market the industry is oligopolistic and firms are assumed to compete in Bertrand fashion. The price a producer receives per unit of exports is the same across all goods belonging to a given sector. Being aware that this export price differs across the three sectors $\left(p_R^E \neq p_M^E \neq p_R^E\right)$, one can call a generic sectoral price as p^E. A consumer at any location can then import any *non-local* good at a price of $p = p^E(1 + \tau)$, the competitive export price with transportation costs added. The combination of scale economies, a small market size at each location, and Bertrand competition, together imply that there is at most one producer of any particular good at each location. That producer sets a limit price of $p^L = p^E(1 + \tau)$ in the local market for importable goods that exactly matches the import price. As for the local price of non-importable goods, this is limited by the possibility that a firms enters in the market and uses the backstop technology to produce it. Therefore this price is, by the zero profit condition, equal to the unit cost of production – $w(i)\gamma$ – where $w(i)$ is individual wage in location i.

Let's now characterize the equilibrium of this economy. Consider first the retail sector. This sector has infinite transport costs so its goods are not traded. Therefore, since the backstop technology is supposed to be the more efficient in autarky, it is used everywhere for retail production. The utility function is Cobb-Douglas, so one knows that in equilibrium, a fraction ρ of the labor force in each location will be employed in retail using the backstop technology, and one can otherwise ignore this sector. Next, consider the joint effect of differences in populations across locations and differences in transportation costs across the two traded sectors, manufacturing

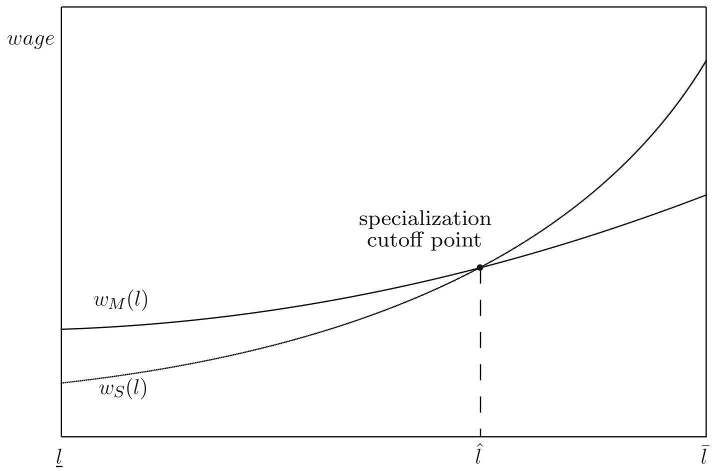


Fig. 1 Bid-wage functions and specialization (20,10)

and services. Each location must produce the non-importable ($k = 1$) varieties of manufactures and services. Production of the non-importable goods and services is carried out using the constant-returns-to-scale backstop technology. The production of non-retail importable goods occurs at firms using the scale technology, and each location will have a single firm using the scale technology for any given product. Because locations differ in population, each location will have a measure of firms using the scale technology for distinct goods within the same sector. The determination of which locations specialize in services and which specialize in manufactures is made on the basis of wages, as the immobile workers will work for the firm that offers the highest wage. The tension between market access and cheap labor makes low (high) transportation cost sectors better suited for small (large) locations.

In particular the authors show that, once one normalizes the export price of manufacturing to one ($p_M^E = 1$) there exists a unique population cutoff $\hat{l}(p_S^E, \tau_S)$, such that the manufacturing and service wages are equal. Locations with $l \geq \hat{l}$ specialize in services; locations with $l < \hat{l}$ specialize in manufacturing. This specialization can be explained by means of Fig. 1 borrowed from Holmes and Stevens (2004b). Rearranging equilibrium equations, one finds that zero-profit wages are function of $l(i)$ and the sector of specialization (which implies a value for p^E and τ):

$$w(p^E, \tau, l(i)) = \frac{p^E}{c'(\tilde{q}(l(i), \tau))} \tag{2}$$

where $c'(.)$ is marginal cost and $\tilde{q}$ is equilibrium output. This function is analogous to the bid-rent function in urban economics (see Mills and Hamilton 1994), except that Holmes and Stevens (2004b) call it a bid-wage function. As in urban economics, the bid-wage function determines what activities occur at what locations. Let

$w_S(p_S^E, \tau_S, l(i))$ denote the bid-wage function for services, and let $w_M(1, 0, l(i))$ denote the bid-wage function for manufacturing (recall that the manufacturing export price is 1 and $\tau_M = 0$). Figure 1 plots these functions. Both equilibrium wage rates increase with location size. This rise occurs because larger locations are relatively more attractive to firms with more price-setting power over a larger local base of consumers. The activity that supports the highest wage is the activity that the location specializes in. For manufacturing, access to low-cost labor is more important than proximity to markets and so the curve w_M dominates w_s for $l < \hat{l}$. By contrast, market access is relatively more important for service that incur in a positive cost τ_s in order to ship goods to other markets and so services concentrate close to consumers' markets ($l \geq \hat{l}$). From this simple Figure, one may draw most of the conclusions I discussed in the previous subsection:

- Consistently with what said about manufacturing as being characterized by less transportation costs than services, industrial activities concentrate far from the center ($l < \hat{l}$) suggesting a negative correlation between specialization in manufacturing activities and consumers' market access.
- The local bidding process for wages implies that they are positively correlated to consumers' markets access for any sector considered (both w_M and w_s increases in population size).
- The distance pattern of wages (which corresponds to the highest value of the two functions in Fig. 1) is a kind of sum of the forces at work and wages over-evaluate both the absolute value and the geographical scope of individual sector externalities.
- If one just regards local specialization, then agglomeration externalities are under-estimated. What one really needs in order to disentangle sector specific externalities is in fact the shape of the two functions w_s and w_M. Nevertheless, simply looking at the size of area in which one industry specialize – i.e., the segment $\hat{l} - \underline{l}(\bar{l} - \hat{l})$ for manufacturing (services) – gives a downward bias estimate of the degree of spatial interdependence among location choices.

3 Data Construction and Sources

In the next section I will be using manufacturing data on US counties in order to perform a regression of a measure of industrial concentration on the major determinants of location in NEG models: access to input–output flows, access to consumers' demand, and local labor costs. The units of analysis are the 3,111 continental US counties excluding the States of Alaska and Hawaii. The reference year for data is 2000 with the exception of the input–output information which refers to 1997.[4] Data on plants' employment and location comes from County

[4] This is the closest year for which input–output tables are available with the NAICS classification.

Business Patterns (CBP) data, used in Holmes and Stevens (2004a), and made available (together with other local databases) by Thomas J. Holmes on his web-site.[5] There are 7,070,048 establishments in the CBP universe for the year 2000. Data are organized in cell counts giving the number of establishments belonging to a given combination of industry (six digit NAICS), employment size class, and county. Data on local employment has thus been reconstructed using the (national) average employment for each size class. It is important to stress that, compared to other data on US plants, there is no problem of missing observation due to confidentiality reasons. All US plants are in the data-set, and for each of them I know the employment size class to which it belongs.

As for input–output flows, I used the (benchmark) Table "Use of Commodities by Industries" coming from the Bureau of Economic Analysis (that refers to US as a whole) for the year 1997. Table 1 lists those sectors (66) that have been used in order to compute both the demand for intermediate use and the input needs of each sector.[6] This commodity based classification is the one actually used in regressions. The database gives a very detailed information on input–output flows in million of US dollars and with a disaggregation that roughly corresponds to the NAICS four digits. Importantly, there is a unique correspondence between the six digits NAICS data on establishments and this classification. All flows among these sectors have been considered. In particular, assuming a Cobb-Douglas production function (like in Combes and Lafourcade 2001), and indicating with r_c the ratio between total input needs and employment in a given industry c, one can derive the following measure of the demand for intermediate use of commodity c in county i ($IntD_{c,i}$):

$$IntD_{c,i} = ln\left(1 + \sum_{c'} \mu^{c,c'} r_{c'} \ emp_{c',i}\right) \tag{3}$$

where $\mu^{c,c'}$ is the share of intermediate use of commodity c in the production of c', and $emp_{c',i}$ is local employment of industry c' in county i. The measure in log in order to give an elasticity meaning to regression parameters. To this respect, adding one eliminates the problem of having the logarithm of zero whenever there is no local demand for intermediate use.[7] All variables in regression will be in log and the trick of adding one will be used for all except wages.[8]

[5] See http://www.econ.umn.edu/126holmes/research.html

[6] With respect to the BEA classification, I have aggregate sectors 2,301, 2,302 and 2,303 into a unique sector 2,300 (Construction). The same applies to sectors 332A and 332B that has been pooled into one industry 3325 (Ordnance and other fabricated metal products). The reason of these changes is to make the commodity classification of the BEA look as close as possible to the NAICS 4 digits. In such a way comparability with other studies should be substantially improved.

[7] A similar approach has been used by Redding and Venables (2004) in their analysis of trade flows.

[8] Whenever there is no wage for a couple sector-county, it is simply considered as missing. Contrary to other regressors, there is in fact no obvious value for a missing wage value.

Table 1 Industry used in the analysis

Industry code	Commodity
1110	Crop products
1120	Animal products
1130	Forestry and logging products
1140	Fish and other nonfarm animals
1150	Agriculture and forestry support services
2110	Oil and gas
2121	Coal
2122	Metal ores
2123	Nonmetallic minerals
2130	Mining support services
2211	Electric power
2212	Natural gas distribution
2213	Water and sewage treatment
2300	Construction
3110	Food products
3121	Beverage products
3122	Tobacco products
3130	Yarn, fabrics, and other textile mill products
3140	Non-apparel textile products
3150	Apparel
3160	Leather and allied products
3210	Wood products
3221	Pulp, paper, and paperboard
3222	Converted paper products
3230	Printed products
3240	Petroleum and coal products
3251	Basic chemicals
3252	Resins, rubber, and artificial fibers
3253	Agricultural chemicals
3254	Pharmaceuticals and medicines
3255	Paints, coatings, and adhesives
3256	Soaps, cleaning compounds, and toiletries
3259	Other chemical products
3260	Plastics and rubber products
3270	Nonmetallic mineral products
331A	Primary ferrous metal products
331B	Primary nonferrous metal products
3315	Foundry products
3321	Forgings and stampings
3322	Cutlery and handtools
3323	Architectural and structural metal products
3324	Boilers, tanks, and shipping containers
3325	Ordnance and other fabricated metal products
3331	Agriculture, construction, and mining machinery
3332	Industrial machinery

(continued)

Table 1 (continued)

Industry code	Commodity
3333	Commercial and service industry machinery
3334	HVAC and commercial refrigeration equipment
3335	Metalworking machinery
3336	Turbine and power transmission equipment
3339	Other general purpose machinery
3341	Computer and peripheral equipment
334A	Audio, video, and communications equipment
3344	Semiconductors and electronic components
3345	Electronic instruments
3346	Magnetic media products
3346	Electric lighting equipment
3352	Household appliances
3353	Electrical equipment
3359	Other electrical equipment and components
3361	Motor vehicles
336A	Motor vehicle bodies, trailers, and parts
3364	Aerospace products and parts
336B	Other transportation equipment
3370	Furniture and related products
3391	Medical equipment and supplies
3399	Other miscellaneous manufactured products

Using the same idea of exploiting the aggregate ratio between input needs and workforce as a way to map workers' distribution into commodities flows, I derive the measure of the supply of commodities for intermediate use to industry c in location i ($IntS_{c,i}$) as:

$$IntS_{c,i} = ln\left(1 + \sum_{c'} \mu^{c',c} r_c \ emp_{c,i}\right) \quad (4)$$

which can also be obtained starting from a Cobb-Douglas production function. Although input–output variables are constructed using our full sample of 66 sectors, estimations will be performed on manufacturing activities only, and in particular to industries with codes between 3130 and 3399 (49 sectors). Agriculture, and mining have in fact been excluded because their location heavily rely on natural resources. Food, beverage, and tobacco industries have instead been dropped from regressions because they are relatively raw-materials intensive and the use of such input may potentially be less well estimated. However, dropped sectors are still used in the construction of input–output variables.

One crucial point, that will receive further attention in the estimations, is the fact that a large fraction of input–output flows has an intra-sector nature. In my sample of 66 industries, roughly 22 % of these flows have the same sector as origin and destination. This is due to the degree of aggregation of the NAICS classification as well as to the phenomenon of outsourcing. The fact that right hand variables like $IntS_{c,i}$ and $IntD_{c,i}$ contain a measure of local industrial concentration (intra-sectoral flows), while local concentration is at the same time being used as dependent variable clearly raises an endogeneity issue. I will explicitly deal with this problem in Sect. 5.

Data on disposable personal income come also from the Bureau of Economic Analysis. Information is organized by county and values are in thousand of US dollars for the year 2000. Using the supplementary information on the share of consumers' expenditure by commodities (coming from the Bureau of Economic Analysis for the year 1997), I constructed the following measure of local demand from consumers:

$$ConsD_{c,i} = ln(1 + \gamma_c Y_i) \tag{5}$$

where γ_c is the share of local disposable income Y_i (measured in million of US dollars) devoted to commodity c.

Sectoral weekly wages in dollars comes from the Bureau of Labor Statistics, and in particular from the current employment statistics survey. Data refers to the year 2000 and are organized by six-digit NAICS industries. Both State and county wages are available but, due to confidentiality constraints, much of the information at the county level is withheld. Therefore, I decided to use State level data, associating to each county the corresponding (logarithm of the) State wages of a particular industry. Furthermore, in order to be compatible with the classification of economic activities presented in Table 1, information has been aggregated by weighting six-digits NAICS wages by the State employment of industries belonging to a given branch.

The use of State level wages is certainly a very rough approximation of the corresponding county earnings. However, since the focus of the chapter is mainly on demand and input–output linkages, wages are used here mainly as a control variable. More importantly, it will turn out that, neglecting them, leads to virtually the same results. Nevertheless, by considering wages one can still gain some useful insight into industries location and that is why I decided to keep them in the analysis.

Data on distances $d_{i,j}$ between any pair (i,j) of US counties are great-circle distances in miles that have been calculated with a GIS software. This represents a large amount of information as there are, excluding internal distances, 3111*((3111–1)/2) = 4,837,605 free elements to handle. Only few studies, like Henderson (2003) and Rosenthal and Strange (2003), have dealt with location of manufacturing and spatial effects with such a detailed geographical level. Information on counties' surface and the portion dedicated to crops production (in square miles) come from ESRI (Environmental Systems Research Institute) database for ArcView, while the water area come from Holmes' databases. Finally, data on the average percentage of possible sunshine days and a dummy variables for whether

Table 2 Summary statistics

	Mean	St. Dev.	Min	Max
Spec	0.2439	0.5797	0	6.0372
IntD	6.1955	2.8770	0	15.7792
IntS	6.4363	2.6086	0	16.4795
ConsD	5.1743	2.6649	0	15.2917
wage	6.6135	0.3019	5.5219	8.1345
Area	965.4524	1,313.6433	1.7314	20,174.7230
Crop surface	230.2999	236.1674	0	2,068.691
Water area	51.7273	207.7853	0	5,424.9820
% Sunshine days	54.7834	12.1476	19	78

Note: Variables Area, Crop surface, Water area and Sunshine refers to original data and are neither ratios nor logarithms. In particular Area, Crop surface, and Water area are expressed in square miles

the county borders the sea coast or borders on one of the great lakes are taken from the Statistical Abstract of the United States, year 2000. Table 2 contains summary statistics on the variables used.

4 Econometric Specification

The main goal of the empirical analysis is to draw a link between local specialization in manufacturing and the major determinants of location in NEG models. In particular, I will focus on intermediates flows (both from supply and demand point of view), consumers' demand and labor costs.

First of all, when dealing with the distribution of firms over space it is evident that it is best analyzed as a process with a latent variable (profitability), in which zeros have a special meaning. This is actually the approach followed in many papers that focus on location choice like Head and Mayer (2004). Only if location in a certain county is profitable will one observe some establishments there, while a negative expected profit will turn into a zero observation. Consequently, this natural censoring of data suggests a Tobit approach in pretty much the same way as trade flows analysis. Indeed, when looking at the distribution of my 49 industries among the 3,111 continental US counties there are a lot of zeros. Of the 152,439 possible combinations only 49,481 (less than one third) couples county-sector have a positive number of establishments. Neglecting all such an information, would result in a great loss in efficiency and could potentially bias results.[9]

[9] Applications of the Logit and Tobit models to the analysis of firms' location and market potential can be found in Combes (2000) and Head and Mayer (2004).

In particular, indicating with $emp_i = \sum_c emp_{c,i}$ local employment in county i, $emp_c = \sum_i emp_{c,i}$ employment in industry c, and $emp = \sum_c emp_c = \sum_i emp_i$ as national employment, the local specialization variable I will use is:

$$Spec_{c,i} = ln(1 + \frac{emp_{c,i}/emp_c}{emp_i/emp}) \tag{6}$$

This index takes a value of zero whenever no plant belonging to industry c has decided to locate in i. If, on the contrary, plants locates there then it takes (positive) increasing values in $emp_{c,i}$. Furthermore, the index gives (by construction) an inter-sectoral comparable measure of specialization that is necessary in order to pull observations together.

Another crucial topic in my analysis is the measurement of spatial effects, i.e., the contribution of neighbor observation in the regression. One solution, followed by Rosenthal and Strange (2003), is to use both the value that a regressor takes in a location as well as its value in neighbor locations, where those latter are identified by means of distance bands matrices. Constructing rings of different size (1 mile, 5 miles, 10 miles, and 15 miles), they first identify locations that are within a given distance from each location i. They then use as regressors the value of each right-hand side variable computed for location i as well as their average for the set of neighbor location belonging to a given distance interval. This methodology leads to having $n + 1$ regressors for each right-hand side variable, where n is the number of distance rings considered. The main advantage of this strategy is that there is no need to evaluate internal distances, i.e. the average distance theoretically covered by goods flows within a given county. The way one should measure such distances is in fact not clear and there is no general consensus in the profession on the way they should be computed. Nevertheless, the two major drawbacks of the methodology of Rosenthal and Strange (2003) are the arbitrarily of the cut-off points used in the construction of the rings and the multi-collinearity among spatial variables.[10]

An alternative strategy, used in the local wages analysis by Amiti and Cameron (2007), is to consider distances between all locations i in order to construct a unique "spatially weighted" variable. Starting from those distance $d_{i,j}$, they use a distance weight function, and in particular the inverse exponential $e^{-\tau d_{i,j}}$, in order to attribute a weight to each observation corresponding to county j before summing them up. Compared to Rosenthal and Strange (2003), they thus use only one variable in order to capture spatial effects, with the estimate of τ giving a measure of their scope. This should reasonably lead to a more accurate measurement of spatial interaction, while allowing to get rid of the multicollinearity problem. However, in order to avoid attributing internal distances, they did not consider in their sum the value that a variable takes in location i. This can indeed lead to a missing variable problem,

[10] The variables constructed with distance bands are in general highly collinear among themselves. This lead to reject the significance of more distant observations because of their poor (biased) t-ratios.

and this is particularly cumbersome in my framework because specialization is, more than local wages, likely to be affected by forces internal to a county. Furthermore, as shown in Mion (2004), the use of the inverse exponential is probably not the best solution, with the polynomial function being more suited for the analysis of trade-based models like NEG ones.

The strategy I will use is to try to get the best from these two alternatives. In particular, I will use all distances (except internal ones) between county pairs (i, j) in order to construct, for each of the three variables that potentially have a distance declining effect (*IntD*, *IntS*, and *ConsD*), a corresponding "spatially weighted" variable (*W_IntD*, *W_IntS*, and *W_ConsD*). To each of these three new measures, I will associate a different distance decay parameter (τ_1, τ_2, and τ_3) that will allow to measure the corresponding geographical scope. In particular, I will use polynomial weights of the type $d_{i,j}^{-\tau}$, with τ thus giving the elasticity with respect to distance. Furthermore, in order to account for externalities within a county, I will include the corresponding county-based observations (*IntD*, *IntS*, and *ConsD*) ad additional regressors. Therefore, the base equation I will estimate is – omitting subscript (i,s) – given by:

$$Spec = \beta_1 IntD + \beta_2 IntS + \beta_3 ConsD + \beta_4 wage + \lambda_1 W\_IntD + \lambda_2 W\_IntS + \lambda_3 W\_ConsD + \varepsilon \quad (7)$$

where, for example, the variable $W\_IntD = W^1 IntD$ is constructed pre-multiplying *IntD* by the spatial weighting matrix W^1 whose generic element is the weight $w_{i,j}^1 = d_{i,j}^{-\tau_1}$, and $w_{i,i}^1 = 0$.[11]

In order to account for the problem of zeros, Eq. 7 will be estimated using maximum likelihood. For any given value of τ_1, τ_2, and τ_3, Eq. 7 has in fact all the features of a censored regression of the Tobit type. Therefore, maximum likelihood is needed and maximization can be achieved quite easily.

Expressions like 7 can be obtained as reduced form equations of a multi-industry NEG models. Nevertheless, I prefer to stick here to a qualitative specification because multi-industries models leads either to complicated long formulas that are difficult to interpret (like in Combes and Lafourcade 2001) or, in order to get something meaningful, one has to impose very strong assumptions (like in Gaigné et al. 2003). A qualitative approach might thus turn out to be more useful if it is coupled with a theoretical ground that gives a clear view of how results should be interpreted. To this respect, I suggest that there is sufficient consensus in the literature about the positive impact of input–output linkages on local concentration, while the model of Holmes and Stevens (2004b) provides sufficient insights to deal with the location of heterogenous sectors and consumers' markets centrality.

[11] For the parameters λ_1, λ_2, and λ_3 to be interpreted as elasticities, spatial variables should be constructed as averages of the corresponding local variables (that are already in log). Therefore, I normalize weights $w_{i,j}$ dividing them by their average row sum. In such a way, $\sum_j w_{i,j}$ will (on average) be equal to one.

5 Estimations

The base regression is reported in Table 3 where I estimate Eq. 7. I also provide several robustness checks in Table 5, while exploring more disaggregated data in Table 6. Standard errors are in parenthesis, while test statistics and information on the sample is given at the bottom of each Table.[12] As one can see, parameters are very precisely estimated and most of them are significant at the 1 % confidence level.

5.1 First Results

Columns (1) and (2) of Table 3 refer respectively to estimates obtained with local and spatially weighted variables only, while in columns (3) and (4) both sets of variables are included together with State dummies. A first striking feature that comes from the comparison of columns (1) and (2) is that local demand and input–output flows (*IntD*, *IntS*, and *ConsD*) have much more predictive power (as measured by both the pseudo R^2 and the Log-Likelihood) than spatial ones (*W_IntD*, *W_IntS*, and *W_ConsD*). The pseudo R^2 corresponding to the regression where only local variables are considered is in fact 0.2045 against the poor 0.0197 of the other. Furthermore, since all variables have been taken in log, one can infer that also the magnitude of the two sets of variables is quite different with local measures having a bigger impact. These results are coherent with those of Rosenthal and Strange (2003) who found little evidence of spatial effects for local specialization. Nevertheless, the scope of spatial spill-overs, as indicated by the distance elasticity parameters τ_1, τ_2, and τ_3, is very large. With an elasticity of about -0.4, input–output flows are such that doubling the distance leads only to a 25 % decrease in their spatial effect. This is a brand new result that suggest some kind of strong non-linearity in the underlying spatial process. Very close interactions turns out to be crucial. However, when considering what is beyond the limit of local markets then distance does not matter so much.

One possible explanation for this result not being found previously is that I am considering here a very large data-set covering all manufacturing activities and with a very detailed geographical level. In fact, as I will show later on, when considering a smaller number of counties/industries then spatial effects are not significant anymore.[13] Furthermore, another element contributing to my positive result is that my estimation strategy addresses the kind of multicollinearity issues among

[12] As one can easily check, the number of observation in estimations is slightly less than the product of the number of industries times the number of counties. This is due to the fact that, in roughly 9 % of the cases, the information on state-level wages is missing because there is no county with positive employment from which an information on wages can be retrieved.

[13] Henderson (2003) find a little scope for spatial spill-overs. However, he considers only 9 three-digit SIC industries and within a subset of 742 of the 3,111 US continental counties.

Table 3 Basic regressions

	Local only (1)	Spatial only (2)	Loc. and Spat. (3)	Non-zero only (4)
IntD	0.1581*		0.2747*	0.1279*
	(0.0023)		(0.0032)	(0.0035)
IntS	0.2305*		0.1921*	0.1029*
	(0.0027)		(0.0030)	(0.0033)
ConsD	−0.0194*		−0.1282*	−0.2544*
	(0.0015)		(0.0029)	(0.0026)
Wage	−0.8127*	−0.4010*	−0.8134*	0.0152
	(0.0132)	(0.0142)	(0.0134)	(0.0129)
W_IntD		0.0696*	0.1618*	0.0628*
		(0.0030)	(0.0040)	(0.0037)
W_IntS		0.2048*	0.0361*	−0.0528*
		(0.0061)	(0.0059)	(0.0057)
W_ConsD		−0.0199*	−0.1282*	−0.2568*
		(0.0015)	(0.0036)	(0.0035)
τ_1		0.4156*	0.5734*	0.5321*
		(0.0376)	(0.0264)	(0.0813)
τ_2		0.4085*	0.5241*	0.5050*
		(0.0313)	(0.0219)	(0.0983)
τ_3		1.0166**	0.4111*	0.4423*
		(0.5063)	(0.0122)	(0.0712)
Estimation method	ML	ML	ML	NLLS
Dummies	No	No	State	State
(Pseudo) R^2	0.2045	0.0197	0.2256	0.2024
Log likelihood	−101,267	−124,800	−99,851	
Prob. LR (F) test	0.0000	0.0000	0.0000	0.0000
Number of industries	49	49	49	49
Number of counties	3,111	3,111	3,111	3,111
Number of (non-missing) obs.	138,419	138,419	138,419	48,195
Number of censored obs.	90,224	90,224	90,224	0

Note: Standard errors in *brackets*. *, ** denote significance at respectively the 1 % and 5 % levels

spatially lagged variables, that are present in Rosenthal and Strange (2003), and that limit the observable extent of spatial interaction.[14]

This picture is further confirmed by estimations in columns (3) and (4), where both spatial and local variables are put together. While maximum likelihood is still used in column (3), I tried to address the problem of high level of censoring in the data in column (4) by using non-zero data only. In particular estimations in column (4) have been obtained with non-linear least squares using observations corresponding to positive values of local industry employment ($Spec_{c,i} \neq 0$).

[14] As mentioned earlier, Rosenthal and Strange (2003) measure proximity effect by introducing many spatial variables; each corresponding to a given distance band. The problem with such an approach is that these variables usually turn out to be very correlated among themselves so biasing the inference.

Table 4 Partial correlation matrix

	IntD	*IntS*	*ConsD*	*wage*	*W_IntD*	*W_IntS*	*W_ConsD*
IntD	1	0.5574	0.3195	0.0163	0.6334	−0.3150	−0.2750
IntS	0.5574	1	0.2913	0.0220	−0.4781	0.5146	−0.2417
ConsD	0.3195	0.2913	1	0.0452	−0.2309	−0.2517	0.6540
Wage	0.0163	0.0220	0.0452	1	−0.1882	0.1739	−0.1493
W_IntD	0.6334	−0.4781	−0.2309	−0.1882	1	0.5722	0.0521
W_IntS	−0.3150	0.5146	−0.2517	0.1739	0.5722	1	0.4234
W_ConsD	−0.2750	−0.2417	0.6540	−0.1493	0.0521	0.4234	1

As one can see, there is no dramatic difference in the two set of estimations, with ML performing better in terms of efficiency. Moreover, estimates of column (3) are quite close to those of columns (1) and (2). As suggested by the partial correlation matrix in Table 4, this is due to the fact that there is no major problem of collinearity among local and spatial variables.

Concerning parameters' signs, all estimations in Table 3 indicate that both demand for intermediate use and input needs are positive with a local (spatial) elasticity varying from 0.10 to 0.30 (0.04–0.20) depending on the specification. Input–output linkages thus seem to be an important element in the location of manufacturing activities. In particular, local flows are those that play a crucial role, suggesting that face-to-face interaction are extremely important not only for service-like activities but also for industrial relationships. By contrast the effect of final demand is, as suggested in a related work by Gaigné et al. (2003), always negative indicating that proximity to large consumers' markets (urban areas) is harmful for manufacturing activities. This finding is in line with the picture I drew from Holmes and Stevens's model, where manufacturing activities are pushed aside from the center by sectors characterized by higher transportation costs. Although all activities would like to be in the center, only those which can afford to pay higher wages and rents can locate close to final demand. On the other hand, the heterogeneity in input–output relations is sufficiently strong for each industry to find an own "ideal place" where it can benefit from the proximity to some strategic sectors.

As for differences in the spatial scope of the three variables considered (*W_IntD*, *W_IntS*, and *W_ConsD*), there does not seem to be any clear pattern in the data. The three corresponding elasticities (τ_1, τ_2, and τ_3) are in fact very close to each other in all cases but column (2). This is consistent with the flows of input–output goods having the same distance cost function that final goods. Indeed, in their analysis of US local trade flows, Hillberry and Hummels (2002) find that the elasticity of trade with respect to distance does not change dramatically if one considers both final and intermediate goods as opposed to final goods only.

Last but not least remains the parameter corresponding to wages. However, before giving any comment, it is better to make here a small digression. At first glance one may simply associate wages to a cost variable and claim that their effect on local specialization should be negative. This is actually the interpretation given by Gaigné et al. (2003). Indeed, in partial equilibrium (i.e., for a given level of local wages), plants have an advantage to locate where labor costs are low. Nevertheless, considering both labor mobility and the fact that profitability has (in general equilibrium) a positive impact on factor prices then one may expect to find a positive relation between local wages and concentration. As plants enjoy the benefits of forming industrial clusters, they raise local labor demand and so wages. Furthermore, as long as labor mobility entails some frictions, then higher wages should also be offered in order to attract newcomers.

What I find in my base estimations is a negative relation between wages and specialization. This suggest that manufacturing as a whole is being drawn to locations with a low labor cost. This picture is indeed coherent with Holmes and Stevens's model where competition for labor and trade costs push manufacturing towards low-wages locations. Nevertheless, this aggregate result may well hide very different sectoral patterns. I will deal with this issue in the next subsection.

5.2 Robustness Checks

Table 5 contains the results of several robustness checks. The first objection that may be raised on my base results is that they do not take into account sectoral heterogeneity. For example, the positive sign of input–output variables may just be due to the fact that, those sectors that are characterized by an high spatial concentration, actually locate where there is a good access to intermediate goods, while other industries simply do not care about this. In order to account for this issue I have re-estimated Eq. 7 in columns (1) of Table 5 by ML while adding sectoral dummies. As one can see, the general picture I have drawn in the previous subsection is re-enforced in this specification. Input–output linkages boost sectoral specialization, while proximity to final demand has a negative impact. Local variables are crucial but, despite their weakness, spatial effects have a large geographical scope.

The only thing that changes by controlling for sectoral heterogeneity is the contribution of wages that becomes positive. This pattern is actually very clear as confirmed by the other robustness regressions in columns (2)–(4). The way I interpret this result is that industries which are more spatially concentrated are actually those which locate the furthest from cities, where wages are low, thus inducing the negative sign of the variable wages in Table 3. However, once controlled for the different concentration trends, then sectoral wages express the value added created by plants' clustering that turns into a positive relation between local wages and specialization. Anyway, it is import to stress that, although I can tell an interesting story about wages, they do not represent a crucial aspect of my analysis for many reasons. The truth is that I introduced them mainly as a control variable. Nevertheless, by neglecting wages qualitative results are virtually unchanged in all specifications I implemented. Furthermore, as I have been forced to use State-level wages, estimates concerning this variables should be taken with some caution.

Table 5 Robustness regressions

	Sectoral heterog.	Natural resources	MSA counties	No int. linkages
	(1)	(2)	(3)	(4)
IntD	0.3275*	0.3357*	0.2374*	0.1061*
	(0.0035)	(0.0032)	(0.0044)	(0.0036)
IntS	0.2025*	0.2017*	0.1410*	0.1392*
	(0.0031)	(0.0026)	(0.0041)	(0.0032)
ConsD	−0.2226*	−0.1861*	−0.0830*	−0.0686*
	(0.0034)	(0.0028)	(0.0029)	(0.0036)
wage	0.2245*	0.2417*	0.2200*	0.3244*
	(0.0231)	(0.0229)	(0.0227)	(0.0333)
W_IntD	0.1916*	0.1867*	0.1764*	0.0863*
	(0.0214)	(0.0239)	(0.0326)	(0.0298)
W_IntS	0.2205*	0.1518*	−0.1027**	0.0534*
	(0.0452)	(0.0456)	(0.0419)	(0.0201)
W_ConsD	−0.1024*	−0.1017*	−0.0561**	−0.0315*
	(0.0230)	(0.0228)	(0.0244)	(0.0147)
τ_1	0.5512*	0.5614*	0.5676*	0.5314*
	(0.0278)	(0.0223)	(0.0486)	(0.0364)
τ_2	0.5547*	0.5837*	0.4912*	0.5671*
	(0.0286)	(0.0297)	(0.0713)	(0.0389)
τ_3	0.5918*	0.5138*	0.4231*	0.5511*
	(0.0313)	(0.0311)	(0.0563)	(0.0401)
Crops		0.0014		
		(0.0026)		
Water		0.0064**		
		(0.0030)		
Sunshine		−0.0126		
		(0.0855)		
Coast/lake border		0.0231*		
		(0.0087)		
Estimation method	ML	ML	ML	ML
Dummies	State, sector	State, sector	MSA, sector	State, sector
Pseudo R^2	0.2670	0.2881	0.2309	0.1583
Log likelihood	93,313	−91,665	−31,505	−108,421
Prob. LR test	0.0000	0.0000	0.0000	0.0000
Number of industries	49	49	49	49
Number of counties	3,111	3,111	820	3,111
Number of (non-missing) obs.	138,419	138,419	38,092	138,419
Number of censored obs.	90,224	90,224	14,909	90,224

Note: Standard errors in *brackets*. *, ** denote significance at respectively the 1 % and 5 % levels

In column (2) I explore a further important issue: missing variables. Clearly, there are other forces that drive the pattern of industrial specialization than agglomeration externalities stemming from input–output flows or final demand. As suggested by Ellison and Glaeser (1999), the most important alternative elements one can think of are factor endowments (nice weather, minerals, water, etc.), and local amenities (ports, road hubs, etc.). As for factor endowments, the fact that I included both agriculture and extraction activities in the computation of input–output flows for manufacturing should partially address the problem. As a further control, I consider three new variables. The first one is $Crops = ln\left(1 + \frac{Crops\,Surface}{Area}\right)$, that gives the (log) of one plus the ratio between the crops surface and the total area of each county. The second one is *Water*, and is constructed in the same way as the other considering the water surface of each county. Finally, the last one is *Sunshine* and gives the average percentage of possible sunshine days. Concerning local amenities, they should already be well captured by State-level dummies. However, I also consider a new dummy *Coast/Lake Border* that indicates whether a county borders on the sea coast or a great lake. As clear from column (2), the introduction of these new controls does not alter results in any significant way, with only *Water* and *Coast/Lake Border* turning out to be significant. However, I do not consider this latter result as evidence against the importance of factor endowments and local amenities. It is in fact more likely that, as suggested before, these forces are already controlled for by input–output flows and State dummies.

The issue of endogeneity is addressed in columns (3) and (4). The first source of endogeneity comes from the simultaneous nature of the location choice. Agglomeration economies enhance plant productivity and profitability, but successful entrepreneurs also seek out productive locations. If overachieving entrepreneurs were disproportionately found in specialized areas, this would cause one to overestimate the relationship between agglomeration economies and local specialization. This issue have been extensively analyzed in Henderson (2003). After experimenting with both 2SLS and GMM, Henderson concludes that controlling for endogeneity through the use of fixed effects is superior. Specifically, drawing on the panel structure of his data, he estimates his local productivity equation including MSA-time specific fixed effects in addition to plant fixed effects. By adding the MSA-time fixed effects the hope is that this will capture the influence of unobserved attributes that might have drawn a given entrepreneur to the area and that might otherwise be correlated with the error term in the estimating equation. Following this idea, I have re-estimated Eq. 7 using counties belonging to MSA only, while including a complete set of MSA dummies. The result is that both local and spatial variables have a lower magnitude as expected. Nevertheless, the qualitative features of the results are virtually the same.

Another source of endogeneity comes from the contemporaneous presence on the right hand side of Eq. 7 of a measure of local concentration of sector *s*. Roughly 22 % of the input–output flows contained in $IntD_{c,i}$ and $IntS_{c,i}$ have in fact sector *c* as both origin and destination. As a solution to this problem, in column (4) of Table 5 I present estimations obtained excluding intra-sectoral flows in the computation of both $IntD_{c,i}$ and $IntS_{c,i}$. As the reader may note, this does not change the sign and significance of both local and spatial demand variables, but reduces their magnitude of about 50 %. Moreover, the pseudo R^2 of the model is also consistently lowered. This suggest that intra-sectoral flows are a key element in the location of an industry and that failing to control for them can overstate the impact of agglomeration externalities.

5.3 Further Results

In Table 6 I basically deal with two further questions. The first one, that is accounted for by estimations in column (1), is whether industries where final demand is a crucial component of total sales exhibit a different behavior with respect to consumers' proximity. Although this question was partially addressed by our treatment of sectoral heterogeneity, the final answer still remains open. In order to draw more precise conclusions, I have thus re-estimated the model by using the sub-sample of the ten industries with the highest ratio between consumers' demand and industry output.[15] As the reader may note, although estimates suggest a more important role for demand variables ($IntD_{c,i}$ and $ConsD_{c,i}$), proximity to consumers still has a negative impact. Moreover, it is interesting to note that when focusing on a sub-sample of industries, like in Henderson (2003) or Rosenthal and Strange (2003), then spatial effects are no longer significant (due to the smaller power of the tests), even when one is using a very detailed spatial disaggregation (like US counties here) and is controlling for censoring. This rather weak effect somehow justifies the fact that previous studies failed to identify the large scope of agglomeration externalities that I find here by using a very big data-set.

The last issue I want to address is the relation between plants' size and agglomeration forces. Columns (2) and (3) shows estimates that refers respectively to the sample of small and big plants. In particular, using the information on the location of plants with at most (more than) 100 employees, I have re-computed the specialization index $Spec_{c,i}$ while using the information on all plants in the construction of right-hand side variables. There are basically two insights that one may draw from the comparison of columns (2) and (3). The first one is that input–output flows (especially at very local level) seem to play a much more important role in the location process of big plants. The second is that small plants can, as suggested by the magnitude of $ConsD_{c,i}$, be found more easily closer to big consumers' markets than big ones. These two results are suggestive of a picture where big manufacturing plants locates far from metropolitan areas in locations where they can maximize their access to strategic inputs, while shipping their goods everywhere. By contrast, local consumers' demand is much more important for small establishments who chose their location not so far from big markets, competing with service activities for both labor and rents. Lafourcade and Mion (2007)

[15] Note that the sampling does not apply to the construction of any variables. In other words, all sectors are still used in order to build both specialization and input–output variables, but only a fraction of the observations (those referring to the 10 industries chosen) are used in the estimation.

Table 6 Further regressions

	Final demand sectors	Small plants location	Big plants location
	(1)	(2)	(3)
IntD	0.4947*	0.2883*	0.9550*
	(0.0067)	(0.0039)	(0.0135)
IntS	0.0597*	0.1588*	0.6312*
	(0.0063)	(0.0035)	(0.0117)
ConsD	−0.1394*	−0.0698*	−0.4588*
	(0.0047)	(0.0034)	(0.0083)
wage	0.0819**	0.0953*	0.3601*
	(0.0328)	(0.0226)	(0.0699)
W_IntD	−0.0150	0.1506*	0.6075*
	(0.0741)	(0.0168)	(0.1049)
W_IntS	0.1101	0.1224*	0.6173*
	(0.0793)	(0.0430)	(0.1324)
W_ConsD	−0.0420	−0.0480	−0.1552*
	(0.0620)	(0.0255)	(0.0371)
τ_1	0.4944**	0.5816*	0.4167*
	(0.2432)	(0.0386)	(0.0464)
τ_2	0.5618	0.4446*	0.4117*
	(0.3441)	(0.0334)	(0.0481)
τ_3	0.5447	0.4847*	0.5926*
	(0.3511)	(0.0368)	(0.0361)
Estimation method	ML	ML	ML
Dummies	State, sector	State, sector	State, sector
Pseudo R^2	0.3405	0.2366	0.3348
Log likelihood	−16,584	−96,357	−46,306
Prob. LR test	0.0000	0.0000	0.0000
Number of industries	10	49	49
Number of counties	3,111	3,111	3,111
Number of (non-missing) obs.	27,046	138,419	138,419
Number of censored obs.	17,168	93,882	123,327

Note: Standard errors in *brackets*. *, ** denote significance at respectively the 1 % and 5 % levels

actually provide some evidence of this kind of pattern for Italian manufacturing data. In particular, they found that big establishments are more influenced by localized externalities, with spatial effects coming essentially from input–output connections. By contrast, small plants are much more widespread and they seem to match the spatial distribution of final demand.

6 Conclusions

In this chapter I explore the role of the demand for intermediates and final goods in the location of US manufacturing sectors. In particular, using data on labor, input–output needs, distances, and wages I regress a measure of industrial concentration at the county level on the major determinants of location in the so-called new economic geography literature: access to input–output flows, access to consumers' demand, and local wages. Particular attention has been devoted to measure the degree of spatial interaction by means of spatially weighted variables that have been estimated with a specific distance decay parameter coming from a polynomial function.

The reference theoretical framework for interpreting data is Holmes and Stevens (2004b) model that I briefly introduce in the chapter. The nicest feature of this framework is that it helps achieve a better understanding of the location of heterogeneous sectors. In particular, it suggests that, despite manufacturing being attracted by centrality, it will locate in peripheral areas because of the higher wages and rents, while service-like activities, for which transport costs are more important, are willing to pay to stay in the core. Moreover, the model also predicts a very different spatial pattern for agglomeration externalities depending on the variable under consideration. In particular, wages analysis should display a large geographical scope, while looking at local specialization should suggest a much smaller spatial gradient.

Results suggest that all variables are extremely relevant at a very localized level. Both input–output demand and supply in a county positively affect the concentration of a an industry in that location. By contrast, consumers' demand has a negative effect and this result is coherent with the model of Holmes and Stevens (2004b), as well as with the findings of Gaigné et al. (2003). As for distance effects, demand and supply coming from other location have a very weak but significant effect, with an elasticity of around 0.5. The relatively stronger importance of local variables is in line with the findings of Rosenthal and Strange (2003), while the surprisingly slowly declining distance effect is a brand new result. This result suggests some kind of strong non-linearity in the underlying spatial process. The breakdown of the agglomeration analysis by establishments size is also a novelty of the chapter. In particular, data suggests that input–output flows (especially at very local level) play a much more important role in the location process of big plants, while small ones can be found more easily close to big consumers' markets.

There are several directions for further research. A first interesting step would be to extend the analysis to services. This exercise would both contribute to a better understanding of local specialization and will allow one to test for the *a priori* positive relation between location of service activities and consumers' demand. However, the great heterogeneity of sectoral data on services, as well as the problem of how well are input–output linkages measured for these activities, would represent the major challenges for such a research. Another topic I do not deal with in my analysis because of a confidential data problems is the spatial distribution of wages at low geographical level and their link with input–output linkages. This issue has partially been addressed by Amiti and Cameron (2007), but further evidence is needed. In particular, Holmes and Stevens' model has quite clear implications about the relative magnitude of agglomeration externalities as measured by specialization and wages. Both analyses lead to misleading results about the "true" scope of such externalities, and the challenge is to find a way to identify what Holmes and Stevens (2004b) define as the bid-wage function of an industry.

References

Amiti M, Cameron L (2007) Economic geography and wages. Rev Econ Stat 89(1):15–29
Combes P-P (2000) Economic structure and local growth: France, 1984–1993. J Urban Econ 47:329–355
Combes P-P, Lafourcade M (2001) Transport costs, geography and regional inequalities: evidence from France. CEPR discussion paper, 2894
D'Costa S, Gibbons S, Overman H, Pelkonen P (2013) Agglomeration and labour markets: the impact of transport investments on labour market outcomes. In: Crescenzi R, Percoco M (eds) Geography, institutions and regional economic performance. Springer, Berlin (forthcoming)
Davis DR, Weinstein DE (1999) Economic geography and regional production structure: an empirical investigation. Eur Econ Rev 43:379–407
Ellison G, Glaeser E (1999) The geographic concentration of industry: does natural advantage explain agglomeration? Am Econ Rev 89:311–316
Fujita M, Krugman P, Venables AJ (1999) The spatial economy: cities, regions and international trade. MIT Press, Cambridge, MA
Gaigné C, Huiban JP, Schmitt B (2003) Labor market and industrial location: evidence from French data. INRA-ENESAD CESAER working paper, 2
Hanson G (2005) Market potential, increasing returns, and geographic concentration. J Int Econ 67(1):1–24
Head K, Mayer T (2004) Market potential and the location of Japanese investment in the European Union. Review of Economics and Statistics
Henderson JV (2003) Marshall's scale economies. J Urban Econ 53:1–28
Hillberry R, Hummels D (2002) Explaining home bias in consumption: the role of intermediate input trade. NBER working paper, 9020
Holmes TJ, Stevens JJ (2004a) Spatial distribution of economic activities in North America. In: Henderson V, Thisse JF (eds) Handbook of regional and urban economics, vol 4. Elsevier-North Holland, Amsterdam
Holmes TJ, Stevens JJ (2004b) Geographic concentration and establishment size: analysis in an alternative economic geography model. J Econ Geogr 4(3):227–250
Krugman P (1991) Increasing returns and economic geography. J Polit Econ 99:483–499
Krugman P, Venables AJ (1995) Globalization and the inequalities of nations. Q J Econ 110:857–880
Lafourcade M, Mion G (2007) Concentration, agglomeration and the size of plants. Reg Sci Urban Econ 37(1):46–68
Mills ES, Hamilton BW (1994) Urban economics, 5th edn. Harper- Collins, New York
Mion G (2004) Spatial externalities and empirical analysis: the case of Italy. J Urban Econ 56(1):97–118
Ottaviano GIP, Tabuchi T, Thisse JF (2002) Agglomeration and trade revisited. Int Econ Rev 43:409–436
Puga D (1999) The rise and fall of regional inequalities. Eur Econ Rev 43:303–334
Redding S, Venables A (2004) Economic geography and international inequality. J Int Econ 62(1):53–82
Rodríguez-Pose A, Comptour F (2013) Do clusters generate greater innovation and growth? An analysis of European regions. In: Crescenzi R, Percoco M (eds) Geography, institutions and regional economic performance. Springer, Berlin (forthcoming)
Rosenthal S, Strange W (2001) The determinants of agglomeration. J Urban Econ 50:191–229
Rosenthal S, Strange W (2003) Geography, industrial organization, and agglomeration. Rev Econ Stat 85:377–393
Rosenthal S, Strange W (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson V, Thisse JF (eds) Handbook of regional and urban economics, vol 4. Elsevier-North Holland, Amsterdam

Agglomeration and Labour Markets: The Impact of Transport Investments on Labour Market Outcomes

Sabine D'Costa, Steve Gibbons, Henry Overman, and Panu Pelkonen

Abstract One of the possible causes of poorer labour market outcomes for workers in peripheral regions is the small size of cities in these regions. Given this possibility, and the difficulty of affecting city size directly, a frequent policy response has been to invest in transport in order to increase access to markets. In this chapter we investigate how local labour markets respond to these potential transport improvements. We use data on individual workers in the UK to assess how area wages respond to better market access and examine whether this variation is due to a changing composition of the labour market or to higher wages for existing workers. Our results indicate that the increase in wages associated with reductions

---

JEL codes: R42; R23

This chapter is based on work for the Northern Way on Strengthening Economic Linkages Between Leeds and Manchester: Feasibility and Implications (full report: http://www.thenorthernway.co.uk/page.asp?id=817). We are grateful to the Northern Way for funding and to a steering group for advice. We thank Lizze Diss at the Department for Transport and Dan Graham at Imperial College London for their help with the ward to ward GTC for driving. We thank Peter Wiener at Steer Davies Gleeve and John Jarvis at Yorkshire Forward for advice on constructing the train counterfactuals.

S. D'Costa
Queen Mary University of London and SERC

S. Gibbons • H. Overman (✉)
LSE and SERC
e-mail: h.g.overman@lse.ac.uk

P. Pelkonen
University of Sussex and SERC

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_13,

in transport times stems from changes in the composition of the workforce and that wage increases for local workers with unchanged characteristics are minimal.

Keywords Transport investment • Agglomeration • Regional labour markets

1 Introduction

There is increasing interest in the role of cities in driving recovery from recession and economic growth more generally. In the UK this partly reflects the fact that, after a long period of relative decline a number of cities have, until recently, experienced improved economic performance (ODPM 2006). At the same time, evidence suggests there may be continued growth in these cities, particularly if the UK economy continues to move from manufacturing to services. Producers of services benefit in a variety of ways when located in cities and research suggests these benefits may be larger than for manufacturers (Rosenthal and Strange 2004). If services do benefit more from cities, a continued shift towards services points towards a future in which economic activity could be concentrated in a (small) number of larger cities.

Amongst UK policy makers this raises a number of questions. Will this growth be concentrated mostly in the South East? If so, is there anything that policy can, or should, do to counteract this? What role might growth in Northern cities play in increasing growth in the wider Northern economy? Which cities in the North might drive this growth and what, if anything, is the appropriate role for policy? This chapter is concerned with the last of these questions.

In this chapter, we consider the role of transport as a way of increasing the size of the local economy. A larger local economy may help firms be more productive. Such agglomeration economies – the beneficial effects of a larger local economy – may arise for a variety of reasons (Duranton and Puga 2004). A large local economy may facilitate sharing of resources (for example of infrastructure such as airports), matching of capacity (for example of workers to firms) or learning (for example transfer of knowledge between firms). Can we say anything about the likely impact of these effects if we improve transport between locations? In this chapter we use labour market data for individual workers to provide a partial answer to this question.

Our research is related to several strands of work that attempt to answer this question. One strand, following Auscher (1989) treats transport infrastructure as public capital and considers the effects using country or regional level aggregate data. Studies that build on this approach are, at best, inconclusive about the impact of transport on productivity (Gramlich 1994; Boarnet 1997). More recently, attention has shifted from the impact of infrastructure as capital to capturing the effects of the transport network. Rice et al. (2006) provide a nice example using data for UK (NUTS 3) regions. More recently, a second strand of research considers the relationship between transport and economic performance by looking at the link between agglomeration (or accessibility) and productivity using micro-data for firms or workers (Melo and Graham 2009; Holl 2010). Melo and Graham (2009) is the closest research to that described here. They estimate the effects of agglomeration economies on wages of workers in the UK finding that a 1 % increase in market potential leads to a 0.1 % increase in wages, an effect halved when taking workers and firms' unobserved heterogeneity into account. These findings are in line with those reported in this chapter although we use more detailed data on actual transport costs in our analysis.

There has been considerable speculation that the small size of cities in peripheral regions may have negative effects on labour market outcomes and that this may help explain their relative under performance. Given that options to directly increase the city size are limited an alternative is to invest in transport to improve access to other markets. This raises two important questions. First, does improving accessibility *cause* changes in labour market outcomes? Second, if so, what kind of changes occur? There is a small literature that addresses the first question (Ahlfeldt and Feddersen 2010). In this chapter we focus on the second question and ask – if observed correlations between accessibility and labour market outcomes capture causal mechanisms, what does this imply about how local labour markets adjust to transport improvements? Following the long running debate on whether policy should focus on people or place (Kain and Persky 1969) we are particularly interested in whether changes in wages are likely to benefit existing residents. To examine these issues we use data on individual workers to see how wages vary with local labour market size. We assess the extent to which these differences arise from changing composition (e.g. large cities have more educated workers) as opposed to higher wages for existing workers. We then use our estimates, coupled with realistic assumptions about policy induced changes in transport costs, to assess the effect of increased integration on labour market outcomes. Specifically, we consider the effects of improving transport links between two of the North's largest cities: Manchester and Leeds. These cities are of interest because, while both have recently experienced strong growth, research finds little evidence of interaction in terms of business connections or commuting (IPEG/CUPS 2008; Lucci and Hildreth 2008) despite the fact that the cities are closely located. More generally, the case study provides important lessons on the magnitudes of possible labour market effects based on specific proposals for a real world transport improvement.

2 Agglomeration and Labour Markets

We are interested in how labour market size affects labour market outcomes, particularly wages. Our starting point is the observation that larger places tend to have higher productivity and wages. Economists refer to the productivity effects associated with increased levels of economic activity as *agglomeration economies*. This chapter focuses on agglomeration economies that arise in production, that is because of the productivity effects of physical proximity. Higher productivity, in turn, tends to lead to higher wages. We refer to this as the effect of better access to economic mass.

The literature emphasises three sources of agglomeration economies: linkages between intermediate and final goods suppliers, labour market interactions, and knowledge spillovers. Input–output linkages occur because firms save transaction costs by locating close to their suppliers and customers. Larger labour markets may allow for a finer division of labour or provide incentives to invest in skills. Finally, knowledge spillovers arise when spatially concentrated firms or workers are more easily able to learn from one another. See Duranton and Puga (2004). In this chapter, we only consider the overall effect of access to economic mass.

As discussed in the introduction, the small size of peripheral economies may have negative implications for labour market outcomes, partly explaining their relative under performance.[1] One argument for improving transport over and above cost savings is therefore that it increases the size of the local labour market. To consider this we use individual data to see how the level and growth of wages are affected by local labour market size. This provides another way of identifying the overall agglomeration effects studied in the research referred to above.

It is increasingly recognized, however, that the *composition* of the labour market may account for a large part of the relationship between wages, productivity and local market size. For example, large cities may attract more educated workers. Because more educated workers also earn more this leads to a positive relationship between city size and wages. When we measure agglomeration economies by looking at how wages change with city size we actually capture the changing *composition* of the labour force. Alternatively, larger cities may make workers more productive whatever their education level. That is, there is a *place-based* effect whereby larger cities pay higher wages. Our research assesses the extent to which the relationship between accessibility (our measure of local labour market size) and wages arises from changing composition as opposed to higher wages for existing workers. We then use our estimates, coupled with realistic assumptions about changes in transport costs, to assess the effect of increased integration on labour market outcomes. This allows us to paint a richer picture of the potential gains, the distribution of effects and the structural changes that might be needed to achieve them.

2.1 Methodology and Data

To assess the magnitude of agglomeration economies we need to see how wages differ with labour market size. We then want to break these overall effects down into those from changing composition versus those from place-based effects. To do this, we need to look at wages for individuals who are otherwise identical but who live in different sized labour markets. Ideally, we would do this by randomly allocating people across places. In reality, fortunately, the UK government does not decide

[1] See, for example, the Manchester Independent Economic Review (2009) which considers this issue.

where people live. People are therefore able to sort across places in non-random ways. If we observed everything about an individual (age, sex, education) that affected their wage, even in the absence of random allocation, we could still identify place effects by comparing wages for people with identical observable characteristics who live in different places. Unfortunately, even with detailed data, we cannot be certain that we observe everything that might affect wages. For example, in our data, we have no information on cognitive abilities or motivation. One way to get round this problem is to follow individuals as they move across places. Providing that ability is fixed over time, if we see the same individual earning more in larger labour markets we may be more confident in attributing this to a place-based effect. Although something may have changed that both affected their earnings potential and their location, in the absence of random allocation, (or a policy change that as good as randomly assigns people) tracking individuals over time is the best we can do to identify place-based effects of changing labour market size.

To do this we need data on where individuals work, their wages and the individual characteristics that might affect wages. We would also like to be able to follow individuals over time, particularly as they move across labour markets. In the UK, such data is available from the Annual Survey of Hours and Earnings (ASHE).[2] ASHE is constructed by the Office of National Statistics (ONS) based on an annual 1 % sample of employees on the PAYE register with workers sampled in multiple years. It provides information on individuals including their work postcodes (which we map to other geographical units using the National Statistics Postcode Directory (NSPD)), earnings including base pay, overtime pay, basic and overtime hours. We use basic hourly wage in our analysis. Individual occupation codes and data on the characteristics of an individual's job (public sector, part time, collective agreement, industry coded by SIC2003) also come from ASHE. ASHE does not provide years of education so we classify workers as belonging to one of four skill groups constructed using a mapping of SOC (Standard Occupational Classification) codes to skill groups.[3]

Information on employment and the industrial composition of areas comes from the Business Structure Database (BSD) which provides an annual snapshot of the Inter-Departmental Business Register (IDBR) accounting for approximately 99 % of economic activity in the UK.[4] For the occupation structures of areas we aggregate the individual data in ASHE. Finally, area proportions of workers belonging to high-skill and intermediate-skill groups are based on LFS (Labour Force Survey) data. More detail on ASHE is provided in Gibbons, Overman and Pelkonen (2010), while information on data cleaning is provided in Northern Way (2009).

We follow existing research by focusing on the relationship between wages and 'access to economic mass' rather than city size. Imposing city boundaries is

---

[2] Previously the New Earnings Survey.

[3] Northern Way (2009) provides more details.

[4] The 99 % coverage was last verified in 2004/05, although there is no reason to think that this is not still the case.

essentially arbitrary whereas measures of access to economic mass treat space as continuous by taking into account access to all other areas discounted by distance or transport cost. We construct two measures of access to economic mass.

The first is based on Generalised Transport Costs (GTC) when driving. Ward to ward generalized costs (driving) were provided by the UK Department for Transport (DfT). They comprise fuel and non-fuel vehicle operating costs and the value of time multiplied by travel time. The data have been averaged for peak and off peak. The exact formulae for these calculations can be found in the DfT's Transport Economics Note (DETR 2001) and more details on our use of GTCs are in Northern Way (2009). The access to economic mass for ward j is calculated by adding up employment in all other wards using inverse-GTC (driving) weighting. That is. $A_{it} = \sum emp_{jt} \times GTC_{ij}^{-1}$ where emp_{jt} is employment in ward j at time t, and GTC_{ij} is the ward-to-ward GTC for driving. To allow employment in ward i to contribute to its own access to economic mass, we set $GTC_{ii} = 0.5 \times GTC_{i1}$ where GTC_{i1} is the minimum ward-to-ward GTC for ward i (i.e. GTC for the "closest" ward). Therefore, a ward is assigned an aggregate of employment in other wards, with employment in more distant places contributing less than employment close by. Each worker is assigned the access to economic mass value for the ward in which their employer is located.[5] Note that this index is identical to the effective density index used by Graham (2006) (although we prefer to refer to it as a measure of *access to economic mass* or *accessibility*).

Our second access to economic mass measure is calculated using train GTC. The train GTC are calculated using data provided to us by the DfT.[6] The final GTC matrix by train is a weighted sum of in-vehicle, wait and walk times (multiplied by the respective time value) and fare matrices. The index is constructed in an identical manner to the index based on driving GTC (although based on ward, not LA data). We refer to the employment accessibility measure based on driving GTC as "Car Accessibility" and that based on train GTC as "Train Accessibility".

2.2 Results

We start by considering a model that captures agglomeration economies, ignoring the distinction between composition and place-based effects. Specifically, we run regressions that explain individual wages as a function of accessibility:

$$\ln(w_{it}) = \alpha_t + \theta \ln(A_{it}) + \varepsilon_{it}$$

where w_{it} is individual i's wage at time t, A_{it} is one, or both, of the accessibility variables, ε_{it} is an error term representing unobservable factors, α_t is a time varying parameter and θ a time invariant parameter (both to be estimated). The alphas

[5] While ASHE contains information on both home and work postcode, NES only provides the latter so we need to base our measure of access to economic mass on work rather than home location.

[6] They stem from Base Year (2004) Rail'Level of Service' skims based on UK Midman rail data.

Table 1 Regressions of wages on accessibility

	[1]	[2]	[3]	[4]	[5]
	Car	Train	Both	Without London	TTWA Employment
ln car accessibility	0.230*		0.084	−0.040	
	(0.092)		(0.122)	(0.035)	
ln train accessibility		0.344***	0.258**	0.217***	
		(0.093)	(0.093)	(0.036)	
ln employment					0.069***
					(0.008)
R^2	0.085	0.086	0.09	0.06	0.085
Observations	1,102,527	1,119,582	1,102,527	884,953	1,119,582

Notes: Dependent variable is log hourly earnings and the explanatory variables are logarithms of car and train accessibility, or log TTWA employment. All estimations are based on panel data for 1998–2007, and include year effects. Standard errors (reported in *brackets*) are clustered at the TTWA level. ***, **, * denote significance at the 1 %, 5 % and 10 % levels respectively

capture the increases in wages over time, while theta captures the impact of accessibility (assumed constant over our relatively short time period). Results are reported in Table 1.

We report results using only Car Accessibility (column 1), only Train Accessibility (column 2) then both together (column 3). When entered separately both are positively and statistically significantly associated with wages. When including both together we find the coefficient on Car Accessibility is positive but insignificant while that on Train Accessibility is both positive and significant. These effects of accessibility remain essentially unchanged if we drop individuals that work in London (column 4). Finally, for comparison we present results based on Travel to Work Area (TTWA) employment rather than accessibility (column 5).

In terms of magnitudes, the coefficient on TTWA employment in column 5 is the easiest to interpret. It tells us that a 10 % increase in TTWA employment is associated with a 0.7 % increase in wages. This is consistent with the existing literature on the effect of city size on productivity which reports the effect of a 10 % increase in city size varying from around 0.2–2 % with most estimates under 1 %. The coefficients on the accessibility measures are harder to interpret because they are calculated using GTC weighting of employment across all areas. Taken at face value, the coefficient of 0.344 for Train Accessibility implies that a 10 % increase in employment in all Local Authorities, or a 10 % reduction in the GTC between all Local Authorities, would increase wages by around 3.4 %. For the moment, we focus on how these coefficients change as we introduce individual characteristics. Later, however, we calculate changes in accessibility consistent with proposed transport interventions which gives a feeling for the magnitude of the wage effects.

As we explained above, the problem with these results is that they do not distinguish between the two different explanations of the positive correlation between accessibility and wage. To separate out these effects, we need to control for the fact that individual characteristics that affect wages may be correlated with accessibility. To do this, we include these individual characteristics in our wage regressions to give:

$$\ln(w_{it}) = \alpha_t + \beta' X_{it} + \theta \ln(A_{it}) + \varepsilon_{it}$$

where X_{it} are individual characteristics, beta is a parameter to be estimated and all other notation is as before. Beta captures the effect of individual characteristics on wages leaving theta to capture the effect of accessibility controlling for composition.

In order to separate composition from place-based effects we want to control for predetermined individual characteristics that are correlated with the accessibility of the places in which they live. These characteristics – e.g. gender – can become correlated with labour market size if individuals with different productivities sort into places of different sizes. Clearly the sex of a worker is not determined by accessibility even if males and females then choose to live in different places so that sex is correlated with accessibility. However, there are some individual characteristics that may at least partly be *determined* by accessibility. If for example good accessibility causes a person to choose a higher paid occupation (which is possible if agglomeration economies cause some occupations to be more prevalent in larger labour markets) then we may want to attribute the resulting effect on wages to accessibility *not* to occupation. Controlling for occupation in our wage regressions will yield estimates of the effect of accessibility that net-out any effects arising from occupational choice.

An additional challenge is that an association between composition and accessibility could arise because better transport connections have evolved between labour markets with more productive workers. This suggests caution because the direction of causality may not run from accessibility to labour market composition, but in the opposite direction. Improving transport would not then be effective in changing composition or raising productivity. Other than controlling for a limited number of other area characteristics we do not address this issue, so our estimates are upward biased and the effects that they imply will never be fully realised by improving transport or otherwise increasing accessibility.[7]

These issues complicate our analysis. In short, we want to control for individual characteristics that can be regarded as predetermined, not determined by accessibility in the place in which a person currently works. Unfortunately, there are some characteristics like occupation, education and industry which are partly predetermined, but may be partly determined by the place in which a person works. If we control for these factors, we control for composition effects arising both through sorting (which we want to eliminate), and through changes in individual characteristics induced by accessibility (which we do not necessarily wish to eliminate).

[7] Results in the academic literature suggest that the issue of reverse causality is likely to be much less important than that of composition. See Melo and Graham (2009) and Combes et al. (2011).

The approach we employ is simply to estimate wage equations using various sets of individual characteristics, whilst recognising that controlling for characteristics that are partly determined by accessibility yields lower bounds to the overall effect of accessibility on wages, whereas failing to control for predetermined characteristics is likely to upward bias our estimates. We start by introducing characteristics that are most likely to be predetermined and then adding in characteristics where we are less certain. Results are reported in Table 2.[8] Column 1 just replicates results from Table 1 where we do not control for any individual characteristics. Column 2 shows what happens when we control for sex, age and age squared which, as argued above, are certainly predetermined. The coefficient on Car Accessibility drops while that on Train Accessibility increases although neither change is statistically significant.

The next individual characteristic that we include is education. Although there is some evidence linking educational outcomes to accessibility, the causal effect (if indeed it is causal) is not large. Given our aggregated skills classification we would argue education as largely predetermined. However, as column 3 makes clear, sorting means that education is quite strongly correlated with accessibility, at least for Train Accessibility. This suggests that higher educated workers get paid higher wages and tend to work in areas with higher accessibility by train. Once we control for education the association between wages and Train Accessibility is considerably weakened.

Next we control for occupation, whether the individual works in the public sector, works part time or is subject to a collective pay agreement. We could think of these characteristics as associated with either the individual or the job. If the latter, it is a little harder to be certain that these characteristics are predetermined. Fortunately this issue is moot as introducing these controls has little effect on the coefficients on the two accessibility variables (column 4). A similar story applies when adding industry controls (column 5).

To summarise, when we control for composition based on the observable characteristics of individuals (and jobs) the effect of accessibility is reduced by between a quarter and a third. So far, however, we have only controlled for the observable characteristics of individuals. Given that we observe individuals over time we can use panel data techniques to control for unobservable characteristics of individuals, such as ability, that might be positively associated with both wages and accessibility. Specifically, we include individual fixed effects to control for time-invariant unobservables. This implies that the effects of accessibility are estimated from individuals that move over time (for individuals that do not move, we cannot be sure whether higher wages are something to do with that individual or with the place in which they work). The specification is:

$$\ln(w_{it}) = \alpha_t + \beta' X_{it} + \theta \ln(A_{it}) + \lambda_i + \varepsilon_{ct}$$

[8] We only report coefficients on the accessibility measures. The full results can be found in an appendix to Northern Way (2009).

Table 2 Regressions of wages on accessibility and other variables

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
ln car accessibility	0.084	0.074	0.071	0.054	0.046	0.069***	0.070***
	(0.122)	(0.118)	(0.080)	(0.066)	(0.058)	(0.016)	(0.021)
ln train accessibility	0.258**	0.277***	0.173**	0.165***	0.170***	0.049***	0.030***
	(0.093)	(0.090)	(0.059)	(0.049)	(0.044)	(0.014)	(0.010)
R^2	0.090	0.218	0.513	0.622	0.638	0.918	0.918
Observations	1,102,527	1,091,551	1,091,551	1,091,551	1,090,528	1,090,528	1,090,528

Notes: Dependent variable is log hourly earnings and the explanatory variables of interest are logarithms of car and train accessibility. All estimations are based on panel data covering years 1998–2007, and include year effects. Column [1] has no controls; [2] adds age, age squared and gender; [3] adds years of education; [4] adds occupational characteristics (1-digit level) and dummies for part-time, public sector and collective wage agreement; [5] adds 1-digit industry controls; [6] adds individual fixed effects; [7] adds area level characteristics as described in the text. Standard errors (reported in *brackets*) are clustered at the travel-to-work area level. ***, **, * denote significance at the 1 %, 5 % and 10 % levels respectively where everything is as defined before, except for the inclusion of individual fixed effects λ_i.

As can be seen from Column 6 the effect on the coefficients on the accessibility measures are considerable. For Train Accessibility the coefficient is decreased by a factor of 3 and now smaller than the coefficient on Car Accessibility (although not significantly so). Car Accessibility is now significant for the first time. Controlling for composition using both the observed and unobserved characteristics of individuals provide the best estimate of the relationship between wages and accessibility. We view these coefficients as the upper bound of the likely effect on individuals who do not change sex, age, education etc. as a result of increasing accessibility.

In the results reported so far, we only allow for place-based effects to be explained by accessibility. It is possible that other area characteristics that are correlated with both accessibility and wages might actually be the source of place-based effects. To consider this we control for a number of additional area based characteristics. Following Wheeler's (2006) work on wage growth for the US these include measures of TTWA industrial and occupational diversity to allow for the possibility that diversity might be more important for wages than size. Industrial diversity of a TTWA j is calculated using a Herfindahl index: $\sum_j (E_{ijt}/E_{it})^2$ where E is employment, j is two-digit industry, i is TTWA and t is year. Occupational diversity is an analogous measure using employment by two-digit occupation instead of SIC. We also include the shares of high and intermediate skills in TTWA working age population (with low skills the omitted category). Finally, we include two digit TTWA industry shares to see if the industrial composition makes any difference.

Column 7 shows what happens when we include these additional area characteristics. The effect of Car Accessibility is essentially unchanged, while that of Train Accessibility falls somewhat further. Detailed results in Northern Way (2009) show that TTWA share of high skills and the share of activity in Other Services are the only two significant area characteristics. These are positively correlated with Train Accessibility which reduces the coefficient on that measure of accessibility. Of course, these results may partly reflect the fact that large places attract lots of skilled workers. Without more evidence on the channels, and given that the coefficients on the access variables do not change too markedly, we prefer to use the results in Column 6 (ignoring other area characteristics) when considering the counterfactuals described below.[9]

Finally, we consider whether effects differ depending on the skill level of workers by running regressions separately for each skill group. Table 3 shows the results for our preferred specification including individual fixed effects. It is interesting to note that the effects of improving both Car and Train Accessibility may be slightly stronger for those with intermediate level skills than higher skills,

[9] Results in Northern Way (2009) show that excluding London does not make that much difference to the Train Accessibility coefficients that are the main focus of our counterfactual analysis below.

Table 3 Regressions of wages on accessibility split by skill group

	Skill group 1	Skill groups 2 and 3	Skill group 4
ln car accessibility	0.054***	0.074***	0.049***
	(0.011)	(0.017)	(0.010)
ln train accessibility	0.003	0.054***	0.019*
	(0.015)	(0.016)	(0.009)
R^2	0.826	0.894	0.863
Observations	46,057	894,873	149,598

Note: Regressions dividing by skill groups. FE reports coefficients from a regression including a full set of individual controls and is equivalent to specification [6] in Table 2. ***, **, * denote significance at the 1 %, 5 % and 10 % levels respectively

while the lower skilled may not benefit at all from increased Train Accessibility. We use the average effects in what follows, ignoring the fact that the effects might differ somewhat across individuals.

2.3 The Labour Market Impacts of Closer Integration

We can now use our results to assess the labour market impact of improving accessibility. To do this we construct a counterfactual Train Accessibility measure based on a 20 min reduction in train travel time between Manchester and Leeds.[10] As already explained, this is an investment that has been the subject of considerable interest from UK agencies concerned with narrowing the gap between the North and South of England. It also provides a natural way of translating the abstract coefficients on our accessibility measures into more concrete estimates of the effects of a real world investment.

We calculate the impact on wages by multiplying the percentage changes in accessibility by the relevant coefficient on Train Accessibility that we reported in Table 2 (repeated in the last row of the table). Results are reported in Table 4. The column marked L-M −20 m gives the percentage change in Train Accessibility produced by the 20 min reduction in journey time. The first column reports the total effects of this change (including any compositional changes). These range from a 2.7 % increase in wages in Wakefield to a 1.06 % increase in Tameside. Column 2 shows what happens as we control for age and sex. The estimate of the percentage wage effect increases slightly because the coefficient on Train Accessibility is slightly higher. Column 3 controls for education which leads to the first big reduction in the estimated size of the effect. Columns 4 and 5 show smaller changes as we first introduce occupation and then industrial controls. Finally column 6 shows the large reduction when we allow for unobservable individual

[10] We also allow for the second round (or knock on) effects on journeys between LAs not directly affected (e.g. Liverpool to Hull) that may see improved journey times as a result of the improved network. Northern Way (2009) provides more details on the construction of counterfactuals.

Table 4 Percentage change in wages for a 20 min reduction in Manchester-Leeds train time

LAD Name	CR	L-M −20 m	[1]	[2]	[3]	[4]	[5]	[6]
Bradford	L	6.59	1.70	1.83	1.14	1.09	1.12	0.32
Calderdale	L	6.05	1.56	1.68	1.05	1.00	1.03	0.30
Craven	L	6.3	1.63	1.75	1.09	1.04	1.07	0.31
Harrogate	L	6.98	1.80	1.93	1.21	1.15	1.19	0.34
Kirklees	L	6	1.55	1.66	1.04	0.99	1.02	0.29
Leeds	L	9.75	2.52	2.70	1.69	1.61	1.66	0.48
Selby	L	6.51	1.68	1.80	1.13	1.07	1.11	0.32
Wakefield	L	10.26	2.65	2.84	1.77	1.69	1.74	0.50
Bolton	M	6.17	1.59	1.71	1.07	1.02	1.05	0.30
Bury	M	6.24	1.61	1.73	1.08	1.03	1.06	0.31
Congleton	M	6.29	1.62	1.74	1.09	1.04	1.07	0.31
High peak	M	5.22	1.35	1.45	0.90	0.86	0.89	0.26
Macclesfield	M	7.84	2.02	2.17	1.36	1.29	1.33	0.38
Manchester	M	10.07	2.60	2.79	1.74	1.66	1.71	0.49
Oldham	M	4.56	1.18	1.26	0.79	0.75	0.78	0.22
Rochdale	M	4.34	1.12	1.20	0.75	0.72	0.74	0.21
Salford	M	4.42	1.14	1.22	0.76	0.73	0.75	0.22
Stockport	M	7.62	1.97	2.11	1.32	1.26	1.30	0.37
Tameside	M	4.12	1.06	1.14	0.71	0.68	0.70	0.20
Trafford	M	6.4	1.65	1.77	1.11	1.06	1.09	0.31
Vale Royal	M	6.21	1.60	1.72	1.07	1.02	1.06	0.30
Warrington	M	6.86	1.77	1.90	1.19	1.13	1.17	0.34
Wigan	M	6.47	1.67	1.79	1.12	1.07	1.10	0.32
Multiply percentage change by			0.258	0.277	0.173	0.165	0.170	0.049

Notes: Table 4 shows percentage change in accessibility for a 20 min reduction in train journey times between Manchester and Leeds (L-M-20 m). Column [1] shows total effects including any compositional changes; [2] controls for age, age squared and gender; [3] controls for years of education; [4] controls for occupational characteristics (1-digit level) and dummies for part-time, public sector and collective wage agreement; [5] controls for 1-digit industry; [6] controls for individual fixed effects. The final row corresponds to the coefficients in columns [1–6] reported in Table 2

characteristics. As a reminder column 6 is our preferred estimate of the effect of increased accessibility controlling for the effects of composition. We see the results range from a high of 0.5 of a percent for Wakefield to a low of 0.2 of a percent for Tameside. As is clear, compositional changes account for the vast majority of the overall effect on wages.[11]

[11] It is also interesting to note that the gains, whether including or excluding compositional effects, vary widely across locations with the (already relatively successful) city centres of Manchester, Wakefield and Leeds gaining most from the reduction in train travel times. This suggests that any growth effect for the region as a whole may come at the expense of widening inequalities within the region. Ferraz, Heddad and Terra (this volume) consider a similar equity-growth trade-off in the very different context of the impact of trade liberalisation in Brazil.

We view this as a fundamental policy message: *if* transport investment has a causal effect on wages that is captured in the correlation between accessibility and wages, then most of the overall wage gains of improving accessibility come from the changing composition of labour markets *not* from improved wages for existing workers (i.e. those that do not change education, occupation, industry or ability in response to increased accessibility). As the composition of the Manchester-Leeds economies shifts towards higher educated, higher ability workers average wages rise by between 1.06 % (Tameside) and 2.65 % (Wakefield). But the gains to existing workers who do not change their characteristics in response to increased integration are considerably smaller. We return to the implications of this below.

2.4 Results: Wage Growth

To reiterate, our results so far suggest that any substantive impact on wage levels from greater integration of labour markets come mostly from changing the composition of individuals and partly from changing the composition of work via effects on industrial structure and occupation. The effects on workers who do not change individual characteristics (education, ability) are quite small. In this sub-section we briefly consider the related question of whether accessibility plays a role in driving individual wage growth rather than levels. That is, we consider the possibility that accessibility is more important for understanding the dynamics of the labour market.

The sample of individuals used to study wage growth is essentially the same as that used for wage levels (some additional trimming eliminates very large growth rates). The dependent variable is annualised percentage wage growth over the period of observation of the individual: $\ln(w_T - w_{t0})/(T - t_0)$ where w_{t0} is the individuals' wage in the first year they are observed and w_T is the wage in the final year. Wage growth is normalized by the number of years $T - t_0$ over which the individual is observed to allow for the fact that we observe individuals for different lengths of time.

We work through the same set of specifications as for wages. Because we are looking at wage growth over a period of years we need to decide which characteristics we measure at the start of the period and which we allow to vary over time. Sex is fixed and we measure age and experience at the start of the period. For the remaining individual and job characteristics we simply take the average over the period for which we observe the individual. We also time-average accessibility and area characteristics for each individual (thus allowing for the fact that individuals may move across TTWAs).

We start by regressing growth in wages on both accessibility measures. Results are reported in column 1 of Table 5. The effects are an order of magnitude smaller than those for wage levels. This is reassuring as large differences in growth rates quickly translate into very large differences in the levels of wages (because of the

Table 5 Regressions of wage growth on accessibility and other variables

	[1]	[2]	[3]	[4]	[5]	[6]
ln car accessibility	−0.018**	−0.003	−0.003	0.006*	0.007**	0.005
	(0.006)	(0.005)	(0.006)	(0.003)	(0.003)	(0.004)
ln train accessibility	0.067**	0.0152**	0.010*	−0.012*	−0.011	−0.001
	(0.008)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)
Observations	248,068	246,125	246,125	246,125	246,125	246,125
R^2	0.00	0.08	0.08	0.10	0.10	0.10

Notes: All models have annualised percentage wage growth over the period of observation of the individual as dependent variables and the explanatory variables of interest are logarithms of car and train accessibility variables. Column [1] has no controls; [2] adds age, age squared and gender; [3] adds years of education; [4] adds occupational characteristics (1-digit level) and dummies for part-time, public sector and collective wage agreement; [5] adds 1-digit industry controls; [6] adds area level characteristics as described in the text. ***, **, * denote significance at the 1 %, 5 % and 10 % levels respectively

"compound interest" nature of wage growth). The meaning of the coefficient of 0.067 on Train Accessibility is that a 10 % improvement in Train Accessibility increases annual wage growth by roughly 0.7% points. We now start to introduce individual characteristics in the same order as for the wage regressions. As before, we only report the coefficients on accessibility. Adding sex, age and age squared (column 2) makes the negative effect on Car Accessibility insignificant and substantially reduces the coefficient on Train Accessibility. Adding education (column 3) has a similar effect. Adding occupational controls (1 digit occupation dummies plus part time, public sector and collective agreement) turns Car Accessibility positive and Train Accessibility negative (column 4). Once we include industry dummies (column 5) we are left with a very small effect of Car Accessibility on wage growth, but no effect from Train Accessibility. When we add in area controls, industrial diversity etc., even the effect of Car Accessibility disappears. Note that the fact that we consider average wage growth over the period means we cannot control for individual unobserved characteristics. These made a large difference for wage levels, but their omission here is of less concern because we do not find particularly strong evidence of an effect of accessibility on wage growth when we control for observed individual and area characteristics.

2.5 Labour Markets and Agglomeration: Conclusions

Our results suggest that closer integration between labour markets may deliver additional benefits in terms of increased area wages. Whether these benefits are actually delivered by transport improvements depends on the extent to which the observed correlation between accessibility and wages is actually capturing the causal impact of transport. We have not done much to address this question in this chapter and the limited literature that does try to assess this causality (see Ahlfeldt and Feddersen 2010 and Gibbons et al. 2010) urges considerable caution in attributing all (or indeed any) of this correlation to the causal effects of transport. Regardless of the causal relationship, the results in this chapter urge further caution for policy makers. While our estimates for a 20 min reduction in train journey times between Manchester and Leeds have wages increasing by between 1.06 % and 2.7 %, nearly all of these wage effects come through the changing *composition* of the workforce (arising through sorting, and/or because people change their characteristics in response to changes in accessibility). The effects for any given individual who does not increase their education or skill levels (the place-based effects) are small at somewhere between 0.20 and 0.50 of a percent. Consistent with this, individual wage growth is faster in places with accessibility, but this effect appears to be driven by the fact that these places tend to have more educated workers. Once we control for this there is essentially no relationship between labour market size and wage growth.

Overall, the findings suggest that the aggregate effects of closer integration may be larger than the individual effects. This aggregated effect relies on structural changes moving the composition of better integrated labour markets towards higher skilled jobs. From a traditional cost-benefit perspective, these effects would *not* be counted as additional for individual projects if, as is likely, they come about because of greater attraction or retention of existing skilled workers. If they occur because existing workers increase their education or skills in response to changing economic opportunities some part of these higher gains may be additional (to the extent that the individual benefits of increasing, say, education, outweigh the costs). Regardless of the mechanism, if increased integration does lead to structural change (and again, we emphasise that this chapter has done little to address the crucial issue of causality) these compositional changes will increase aggregate output in better connected labour markets, and this will be of interest to policy-makers interested in the performance of the better connected places. In our case study, the estimated impact of closer integration between Manchester and Leeds is dependent on induced changes in the composition of the population. It represents an upper bound of the possible effects as we cannot rule out the possibility that some of this effect runs *from* the composition of the labour market *to* lower transport costs (rather than vice versa). We find evidence that the effect on wages for individuals who do not change their personal or job characteristics are small (between 0.2 % and 0.5 %). This modest impact on the wages of workers whose characteristics remain unchanged is likely to be offset or even reversed by induced increases in the cost of living.

References

Ahlfeldt G, Feddersen A (2010) From periphery to core: economic adjustments to high speed rail. Processed, LSE

Aschauer DA (1989) Is public expenditure productive? J Monet Econ 23:177–200

Boarnet MG (1997) Highways and economic productivity: interpreting recent evidence. J Plann Lit 11(4):476–486

Combes P-P, Duranton G, Gobillon L (2011) The identification of agglomeration economies. J Econ Geogr 11(2):253–266

DETR (2001) Transport economics note. DETR, London
Duranton G, Puga D (2004) Micro-foundations of urban agglomeration economies. In: Henderson JV, Thisse JF (eds) Handbook of urban and regional economics, vol 4. Elsevier, Amsterdam, pp 2063–2117
Gibbons S, Lyytikainen T, Overman HG, Sanchis R (2010) Productivity and employment impacts of agglomeration: evidence from transport improvements, SERC/LSE/UKTRC working paper
Gibbons S, Overman HG, Pelkonen P (2010) Wage disparities in Britain: people or place? SERC/LSE discussion paper 0060
Graham D (2006) The localization of service industries in Britain. Working paper, Centre for Transport Studies, Imperial College, London
Gramlich EM (1994) Infrastructure investment: a review essay. J Econ Lit Am Econ Assoc 32 (3):1176–1196
Holl A (2010) Market potential and firm-level productivity in Spain. Processed, FEDEA, Madrid
IPEG/CUPS (2008) The Northern connection, report for the Northern Way. http://www.thenorthernway.co.uk/displaypagedoc.asp?id=516
Kain J, Persky J (1969) Alternatives to the gilded ghetto. Public Interest 14:74–83
Lucci P, Hildreth P (2008) City links. Centre for Cities Report for the Northern Way
Manchester Independent Economic Review (MIER) (2009) The case for agglomeration economies. http://www.manchester-review.org.uk/
Melo PC, Graham DJ (2009) Agglomeration economies and labour productivity: evidence from longitudinal worker data for GB's travel to work areas. SERC discussion paper no 31
Northern Way (2009) Strengthening economic linkages between Leeds and Manchester: feasibility and implications http://www.thenorthernway.co.uk/page.asp?id=817
ODPM (2006) State of the cities report. http://www.communities.gov.uk/publications/regeneration/state4
Rice P, Patacchini E, Venables AJ (2006) Spatial determinants of productivity: analysis for the regions of Great Britain. Reg Sci Urban Econ 36(6):727–752
Rosenthal SS, Strange WC (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson JV, Thisse JF (eds) Handbook of urban and regional economics, vol 4. Elsevier, Amsterdam, pp 2119–2171
Wheeler CH (2006) Cities and the growth of wages among young workers: evidence from the NLSY. J Urban Econ 60:162–184

Firm Capabilities and Cooperation for Innovation: Evidence from the UK Regions

Simona Iammarino, Mariacristina Piva, Marco Vivarelli, and Nick von Tunzelmann

Abstract This chapter focuses on the relationship between firms' technological competencies and capabilities and different forms of cooperation for innovation by combining the analysis of both micro and meso levels, i.e. the level of the firm and of the geographical region. Our findings, based on the Fourth UK Community Innovation Survey (CIS), provide new insights regarding the relationship between cooperative linkages for innovation and firms' technological status. Firstly, the distinction between competencies and capabilities adopted in this chapter seems appropriate for going beyond the rather simplistic dichotomy of 'innovative' versus 'non-innovative' firms commonly used in interpreting CIS data. Secondly, we find that the analysis for the UK as a whole masks stark regional differences in terms of intra- and extra-region collaborative linkages and firms' technological status.

Keywords Cooperation for innovation • Firms' technological competencies and capabilities • UK regions

---

JEL Classification: O30, R12

S. Iammarino
London School of Economics and Political Science, London, UK
e-mail: s.iammarino@lse.ac.uk

M. Piva • M. Vivarelli
Università Cattolica del Sacro Cuore, Piacenza, Italy
e-mail: mariacristina.piva@unicatt.it, marco.vivarelli@unicatt.it

N. von Tunzelmann
Science and Technology Policy Research (SPRU), University of Sussex, Brighton, UK
e-mail: G.N.Von-Tunzelmann@sussex.ac.uk

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_14,

1 Introduction

Following evolutionary views of technical change, technological capabilities can be considered as the outcome of in-house technological competences and of complex interactions among individuals, firms and organisations within a specific socio-economic and institutional framework. Among these interactions, a crucial role is played by the cooperation in R&D and innovative activities, where such a cooperation can involve other firms in the same sector, suppliers, customers, consultants and scientific institutions like universities and public labs.

A shortcoming of most previous studies that have investigated technological capabilities is that these latter are seen at the same time as inputs and outcomes (among others Westphal et al. 1990; Romijn 1999; Wignaraja 2002). In this chapter it is argued that, in order to evaluate firms' technological capabilities, variables related to outcomes such as the introduction of new products and processes are more appropriate. This distinction comes mainly from the differentiation between competences and capabilities introduced by von Tunzelmann and Wang (2003). Whilst competences are understood as enhanced inputs to produce goods and services, capabilities generally involve learning, the accumulation of new knowledge, and the integration of behavioural, social and economic factors, as adapted to specific contexts. Consequently, capabilities are to be taken as the results of adaptive learning processes that are sustained through a variety of external connections and sources for innovation (von Tunzelmann and Wang 2003, 2007), at least partially embedded in the regional environment of the firm. As capabilities directly promote better and/or cheaper goods and services, we can index them by their embodiment in new products and processes (von Tunzelmann 2009a).

This chapter investigates the relationship between firms' technological status – in terms of its competences and capabilities – and different forms of cooperation for innovation by combining the analysis of both micro and meso levels, i.e. the level of the firm and of the geographical region. In particular, the aim is to provide an answer to the following questions: To what extent are the patterns of inter-organisational cooperation for innovation associated with firms' technological status? And is such a relationship influenced by the regional location of firms?

The chapter is structured into five sections. The following section briefly summarises the literature on firm-level technological capabilities, innovation linkages and sources external to the firm, emphasising the relevance of the regional environment as an appropriate dimension to study such relationships. Section 3 provides an overview of the data, methodology and assumptions underlying our empirical analysis. Section 4 discusses the results at the firm level, taking into account both the overall national context and the specific regional environments. Section 5 offers some concluding remarks and implications of the research here carried out.

2 Technological Competences and Capabilities, Firms and Regions

Over the last three decades, following Nelson and Winter' s seminal book (1982), contributions in the area of firm-specific capabilities have proliferated in and around resource-based views, evolutionary economics, the economics and history of technical change, strategic management and, more recently, evolutionary economic geography. The term capabilities has been used variously across different levels of systems from individual to global, to describe a large variety of processes (e.g. 'social capabilities' for growth, see Abramovitz 1986) and a variety of functions. In this chapter, technological capabilities at the micro-level are defined as the knowledge and skills that the firm needs in order to acquire, use, adapt, improve and create technology, interacting with the external environment (e.g. Lall 1992; Malerba 1992; Bell and Pavitt 1993, 1995). The main extensions to traditional static notions of capabilities involve both interactive and dynamic capabilities. Formally, the interactive dynamic capabilities of firms represent the extent to which the change in their productive capabilities influences or is influenced by the change in the capabilities of other external actors – i.e. consumers, clients, suppliers, universities, etc. – in real time or over historical periods (von Tunzelmann and Wang 2007; von Tunzelmann 2009a).

A crucial distinction between competences and capabilities has been introduced by von Tunzelmann and Wang (2003). Competences are understood as stemming from inputs to produce goods and services – in this sense they are pre-set attributes of individuals and firms, with the enhancements typically produced either in-house or by a different organisation. For example, one may think of firm's endowment of adequate skills as the necessary internal competences to obtain value from R&D and innovation investments (see Piva and Vivarelli 2009). By the same token, the recruitment of university graduates may be intended as the necessary internal competences for small and medium enterprises (SMEs) and new technology-based firms (NTBFs) that want to obtain values from external spillovers (see Acs et al. 1994; Audretsch and Vivarelli 1994; Arrighetti and Vivarelli 1999).

Capabilities instead involve both internal and external learning, accumulation of new knowledge on the part of the firm, and the integration of behavioural, social and economic factors into a specific set of outcomes. Consequently, capabilities are to be taken as the results of adaptive learning processes that, in their collective dimension, can be highly localised, giving rise to 'system' capabilities, i.e. referring to a specific spatial and industrial setting such as a regional innovation system (von Tunzelmann 2009b). For instance, an endowment of highly qualified human resources is not a capability per se, but a resource that, through learning, may become a source of technological capabilities for the firm or the system as a whole.

In other words, while technological competences are prerequisites or resources for innovation activity, technological capabilities correspond to knowledge that, through learning and processing, is ready to be incorporated into new products and processes (von Tunzelmann and Wang 2003, 2007). Thus, a firm with technological capabilities does possess competences, though a firm with competences does not necessarily have technological capabilities. For example, a pharmaceutical firm endowed with an adequate R&D lab and performing research on a new vaccine is a firm with technological competences, while a competitor already testing a new vaccine on patients is a firm with technological capabilities.

The study of technological capabilities at the micro-level, as pursued by the data in this chapter, sets the firm at the centre of the analysis (Bell 1984; Bell and Pavitt 1995; Hobday 1995). In addition to interactions and organisational behaviours within the enterprise, the micro-level approach focuses on one-way knowledge and resources flows from external sources of knowledge into the firm. Firms can learn horizontally, that is from spillovers from other producers and competitors, or vertically, by interacting with upstream suppliers and downstream users, as well as from independent research carried out in the regional, national or international science and technology system by universities and research institutes. Such an external learning will need to be absorbed in various ways through the firm's 'absorptive capacity' (see Cohen and Levinthal 1990).

The empirical estimation of such complex internal and external learning processes would require panel micro data with detailed characteristics of firms' internal routines and practises across time, still rarely available in firm-level innovation surveys such as the Community Innovation Survey (CIS) used here. Therefore, the simplifying assumption is that of a sequential logic between the categories of competencies and capabilities: in other words, the know-how accumulated through actual learning and experience in the production of outputs has to lie beyond mere 'potential' – or competencies; hence capabilities are associated with 'realisations' – or the ability of firms to handle change (von Tunzelmann 2009a).

On the other hand, the importance of contextual factors and systemic interactions in the process of generation and diffusion of innovation has long been recognised as a key determinant of the technological and economic performance of firms, countries and regions (see for example Lundvall 1988; von Hippel 1988; Cooke et al. 1997). The significance of the regional dimension of innovation systems has emerged as the logical consequence of the interactive model of innovation (Kline and Rosenberg 1986), which indeed puts the emphasis on the relations with knowledge sources external to the firm. More recently, the 'open innovation' model (Chesbrough 2003; Laursen and Salter 2006) has complemented the innovation system perspective by reinforcing the view that innovative firms draw knowledge from a variety of external sources and linkages, integrating them into their own routines and learning processes, thus achieving more advanced technological capabilities. In this approach, the latest applications of the capabilities framework to regional innovation systems have emphasised that regions can be considered as spatial congregations of actors – i.e. suppliers, producers, consumers, public organisations, etc. – each with its own unique level of competences and capabilities, and all embedded in a specific regional institutional setting (von Tunzelmann 2009b).

It follows that, at the meso-level, regional competences and capabilities cannot be considered merely as the sum of those developed in isolation by individual firms thereby located (Lall 1992; Iammarino et al. 2008; von Tunzelmann 2009b). A region embeds many systemic elements external to the firm that influence its technological capabilities and growth (e.g. Cooke et al. 1997; Howells 1999; Revilla Diez 2000; Cooke 2001; Evangelista et al. 2002; Iammarino 2005; Rodriguez-Pose and Comptour, in this book). While the individual firm can regard some of these as exogenous – for instance, the number of graduate students produced in the area – for the region itself this is not so. Nonetheless, the development of regional capabilities shares many of the features of the micro-level: regional learning is a long, uncertain and costly process, showing high path-dependence and cumulativeness. Thus, idiosyncratic regional structural and institutional features, networks and cooperative agreements emerge and have an influence on firms' R&D and innovative competences and capabilities, eventually resulting into a specific regional innovation pattern (Kleinknecht and Poot 1992; Cooke et al. 1997; Hassink 1997).

In spite of the rather copious empirical literature on cooperation linkages and innovativeness at the firm level (e.g. Cassiman and Veugelers 2002; Vanhaverbeke et al. 2002; Criscuolo and Haskel 2003; Belderbos et al. 2004a; Laursen and Salter 2004, 2006; Faems et al. 2005), surprisingly much less evidence has emerged on the relationship between different forms of collaborative innovative linkages and firms' technological status taking into account the environment of the firm, i.e. its regional location (for an exception, see Simonen and McCann 2008). In this chapter, the regional dimension of the relationship between cooperation for innovation and firms' technological status will be investigated in two ways. On the one hand, we will be able to distinguish between cooperation with local versus non-local partners; on the other hand, the aggregate empirical analysis for the UK will be split into macro-regions, in search for the possible territorial peculiarities that the literature suggests as very likely.

3 Data and Methodology

This paper uses data from the UK Innovation Survey 2005 (as part of the fourth iteration of the wider Community Innovation Survey – CIS4 – covering EU countries), which refer to the period 2002–2004. The survey sampled over twenty-eight thousand UK enterprises with 10 or more employees, had a wide sectoral coverage including both manufacturing and service sectors, and was stratified by Government Office Region in England, along with Scotland, Wales and Northern Ireland. The final representative sample consists of 16,445 firms.

Following the conceptualisation discussed above, technological capabilities at the firm level are signalled by the introduction of a product and/or process innovation. In other words, to identify firms with technological capabilities in the period of reference of the CIS4 we use the strict (output-oriented) definition of innovators, as in most of the previous literature on the CIS.[1] Such a definition (based on questions 5 and 9 of the UK questionnaire) applies if, during the period 2002–2004, the enterprise introduced a new or significantly improved product (either a good or a service) and/or new or significantly improved processes for the production or supply of its products. In such a case the respondent is here classified as a firm with technological capabilities.[2] If instead the enterprise has invested in innovative inputs (on the basis of question 13),[3] but without achieving any innovative output (new product or new process) in the relevant period, it is classified as a firm with technological competences. Finally, if the enterprise has declared neither innovative output nor investment in innovative inputs, it is classified as a technologically inactive firm in the period analysed by the UK CIS4.

Thus, our dependent variable – the firm's technological status – is a categorical ordered variable which assumes the following values: 0 in the case of a *technologically inactive firm*; 1 in the case of a *firm with technological competences*; 2 in the case of a *firm with technological capabilities*. Table 1 displays the regional distribution of the three categories of firms. It is interesting to note that the shares of firms belonging to each group are rather equally distributed across regions.

The patterns of cooperation for innovation – our main interest in analysing the factors influencing the technological status of firms – are based on question 18 of the UK questionnaire, which, in line with the Eurostat standardised questionnaire, is devoted to the types of cooperation partners used by the respondent firms, and their location. In fact, cooperation in R&D and innovative activities may imply a variety of different partners ranging from firms within the same corporate group, customers, suppliers, competitors and institutional partners such as universities and public labs (e.g. Fritsch and Lukas 2001; Miotti and Sachwald 2003; Belderbos

[1] As illustrated by D'Este et al. (2008), there are several reasons why the use of such a definition is appropriate. First, it helps towards separating invention from innovation by requiring new products and processes to be of economic value, as shown by the commercialisation requirement (i.e. introduction to market). Second, it is consistent with the standard definition of innovation provided by the Oslo Manual (OECD 2005). Finally, it helps separating the firm's efforts in innovative activities (as measured, for instance, by its investment in R&D-related activities) from the outputs of those activities (as reflected by the market introduction of new products): thus, it is congruent with the distinction between competences and capabilities adopted here.

[2] Strictly speaking and as already implied, the product or process innovation thus detected does not amount to the relevant capability – it is the accumulated ability to 'know' (learn) how to effect such an innovation that is the 'capability' in the proper sense. For obvious reasons we do not impose this distinction here. As stated above, the observed product/process innovations are direct indexes of these capabilities.

[3] Question 13 in the UK CIS questionnaire asked whether, in the reference period, the firm engaged in any of the following seven innovation activities: (1) intramural R&D; (2) acquisition of R&D; (3) acquisition of machinery, equipment and software to produce new or significantly improved products; (4) acquisition of external knowledge (e.g. licensing of patents); (5) training of personnel for the development or introduction of innovations; (6) expenditure on design functions for the development of new or improved products or processes; and (7) expenditures on activities for the market preparation and introduction of new or significantly improved products (including market research and launch advertising).

Table 1 Regional distribution of the categorical dependent variable on firms' technological status – number of firms and relative percentages (in brackets)

	UK	Northern England	Midlands	Eastern England	London	Southern England	Wales	Scotland	Northern Ireland
Technologically inactive firm	5,308	1,214	891	461	540	910	348	425	519
(*value 0*)	(35.03)	(34.82)	(34.67)	(35.60)	(36.58)	(32.32)	(34.39)	(36.51)	(38.87)
Firm with technological competences	4,105	983	675	326	357	796	293	335	340
(*value 1*)	(27.09)	(28.20)	(26.26)	(25.17)	(24.19)	(28.28)	(28.95)	(28.78)	(25.47)
Firm with technological capabilities	5,740	1,289	1,004	508	579	1,109	371	404	476
(*value 2*)	(37.88)	(36.98)	(39.07)	(39.23)	(39.23)	(39.40)	(36.66)	(34.71)	(35.66)
Total	**15,153**	**3,486**	**2,570**	**1,295**	**1,476**	**2,815**	**1,012**	**1,164**	**1,335**
	(100)	**(100)**	**(100)**	**(100)**	**(100)**	**(100)**	**(100)**	**(100)**	**(100)**

et al. 2004b; Piga and Vivarelli 2004; Laursen and Salter 2006). Seven types of partner are listed in the CIS questionnaire: (A) other enterprises within the firm's group; (B) suppliers of equipment, materials, services or software; (C) clients or customers; (D) competitors or other enterprises in the firm's industry; (E) consultants, commercial labs, or private R&D institutes; (F) universities or other higher education institutions; (G) government or public research institutes.[4]

Dummies were created for each of the *cooperation partners* reported in question 18, aggregating universities/other higher education institutions and government/public research institutes into one category. Two location levels for each of the six partners were taken into account: local/regional, that is within approximately 100 miles of the surveyed enterprise (as defined in the CIS questionnaire); extra-regional (i.e. national/international). The total number of collaboration dummies is thus 12.

In order to have reasonably homogeneous macro-regions, the eight UK *regions* considered are defined following partial aggregations of NUTS 1 regions: Northern England (North East, North West, Yorkshire and the Humber), Midlands (East Midlands, West Midlands), Eastern England, London, Southern England (South East, South West), Wales, Scotland, and Northern Ireland. The size of our final sample is 15,153 firms, owing to the presence of missing values in the patterns of collaboration. We controlled for the geographical representativeness of our final sample, which turned out to be not statistically different from that of the original sample.

While the relationship between the different forms of cooperation and the technological status of the firm is at the core of our investigation, other firm-level factors must be considered. In particular, our control variables are the following:

- *Size:* firm size in terms of employment (continuous variable).
  The Schumpeterian notion that large firms are more likely both to undertake and to succeed in innovative activities has constituted a constant theme in the literature (Schumpeter 1943). Such a notion has been initially challenged from a theoretical point of view (Arrow 1962), and then proposed again in terms of scale and scope economies in R&D investments (Cohen and Klepper 1996). In the last few decades mixed empirical evidence has been found to support the Schumpeterian hypothesis (e.g. Cohen and Levin 1989; Kleinknecht and Reijnen 1991; Audretsch 1995; Breschi et al. 2000).
- *Group:* whether the firm is part of an enterprise group (dummy).
  Various studies have recognised that the group form of organisation tends to play an important role in promoting and supporting innovation (see, for instance, Filatotchev et al. 2003; Piga and Vivarelli 2004).

[4] It is important to stress that, from a theoretical perspective, our focus is on the influence of cooperation on the firm technological status, i.e. technological capabilities versus competences versus inactivity, and not on cooperation as a determinant of innovation. Moreover, as from the CIS questionnaire, the cooperation is in fact intended to be "cooperation for innovation activities" on the basis of the 'conceptualization' of the CIS and the Oslo Manual that look at innovation as an interactive process.

- *Internationalisation:* the extent of internationalisation of the markets served by the firm, in terms of whether the firm sells products/services outside the national market (dummy).
  This is based on the premise that global competition can spur innovation, competences and capabilities, while technologically inactive firms are doomed to be excluded from the international arena (e.g. Archibugi and Iammarino 1999; Narula and Zanfei 2003).
- *Start-up:* whether the firm was established after 1st January 2000 (dummy).
  The debate on the so-called New Technology Based Firms (NTBFs) points out how – at least in some sectors – young companies may be at the core of the innovation process (see, for instance, Storey and Tether 1998; Colombo et al. 2004; Colombo and Grilli 2005).
- *Human capital:* firm-specific skills in terms of proportion of employees educated to degree level or above (continuous variable).
  Human capital is seen as complementary to innovation, constituting per se a competence of the firm, and generating a super-additive effect in terms of both innovative and economic performance (e.g. Acemoglu 1998; Machin and van Reenen 1998; Piva and Vivarelli 2004; Piva et al. 2005).

The analysis also includes regional dummies in the aggregated model, and sectoral dummies in all specifications.[5] The list of all variables used in the analysis and some descriptive statistics are reported in Appendix 1.

In accordance with the nature of the dependent variables, ordered logistic regressions were run.[6] The following specification has been tested for the country as a whole:

$$\begin{aligned}\mathit{Technological\ status} = {} & \alpha + \beta_1 \log(\mathit{employment}) + \beta_2(\mathit{group}) \\ & + \beta_3(\mathit{internationalisation}) + \beta_4(\mathit{start-up}) \\ & + \beta_5(\mathit{human\_capital}) + \beta_6(\mathit{cooperation\_dummies}) \\ & + \beta_7(\mathit{sectoral\_dummies}) + \beta_8(\mathit{regional\_dummies}) + \varepsilon\end{aligned}$$

[5] We followed the clustering criteria used by the (then) Department for Innovation, University and Skills (DIUS), now Department for Business, Innovation, and Skills (BIS): Primary sector; Engineering-based manufacturing; Other manufacturing; Construction; Retail & distribution; Knowledge-intensive services; Other services.

[6] As the dependent variable is an ordered one, we opted for the ordered logistic model. However, multinomial logistic regressions were also run with the category of technologically inactive firms as the reference (category 0). The results from the multinomial, and the estimated predicted probabilities of both the multinomial and the ordered logistic models, supported our choice of the latter, as the probability distribution between the two estimation methods is not substantially different. Furthermore, a Brant test to verify the parallel regression assumption (also called the proportional odds assumption) was performed after the ordered model and – where feasible in the regional models – it provided evidence that the parallel regression assumption has not been violated.

To explore whether and to what extent the relationship between technological competences and capabilities and different cooperation patterns is region-specific, the same model was also estimated for each of the eight UK regions.

4 Competence, Capabilities and Cooperation for Innovation: Results

This section reports the results on the differences across firms and regions in terms of competences and capabilities for innovation, once we explicitly consider a number of factors, and particularly collaborative linkages, which may influence the dependent variable.

First of all, it is important to note that our results are especially driven by the category of firms with technological capabilities. In fact, the cut-off as between technologically inactive firms and those with competences (enhanced inputs but lacking technological outputs) is generally not statistically significant in the ordered logistic regressions for the country as a whole and for any region in the regional estimations.[7] On the other hand, the few peculiar features that characterise the group of firms with technological competences with respect to the other two firm groups support the conjecture that these firms represent somehow an intermediate innovative behaviour (see also D'Este et al. 2008). This sits comfortably with our choice of distinguishing the three categories of firms according to their technological status.

Table 2 shows the results for the aggregate model – i.e. for the UK as a whole. In line with the theoretical expectations discussed above, all the variables related to firm characteristics are highly significant at 1 % level (with the exception of start-up, with a level of significance of 5 %), indicating a positive impact on the likelihood of firms to be classified as firms with technological capabilities.[8] As far as the independent variables concerning cooperation partners are concerned, linkages with other enterprises within the group, suppliers, clients, and public research and higher education institutions are all positive and highly significant at both local and extra-regional level.

Consistently with the previous literature, both the dummy for belonging to a group and the two dummies indicating cooperation with firms within the same corporate group turn out to be statistically significant at 1 % level, pointing to the strategic role of corporate relationships in enhancing the technological status of the

[7] Equivalent results hold in the unreported multinomial logit regressions, where the coefficients for firms with technological competences (category 1) are in general smaller and/or with lower significance levels than those for firms with technological capabilities (category 2).

[8] In order to verify the relevance of the human capital regressor in the ordered logit estimate, we run the model also excluding the variable: the results are confirmed, indicating that endogeneity issues with respect to this variable are not serious.

Table 2 Determinants of firms' technological status, UK with regional dummies. Ordered logistic regression. Categorical ordered dependent variable: 0 = technologically inactive firm; 1 = firm with technological competences; 2 = firm with technological capabilities

(1)		(2)	
Ln(Employment)	0.12*** (9.60)	*Sectoral dummies*	Yes***
Group	0.23*** (6.27)	Primary sector	−1.06*** (7.67)
Internationalisation	0.56*** (14.59)	Engineering-based manuf	–
Start-up	0.10** (2.26)	Other manufacturing	−0.01 (0.20)
Human capital	1.41*** (16.99)	Construction	−0.97*** (13.87)
		Retail and distribution	−0.82*** (13.54)
C **Cooperation partners for innovation**		Know.-intensive services	−0.28*** (4.37)
A: other enterp. within group LOCAL	0.40*** (2.90)	Other services	−0.72*** (12.66)
A: other enterp. within group NON LOCAL	0.32*** (2.62)		
B: suppliers LOCAL	0.44*** (3.43)	*Regional dummies*	Yes***
B: suppliers NON LOCAL	1.16*** (10.81)	North England	0.11 (1.63)
C: clients LOCAL	0.44*** (3.38)	Midlands	0.14* (1.95)
C: clients NON LOCAL	0.69*** (6.16)	Eastern England	0.09 (1.12)
D: competitors LOCAL	−0.27* (1.69)	London	−0.16** (2.11)
D: competitors NON LOCAL	0.05 (0.38)	South England	0.20*** (2.88)
E: consultants LOCAL	0.14 (0.92)	Wales	0.13 (1.54)
E: consultants NON LOCAL	−0.02 (0.12)	Scotland	–
F + G: universities&pub.res. LOCAL	0.46*** (3.47)	Northern Ireland	0.14* (1.80)
F + G: universities&pub.res. NON LOCAL	0.34** (2.43)		

(continued)

Table 2 (continued)

(1)	(2)
LR χ^2(d.f.)	χ^2(30) 3,854***
Pseudo R^2	0.12
Observations	15,153

Notes

In brackets: z- statistics; * = 10 % significant; ** = 5 % significant; *** = 1 % significant

In column (1) the control and the cooperation regressors are reported; in column (2) the seven sectoral dummies (DIUS sectoral classification – Engineering-based manufacturing is the reference case) as well as the eight regional dummies (NUTS1 regional aggregations – Scotland is the reference case) are reported. See Appendix 1 for definitions of variables

Yes*** for sectoral and regional dummies reporting that they are, respectively, jointly significant at 1 % level

individual firms involved. Interestingly enough, the enforcement role of cooperation appears more obvious for firms in the same corporate group and also located at a short distance.

Geographical distance instead does not seem to play a dominant role in the vertical cooperative links, where actually the coefficients related to non local suppliers and customers give a result that is higher in magnitude than those related to local partners. However, the vertical relationships themselves come out as highly significant in all four cases.

By contrast, horizontal cooperative links with competitors and consultants do not bear a significant impact on enhancing the probability that a firm achieves either technological competences or capabilities. In other words, while innovation seems to be reinforced by collaborations along the value or supply chains, once we turn our attention to the horizontal dimension rivalry seems to dominate. It is interesting to note, though, that cooperation with extra-regional consultants turns out to have a much stronger effect in increasing the likelihood of being a firm with technological competences. This can be explained as that firms at the stage of investing in innovation inputs tend to rely much more on the advice provided by external consultants in order to achieve commercial success.

Cooperation with universities and public research institutes turns out to be significant and positively affected by the close proximity of the involved partners. This is not surprising, due to the localised nature of labour markets and the fact that university-industry innovative linkages often occurs through the hiring of graduates by local firms.

In line with expectations, the high significance of the sectoral dummies indicates that the industrial structure matters in affecting the technological status of the firm: not surprisingly, the impact is greatest in engineering-based manufacturing, which is the chosen reference category. The regional dummies in the aggregate regression are also jointly significant at 1 % level, but with rather pronounced variations in terms of sign, magnitude and significance of the coefficients. This result provides

Table 3 Determinants of firms' technological status, regions of UK. Ordered logistic regression: regional analysis. Categorical ordered dependent variable: 0 = technologically inactive firm; 1 = firm with technological competences; 2 = firm with technological capabilities

	(1) Northern England	(2) Midlands	(3) Eastern England	(4) London	(5) Southern England	(6) Wales	(7) Scotland	(8) Northern Ireland
Ln(Employment)	0.12^{***}	0.15^{***}	0.16^{***}	0.10^{***}	0.11^{***}	0.08	0.07	0.14^{**}
	(4.62)	(4.82)	(3.55)	(2.98)	(4.27)	(1.34)	(1.48)	(2.53)
Group	0.22^{***}	0.20^{**}	0.26^{**}	0.45^{***}	0.13	0.21	0.27^{**}	0.21
	(2.01)	(1.56)	(2.79)	(2.18)	(1.99)	(4.05)	(1.52)	(1.37)
Internationalisation	0.61^{***}	0.56^{***}	0.66^{***}	0.32^{**}	0.61^{***}	0.88^{***}	0.75^{***}	0.39^{***}
	(7.34)	(5.84)	(4.90)	(2.77)	(6.61)	(5.43)	(5.35)	(3.23)
Start-up	0.26^{***}	0.04	0.08	0.26^{*}	−0.07	0.03	0.17	0.07
	(2.88)	(0.39)	(0.49)	(1.73)	(0.62)	(0.16)	(1.00)	(0.44)
Human capital	1.89^{***}	1.58^{***}	1.76^{***}	1.06^{***}	1.41^{***}	0.82^{**}	0.93^{***}	1.80^{***}
	(9.26)	(7.02)	(5.62)	(5.79)	(7.26)	(2.36)	(3.26)	(5.44)
Cooperation partners for innovation								
A: other enterp. LOCAL	0.22	0.13	0.43	0.32	0.30	0.89^{*}	1.35^{***}	1.03
	(0.73)	(0.39)	(0.87)	(0.75)	(0.89)	(1.73)	(2.92)	(1.60)
A: other enterp. NON LOCAL	0.41^{*}	0.55	−0.20	0.03	1.11^{***}	−0.15	−0.24	0.82
	(1.67)	(1.47)	(0.46)	(0.11)	(3.61)	(0.28)	(0.49)	(0.88)
B: suppliers LOCAL	0.67^{**}	0.63^{*}	0.65	0.78^{**}	−0.44	0.91^{*}	0.66	0.31
	(2.45)	(1.77)	(1.51)	(2.05)	(1.47)	(1.68)	(1.53)	(0.62)
B: suppliers NON LOCAL	0.83^{***}	1.51^{***}	1.52^{***}	1.40^{***}	1.23^{***}	1.01^{**}	0.79^{**}	1.71^{***}
	(3.76)	(4.36)	(3.91)	(4.59)	(5.19)	(2.25)	(1.99)	(3.27)
C: clients LOCAL	−0.10	1.00^{***}	0.99^{**}	0.09	0.82^{***}	0.37	0.10	0.90^{*}
	(0.36)	(2.93)	(2.14)	(0.21)	(2.58)	(0.73)	(0.22)	(1.72)
C: clients NON LOCAL	0.56^{**}	1.37^{***}	0.81^{**}	0.44	0.27	0.16	0.96^{**}	2.24^{***}
	(2.49)	(4.36)	(2.07)	(1.36)	(1.00)	(0.33)	(2.41)	(2.96)
D: competitors LOCAL	−0.02	−0.39	−0.16	−0.39	−0.60	−0.58	−0.40	−0.70
	(0.07)	(0.98)	(0.26)	(0.83)	(1.36)	(1.01)	(0.75)	(0.11)
D: competitors NON LOCAL	0.20	-0.99^{***}	−0.71	0.70^{*}	1.03^{***}	0.48	−0.56	−0.69
	(0.76)	(2.88)	(1.60)	(1.69)	(3.30)	(0.99)	(1.16)	(0.96)

(continued)

Table 3 (continued)

	(1) Northern England	(2) Midlands	(3) Eastern England	(4) London	(5) Southern England	(6) Wales	(7) Scotland	(8) Northern Ireland
E: consultants LOCAL	0.65^{**}	−0.08	−0.45	−0.22	0.73^{*}	0.49	−0.08	−0.92
	(1.97)	(0.20)	(0.77)	(0.52)	(1.78)	(0.76)	(0.17)	(1.38)
E: consultants NON LOCAL	0.61^{**}	-0.75^{**}	0.20	−0.29	−0.24	0.38	0.36	1.02
	(2.04)	(2.06)	(0.39)	(0.68)	(0.74)	(0.70)	(0.66)	(1.04)
F + G: universities&pub.res. LOCAL	0.81^{***}	0.04	0.62	0.10	0.47	−0.03	0.56	0.54
	(3.17)	(0.11)	(1.16)	(0.22)	(1.29)	(0.08)	(1.19)	(1.03)
F + G: universities&pub.res. NON LOCAL	0.08	0.91^{**}	0.47	0.24	0.26	0.66	0.50	-1.65^{*}
	(0.30)	(2.50)	(1.04)	(0.59)	(0.75)	(1.10)	(1.04)	(1.77)
Sectoral dummies	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Primary sector	-1.13^{***}	-0.84^{**}	−0.73	-1.84^{***}	-0.70^{*}	−0.97	-1.23^{***}	-1.29^{***}
	(3.20)	(2.10)	(1.46)	(3.76)	(1.89)	(1.60)	(3.91)	(3.17)
Engineering-based manuf.								
Other manufacturing	1.18	−0.02	0.06	-0.70^{**}	0.08	0.02	−0.22	−0.19
	(1.57)	(0.18)	(0.30)	(2.40)	(0.55)	(0.08)	(0.99)	(0.90)
Construction	-0.66^{***}	-0.98^{***}	-1.16^{***}	-1.60^{***}	-1.06^{***}	-1.33^{***}	-0.87^{***}	-0.97^{***}
	(4.67)	(5.95)	(4.49)	(4.99)	(6.21)	(5.02)	(3.52)	(4.22)
Retail and distribution	-0.65^{***}	-0.82^{**}	-0.71^{***}	-1.40^{***}	-0.73^{***}	-1.02^{***}	-0.74^{***}	-1.07^{***}
	(5.23)	(5.79)	(3.45)	(4.92)	(5.09)	(4.43)	(3.19)	(5.18)
Knowledge-intensive services	−0.19	-0.32^{**}	-0.40^{*}	-1.04^{***}	-0.24^{*}	−0.14	−0.20	0.16
	(1.44)	(2.17)	(1.80)	(3.80)	(1.65)	(0.58)	(0.87)	(0.61)
Other services	-0.48^{***}	-0.81^{***}	-0.89^{***}	-1.32^{***}	-0.71^{**}	-0.80^{***}	-0.74^{***}	-0.74^{***}
	(4.22)	(6.14)	(4.45)	(4.91)	(5.23)	(3.76)	(3.52)	(3.45)
LR χ^2(d.f.)	χ^2(23) 891***	(23) 700***	(23) 418***	(23) 323***	(23) 801***	(23) 298***	(23) 293***	(23) 341***
Pseudo R^2	0.12	0.13	0.15	0.10	0.13	0.14	0.12	0.12
Observations	3,486	2,570	1,295	1,476	2,815	1,012	1,164	1,335

Notes
In brackets: z- statistics; * = 10 % significant; ** = 5 % significant; *** = 1 % significant
Seven sectoral dummies are included (DIUS sectoral classification)
Yes*** for sectoral dummies reporting that they are jointly significant at 1 % level

further justification to our choice to investigate the regional location, that may have an important role in moulding technological firms' competences and capabilities.

Table 3 provides the estimates for the individual UK regions. As we can see, the cooperation variables show remarkable differences in influencing the technological status of firms, according to the specific regional context.

London for instance, emerges as a peculiar case in terms of the relationship between cooperation for innovation and firms' technological status. The only type of cooperative linkage that turns out to be significant is that with suppliers at both local and, even more, extra-regional level. This would suggest that a firm's location in the metropolitan area does not tend to entail a strong relationship between collaborative linkages and its technological capabilities (see Schienstock 2009, for similar results on firms' capabilities in city-regions). Indeed, the London metropolitan area has to be understood more as an "accumulation node" of global economic and financial transactions, rather than a self-contained regional system. Therefore London's boundaries per se do not set up significant interactions for local firms in terms of either markets or social organisations (Budd 2006), but rather comprise an array of control and management nodes for global transactions and businesses (Newman and Thornley 2005). In this context, London firms' innovative strengths are driven mostly by high levels of variables such as size, corporate group, foreign sales and, especially, human capital.

Considering Southern England, the impact of collaborative relations on firms' capabilities is much more evident than for London, though the strongest of these links are again mostly non local. The only notable exception is intra-regional collaborations with clients, which strongly increase the probability of the firm being classified as having technological capabilities. The peculiar strongly positive effect of horizontal collaboration with extra-regional competitors may perhaps be explained by an 'M4 corridor effect' that could reflect national and international strategic technological alliances among large and/or multinational enterprises whose headquarters are located elsewhere.

In the Eastern England region, the likelihood that firms display technological capabilities is positively influenced by collaborations with extra-regional suppliers and with both local and non-local clients. This latter effect might be justified by the presence of the 'Cambridge cluster', whose successful innovative performance is however counterbalanced by some lagging behind of the rest of the region (Gray et al. 2006).

As far as vertical cooperation is concerned, the results for the Midlands are similar to those for Eastern England, though with more significant linkages with clients and an additional positive impact of local suppliers on firms' technological capabilities (significant at 10 % level). A peculiar feature of the Midlands is the positive effect on firms' capabilities of linkages with universities and public research located outside the region. Indeed, in recent decades the Midlands – characterised until the 1970s and '1980s as the Fordist heartland of the country, particularly for automotive and metal manufacturing, and by coal-based industry – have gone through a period of post-industrial economic restructuring which, more recently, has sparked local innovation potential and connectivity among firms

(AWM 2004; Green and Berkeley 2006; Hardill et al. 2006). This path of evolution of the regional industrial structure may also underlie the negative and significant effect on firms' competences and capabilities of linkages with competitors and consultants external to the region, suggesting that the firms located in the Midlands may be facing difficulties in the international technological race.

Northern England turns out to be the UK region where the association between the different independent variables and firm technological status is the most striking. For example, the start-up variable is strongly significant and positively affecting the probability of firms having technological capabilities, driving the effect at the country level at large. This might be interpreted in terms of a process of gradual replacement of declining and mature industries and shifts of the regional industrial structure towards more advanced manufacturing and service sectors. This seems to be further supported by the positive and significant sign of cooperative linkages with both regional and extra-regional private research and consultants, which again determines the result at the national level. Remarkably, the North is also the only region where the probability of being a firm with technological capabilities is strongly increased by collaborations with local public research institutes and universities, once more driving the result for the UK as a whole. Such a result should be interpreted in the light of the 'Northern Way' strategy implemented for the three Northern English regions (North East, North West and Yorkshire and Humber) since 2000. The 'Northern Way' is mainly aimed at strengthening intra-regional coordination in economic and social development efforts, with a strong emphasis on the local knowledge base, and the local integration of innovation, research and education and training. The remarkable concentration of high-rank universities in the region (among others Manchester, Newcastle, Leeds, Sheffield) has acted as one of the main pillars of this strategy (e.g. Byrne and Benneworth 2006; Wilson and Baker 2006; Gore and Jones 2006).

A remarkable multinational presence and a high degree of openness[9] – as also highlighted by the magnitude of the coefficients of the Internationalisation variable – underlie the patterns of collaboration for innovation in the regions of Wales (Cooke et al. 1994, 1998; Arndt and Sternberg 2000; De Laurentis 2006) and Scotland (Raines et al. 2001; De Laurentis 2006). In Wales, only the cooperation with non-local suppliers influences the likelihood that firms are in the category of those with technological capabilities. Similarly to Wales, Scotland shows a positive impact of cooperation with extra-local suppliers and clients, possibly due to the corporate vertical integration of the multinational enterprises located in the region (Turok 1997; Raines et al. 2001).

A similar pattern to Scotland emerges for Northern Ireland's vertical cooperative linkages, with strongly positive coefficients for collaborations with extra-regional suppliers and clients. This is not surprising, due to the strong economic integration of Northern Ireland with the other UK regions.

[9] Both towards the other UK regions and towards the international markets.

5 Conclusions

The aim of this chapter has been to investigate the relationship between different forms of collaborative linkages for innovation and firms' technological competence and capabilities, considering in particular the role of the environment of the firm in the form of its regional location.

Our findings indicate that highly significant results obtained for the UK as a whole actually mask considerable differences among the regions. In particular, the findings show remarkable regional specificities in terms of the association between collaborative patterns and technological capabilities at the firm level. For instance, UK regions such as the Midlands, and even more Northern England, show the greatest evidence of utilising a richer variety of collaborative linkages at the firm level to restructure their regional systems of innovation and enhance their technological capabilities. On the contrary, the highly globalised metropolitan region of London displays a weak association between cooperative patterns and the technological status of firms located there. By the same token, local networking is also less crucial in less central – but highly open – regions such as Scotland, Wales and Northern Ireland.

The main suggestion for regional analysis is that the scope for interaction varies greatly among regions and in some contexts is potentially huge, provided that private and public resources are devoted to identifying and facilitating the most effective linkages for the observed region. In other words, managing regional interactions and cooperation in order to enhance firms' technological capabilities is not a free lunch.

In particular, regional policy should start from the distinction drawn between technological competences and capabilities, being aware that simply marshalling the resources – i.e. increasing innovation inputs – cannot be enough, and finally recognizing that fuelling the link between cooperative networking and firms' capabilities is a policy target that has to be tailored to the specific features of a given regional economic and innovation system.

On the whole, taking into account the economic and social structures of the different UK regions, policies which prove to be successful in one region may not automatically be effective in other contexts; this calls for a targeted innovation policy, bearing in mind the geographical and sectoral structures which characterise the potential beneficiaries.

Acknowledgments The authors gratefully acknowledge the funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 216813.

Appendix 1: List of Variables

Name	Nature	Mean	Standard deviation
Dependent variable			
Technological status of the firm	Categorical ordered (N = 15,153)	1.028	0.853
Technologically inactive firm = 0	N = 5,308		
Firm with technological competences = 1	N = 4,105		
Firm with technological capabilities = 2	N = 5,740		
Independent/control variables			
Cooperation partners	Dummies		
A: other enterp. LOCAL		0.036	0.187
A: other enterp. NON-LOCAL		0.059	0.235
B: suppliers LOCAL		0.046	0.209
B: suppliers NON-LOCAL		0.091	0.287
C: clients LOCAL		0.050	0.217
C: clients NON-LOCAL		0.084	0.278
D: competitors LOCAL		0.026	0.160
D: competitors NON-LOCAL		0.051	0.220
E: consultants LOCAL		0.030	0.170
E: consultants NON-LOCAL		0.049	0.216
F + G: universities&pub.res. LOCAL		0.040	0.195
F + G: universities&pub.res. NON-LOCAL		0.047	0.212
Size: Ln(Employment)	Continuous	4.043	1.504
(Number of employees)		(276.192)	(1403.906)
Group	Dummy	0.358	0.479
Internationalisation	Dummy	0.342	0.474
Start-up	Dummy	0.150	0.357
Human capital	Continuous	0.127	0.230
Sectors	Dummies		
Primary sector		0.015	
Engineering-based manuf		0.137	
Other manufacturing		0.175	
Construction		0.093	
Retail and distribution		0.165	
Knowledge-intensive services		0.169	
Other services		0.246	
Regions	Dummies		
North England (North East, North West, Yorkshire and the Humber)		0.230	
Midlands (East Midlands, West Midlands)		0.170	
Eastern England		0.085	
London		0.097	
South England (South East, South West)		0.186	
Wales		0.067	

(continued)

Name	Nature	Mean	Standard deviation
Scotland		0.077	
Northern Ireland		0.088	

References

Abramovitz M (1986) Catching up, forging ahead, and falling behind. J Econ Hist 46:385–406

Acemoglu D (1998) Why do new technologies complement skills? Directed technical change and wage inequality. Quart J Econ 113:1055–1090

Acs Z, Audretsch DB, Feldman M (1994) The recipients of R&D spill-overs: firm size and innovation. Rev Econ Stat 76:336–340

Archibugi D, Iammarino S (1999) The policy implications of the globalisation of innovation. Res Policy 28:317–336

Arndt O, Sternberg R (2000) Do manufacturing firms profit from intraregional innovation linkages? An empirical based answer. Eur Plann Stud 8:465–485

Arrighetti A, Vivarelli M (1999) The role of innovation in the post-entry performance of new small firms: evidence from Italy. South Econ J 65:927–939

Arrow K (1962) Economic welfare and the allocation of resources for invention. In: Nelson R (ed) The rate and direction of inventive activity. Princeton University Press, Princeton

Audretsch DB (1995) Innovation and industry evolution. MIT Press, Cambridge, MA

Audretsch DB, Vivarelli M (1994) Small firms and R&D spillovers: evidence from Italy. Rev Econ Ind 67:225–237

AWM (2004) Advantage West Midlands. www.advantagewm.co.uk

Bell M (1984) Learning and the accumulation of industrial technological capability in developing countries. In: Fransman M, King K (eds) Technological capability in the third world. Macmillan, London, pp 187–209

Bell M, Pavitt K (1993) Technological accumulation and industrial growth: contrast between developed and developing countries. Ind Corp Change 2:157–210

Bell M, Pavitt K (1995) The development of technological capabilities. In: Haque IU (ed) Trade, technology and international competitiveness. World Bank, Washington, pp 69–101

Belderbos R, Carree M, Lokshin B (2004a) Cooperative R&D and firm performance. Res Policy 33:1477–1492

Belderbos R, Carree M, Diederen B, Lokshin B, Veugelers R (2004b) Heterogeneity in R&D cooperation strategies. Int J Ind Organ 22:1237–1263

Breschi S, Malerba F, Orsenigo L (2000) Technological regimes and Schumpeterian patterns of innovation. Econ J 110:388–410

Budd L (2006) London: from city-state to city-region? In: Hardill I, Benneworth P, Baker M, Budd L (eds) The rise of the English regions? Routledge, London, pp 245–266

Byrne D, Benneworth P (2006) Where and what is the North East of England? In: Hardill I, Benneworth P, Baker M, Budd L (eds) The rise of the English regions? Routledge, London, pp 107–119

Cassiman B, Veugelers R (2002) R&D cooperation and spillovers: some empirical evidence from Belgium. Am Econ Rev 92:1169–1184

Chesbrough H (2003) Open innovation. Harvard Business School Press, Boston

Cohen WM, Klepper S (1996) A reprise of size and R&D. Econ J 106:925–951

Cohen WM, Levin RC (1989) Innovation and market structure. In: Schmalensee R, Willig R (eds) Handbook of industrial organization. Elsevier Science, Amsterdam, pp 1059–1107

Cohen WM, Levinthal DA (1990) Absorptive capacity: new perspective on learning and innovation. Adm Sci Q 35:128–152
Colombo MG, Delmastro M, Grilli L (2004) Entrepreneurs' human capital and the start-up size of new technology-based firms. Int J Ind Organ 22:1183–1211
Colombo MG, Grilli L (2005) Founders' human capital and the growth of new technology-based firms: a competence-based view. Res Policy 34:795–816
Cooke P (2001) Regional innovation systems, clusters, and the knowledge economy. Ind Corp Change 10:945–974
Cooke P, Gómez Uranga M, Etxebarria G (1997) Regional innovation systems: institutional and organisational dimensions. Res Policy 26:475–491
Cooke P, Heidenreich M, Braczyk H-J (eds) (1998) Regional innovation systems, 2nd edn. Routledge, London
Cooke P, Morgan K, Price A (1994) The welsh renaissance: inward investment and industrial innovation. Centre for Advanced Studies, Cardiff
Criscuolo C, Haskel JE (2003) Innovation and productivity growth in the UK: evidence from the CIS2 and CIS3. Centre for research into business activity working paper
de Laurentis C (2006) Regional innovation systems and the labour market: a comparison of five regions. Eur Plann Stud 14:1060–1084
D'este P, Iammarino S, Savona M, von Tunzelmann N (2008) What hampers innovation? Evidence from the UK CIS4, SPRU electronic working paper series (SEWPS), no. 168
Evangelista R, Iammarino S, Mastrostefano V, Silvani A (2002) Looking for regional systems of innovation: evidence from the Italian innovation survey. Reg Stud 36:173–186
Faems D, van Looy B, Debackere K (2005) The role of inter-organizational collaboration within innovation strategies: towards a portfolio approach. J Prod Innov Manage 22:238–250
Filatotchev I, Piga C, Dyomina N (2003) Network positioning and R&D activity: a study of Italian groups. R&D Manage 33:37–48
Fritsch M, Lukas R (2001) Who cooperates on R&D. Res Policy 30:297–312
Gore T, Jones C (2006) Yorkshire and the Humber. In: Hardill I, Benneworth P, Baker M, Budd L (eds) The rise of the English regions? Routledge, London, pp 137–155
Gray M, Martin R, Tyler P (2006) The East of England: a nebulous region in transition. In: Hardill I, Benneworth P, Baker M, Budd L (eds) The rise of the English regions? Routledge, London, pp 184–203
Green A, Berkeley N (2006) The West midlands: the 'Hinge' in the middle. In: Hardill I, Benneworth P, Baker M, Budd L (eds) The rise of the English regions? Routledge, London, pp 156–167
Hardill I, Bentley C, Cuthbert M (2006) The east Midlands: the missing middle? In: Hardill I, Benneworth P, Baker M, Budd L (eds) Rise of the English regions? Routledge, London, pp 168–183
Hassink R (1997) Localized industrial learning and innovation policies. Eur Plann Stud 5:279–282
Hobday M (1995) Innovation in East Asia: the challenge to Japan. Edward Elgar, London
Howells J (1999) Regional systems of innovation? In: Archibugi D, Howells J, Michie J (eds) Innovation policy in a global economy. Cambridge University Press, Cambridge, pp 67–93
Iammarino S (2005) An evolutionary integrated view of regional systems of innovation: concepts, measures and historical perspectives. Eur Plann Stud 13:495–517
Iammarino S, Padilla R, von Tunzelmann N (2008) Technological capabilities and global–local interactions: the electronics industry in two Mexican regions. World Dev 36:1980–2003
Kleinknecht A, Poot TP (1992) Do regions matter for R&D? Reg Stud 26:221–232
Kleinknecht A, Reijnen JON (1991) More evidence on the undercounting of small firm R&D. Res Policy 20:579–587
Kline SJ, Rosenberg N (1986) An overview of innovation. In: Landau R, Rosenberg N (eds) The positive sum strategy: harnessing technology for economic growth. National Academy Press, Washington DC, pp 275–306
Lall S (1992) Technological capabilities and industrialization. World Dev 20:165–186

Laursen K, Salter AJ (2004) Searching low and high: what types of firms use universities as a source of innovation? Res Policy 33:1201–1215
Laursen K, Salter AJ (2006) Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms. Strat Manage J 27:131–150
Lundvall BA (1988) Innovation as an interactive process: from user-producer interaction to the national system of innovation. In: Dosi G, Freeman C, Nelson R, Silverberg G, Soete L (eds) Technical change and economic theory. Pinter, New York, pp 349–369
Machin S, van Reenen J (1998) Technology and changes in the skill structure: evidence from seven OECD countries. Quart J Econ 113:1215–1244
Malerba F (1992) Learning by firms and incremental technical change. Econ J 102:845–859
Miotti L, Sachwald F (2003) Co-operative R&D: why and with whom? An integrated framework of analysis. Res Policy 32:1481–1499
Narula R, Zanfei A (2003) The international dimension of innovation. In: Fagerberg J, Mowery DC, Nelson RR (eds) The oxford handbook of innovation. Oxford University Press, Oxford, pp 318–345
Nelson RR, Winter SG (1982) An evolutionary theory of economic change. Harvard University Press, Cambridge, MA
Newman A, Thornley N (2005) Planning world cities: globalization and urban politics. Palgrave Macmillan, New York
OECD (2005) Oslo manual. OECD, Paris
Piga C, Vivarelli M (2004) Internal and external R&D: a sample selection approach. Oxf Bull Econ Stat 66:457–482
Piva M, Santarelli E, Vivarelli M (2005) The skill bias effect of technological and organisational change: evidence and policy implications. Res Policy 34:141–157
Piva M, Vivarelli M (2004) The determinants of the skill bias in Italy: R&D, organisation or globalisation? Econ Innov New Technol 13:329–347
Piva M, Vivarelli M (2009) The role of skills as a major driver of corporate R&D. Int J Manpower 30:835–852
Raines P, Turoki I, Brown R (2001) Growing global: foreign direct investment and the internationalization of local suppliers in Scotland. Eur Plann Stud 9:965–978
Revilla Diez J (2000) The importance of public research institutes in innovative networks-empirical results from the metropolitan innovation systems Barcelona, Stockholm and Vienna. Eur Plann Stud 8:451–463
Rodriguez-Pose A, Comptour F (2012) Evaluating the role of clusters for innovation and growth in Europe. In: Crescenzi R, Percoco M (eds) Geography, institutions and regional economic performance, Springer-Verlag, Berlin
Romijn H (1999) Acquisition of technological capability in small firms in developing countries. Macmillan, Basingstoke
Schumpeter J (1943) Capitalism, socialism and democracy. Unwin University Press, London
Schienstock G (2009) Organisational innovations and new knowledge management practices, their diffusion and influence on firm performance: results from a Finnish firm survey", IAREG working paper, No. 1.2.d, April
Simonen J, McCann P (2008) Firm innovation: the influence of R&D cooperation and the geography of human capital inputs. J Urban Econ 64:146–154
Storey DJ, Tether BS (1998) New technology-based firms in the European Union: an introduction. Res Policy 26:933–946
Turok I (1997) Linkages in the Scottish electronics industry: further evidence. Reg Stud 31:705–711
Vanhaverbeke WPM, Duysters GM, Beerkens BE (2002) Technology capability building through networking strategies within high-tech industries. Academy of Management Papers, Denver
von Hippel E (1988) The sources of innovation. Oxford University Press, New York
von Tunzelmann N (2009a) Competencies and capabilities: a reassessment. Econ Pol, special issue on technologies and organizations, 3 Dec 2009

Von Tunzelmann N (2009b) Regional capabilities and industrial regeneration. In: Farshchi M, Janne O, Mccann P, Farshchi M, Janne O, Mccann P (eds) Technological change and mature industrial regions: firms, knowledge and policy. Edward Elgar, Cheltenham, pp 11–28
von Tunzelmann N, Wang Q (2003) An evolutionary view of dynamic capabilities. Econ Appliquée 6:33–64
von Tunzelmann N, Wang Q (2007) Capabilities and production theory. Struct Change Econ Dynam 18:192–211
Westphal LE, Kritayakirana K, Petchsuwan H, Sutabutr, Ythavong Y (1990) The development of technological capability in manufacturing: a macroscopic approach to policy research. In: Evenson RE, Ranis G (eds) Science and technology: lessons for development policy. Intermediate Technology Publications, London, pp 81–134
Wignaraja G (2002) Firm size, technological capabilities and market-oriented policies in Mauritius. Oxf Develop Stud 30:87–104
Wilson C, Baker M (2006) The North West: cultural coherence and institutional fragmentation. In: Hardill I, Benneworth P, Baker M, Budd L (eds) The rise of the English regions? Routledge, London, pp 120–136

Part IV
Geography in Motion: Trade, FDI and Migrations

Assessing Regional Economic Performance: Regional Competition in Spain Under a Spatial Vector Autoregressive Approach

Miguel A. Márquez, Julián Ramajo, and Geoffrey J.D. Hewings

Abstract After a wave of empirical literature on regional competition focused on issues related to regional convergence, work developed recently has tried to address some of the shortcomings of the previous literature, developing a series of alternative approaches that center their attention on the assessment of regional economic performance. These alternative methodologies embrace a complex set of space-time interactions that take into account that a single region's economic performance affects and is affected by other regions. In this context, and as an original contribution of this chapter, a spatial vector autoregressive (SpVAR) model for the Spanish regions during the period 1955 –2009 is presented. The SpVAR model considers spatial as well temporal lags of the variables. The estimated SpVAR is used to calculate impulse responses that provide insights about the effects of shocks to relative regional productive capacity on different regions. The empirical results suggest that the existence of trade linkages have had different significant impacts on the production shares of the 17 Spanish regions, but competition between regional economies prevails.

Keywords Regional growth • Regional competition • Spatial VAR

JEL codes: C31, C33, J24, O18, R11

M.A. Márquez (✉) • J. Ramajo
Department of Economics, University of Extremadura, Badajoz , Spain-EU

Regional Economics Applications Laboratory (REAL), University of Illinois, Urbana-Champaign, IL, USA
e-mail: mmarquez@unex.es

G.J.D. Hewings
Regional Economics Applications Laboratory (REAL), University of Illinois, Urbana-Champaign, IL, USA

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_15,

1 Introduction

Spurred by the initial contributions of Barro and Sala-i-Martin (1991), there has been a significant increase in attention to regional competition (see Batey and Friedrich 2000b), initially focused on issues related to regional convergence. While there is clear evidence, derived from many alternative econometric specifications, that a single region's economic performance affects and is affected by other regions, there is less agreement on (1) appropriate ways to measure the interactions between regions and (2) the way in which both complementary and competitive linkages and externalities in general may be modeled. The empirical evidence shows that by adopting more dynamic specifications, regional economic performance can be seen to embrace a complex set of space-time interactions. In these recent approaches, the main underlying hypothesis centers on the existence of regional growth spillovers, where the dynamics of one regional economy influences growth of neighboring regional economies (Cheshire and Carbonaro 1996).

As a point of departure, it is essential to evaluate regional economic performance (Armstrong and Taylor 2000); the importance of this type of research is clear. This evaluation has long been assumed a central role in the context of the European Union (López-Bazo et al. 1999). Nevertheless, little empirical economic analysis has been completed on the regional economic performance of regions as economic units. The empirical literature to date has focused mainly on the growth of regional output per capita, that is, changes in economic welfare (see, for example, Ramos et al. 2010). The major shortcoming of using regional output per capita as an indicator of disparity comes from the statistical point of view of associating better conditions of relative wellbeing with movements of population from an area (Capello 2009a). This paper attempts to add a new perspective to the empirical literature by focusing on the growth of relative regional productive capacity. Thus, the empirical results could be regarded as complementary to explorations of convergence in properties of regional economies (see Sala-i-Martin 1996, and Cuadrado-Roura 2001).

As stated by Capello (2009b), the conceptual interpretation of spatial spillovers highlighting the role played by geographical proximity should be enriched making reference to territorialized channels through which spillover effects are spread around. In this sense, growth spillovers would make reference to influences from a single regional economy to the growth of neighboring regional economies through trade linkages and market relationships. From this view, the performance of a regional economy as a whole in terms of its relative size will be determined mainly by external conditions based on both demand and supply effects. As demand for the region's outputs and supply of inputs to the productive activity of the region are the fundamental external factors affecting the overall level of activity, the existence of trade linkage conditions could induce different regional economic results. Obviously, the performance of a regional economy would also be influenced by demand and supply relationships within the region itself.[1]

[1] For example, regional performance can be significantly affected through changes in intraregional input endowments and intersectoral relationships.

In this work, the guiding question is the following: How does the strength of regional trade linkages affect the relative productive capacity of a region within a regional system? In the context of regional economic competition, this is important, since the answer could uncover the patterns of competitive relationships related to trade linkages among regions as economic units in a regional economic system. Trade linkages turn out to be a major source of interaction between regions, generating complementarities (through supply chains associated with the production of final goods and services) that often transcend simple notions embodied in standard weight matrices (see Hewings 2008). Further, the aggregate trade interactions between regions often mask considerable heterogeneity in the sectoral dependencies, generating further complexities in the behavior of regional business cycles. In many countries, interregional trade data are not available, necessitating some nonsurvey methods to estimate the interactions. In the present case, this was accomplished by the construction of a simple empirical dynamic model inspired from a synthesis between the traditional version of the pure gravity type model and the empirical regional curves provides by Márquez, et al. (2006). The estimated econometric model will reveal evidence about how the relative regional economic sizes are connected and interdependent.

Thus, and as original contribution of this chapter, a spatial vector autoregressive (SpVAR) model for the Spanish regions during the period 1955–2009 is presented. The SpVAR model considers spatial as well temporal lags of the variables. The estimated SpVAR is used to calculate impulse responses that provide insights about the effects of shocks to relative regional productive capacity on different regions. The empirical results suggest that the existence of trade linkages have had different significant impacts on the relative productive capacity for the 17 Spanish regions, but competition between regional economies prevails.

The rest of the paper is organized as follows. After the theoretical background (Sect. 2) and an exploratory analysis of the Spanish regional economic system (Sect. 3), the core of the paper (Sect. 4) presents an empirical dynamic space-time model. The usage of spatial vector autoregressive -SpVAR- specifications allows examination of the temporal evolution of the Spanish *regional shares*, providing the effects of shocks in relation to the size of regions. Sect. 5 provides some concluding remarks.

2 Conceptual Background and Motivation

In this section, the concept of regional competition is addressed and then some discussion is provided for the theoretical background about regional growth and regional trade linkages that sustains our empirical approach. Recently, the question about whether externality effects among regions are producing positive results on the outcome of regional growth processes has attracted the attention of a large number of studies (see, for example, Márquez, et al. 2010). Even though the recognition of these external effects among regions would imply the consideration of regional competition as a main topic on the research agenda of regional economists, the issue of regional competition is still generally omitted. The competitive relations among regions have been neglected because, drawing on the convergence predictions of the traditional neoclassical model of growth, differences in welfare across regions have been one of the main focuses of empirical research. Thus, following the contributions of Barro (1991) and Barro and Sala-i-Martin (1991, 1992), a large number of studies have analyzed whether incomes are converging across regional economies. This key prediction of the neoclassical growth model (spatial disparities in regional per capita income should converge over the long run) has been tested by a large number of empirical works (see Ramajo et al. 2008).

Nevertheless, the analysis of regional growth of output per capita focuses on the changes in economic welfare, while these empirical methods fail to properly account for regional economic performance. There is an extended perception that regional welfare is associated with the success of the region as an economic unit; however, this is not so. For example, consider the case of the Basque Country, a region in Spain that during the period 1955–2007 lost part of its share of regional productive capacity within the Spanish regional economic system, but it improved its output per capita. In this context, the growth of relative regional productive capacity is a complementary perspective that needs to be addressed.

In considering this issue, and following among others Glaeser et al. (1995), population movements would be the best measure of relative spatial welfare differences. Nonetheless, as noted in Cheshire and Magrini (2009), this argument would not be sustainable in a European context, where population mobility still seems to be confined within national borders; thus, changes in welfare levels could be measured by mean differences in the growth of real incomes (real GDP percent). Although the use of the regional shares of GDP would be more relevant to the analysis of the sources of productivity rather than to the analysis of income growth, in general, one would expect that they should be highly correlated.

The present chapter will develop a procedure that will detect significant movements in the evolution of the regional shares within a regional system, complementing the analysis conducted in the 'per capita literature.' This way, a competitive perspective arises that follows the notion of regional competition stated by Parr (1978, p. 122): "Broadly speaking, regional competition may be regarded as the market process by which economic activities or employed factors of production are allocated through time among the regions of a nation. Phrasing this in terms of national income, regional competition represents the process by which the gross national product (GNP) is distributed among regions. The overall competitiveness of a particular region can thus be measured by the region's share of the GNP, although a more useful view of a region's competitiveness might be the extent to which it is able to maintain or increase its share of the GNP through time". The procedure adopted implies inter-regional competition and, in the line of the models of competitive growth (Richardson 1973), national growth is assumed. Following the theoretical background of the models of competitive growth identified by Richardson, a top-down approach is adopted (disregarding the effects of the regional growth on the national level performance), considering that if a regional economy increases its share one or more of the rest of the regional economies have to reduce their shares. Under these conditions, a top-down non-structural approach was adopted.

At regional level, it is interesting to know if the existence of interactions among regional economies is producing a specific form of the regional structural competition pattern, contributing to an improvement in decision-making for policymakers (see Friedrich 1987). The issue of measurement of regional competition is a complex one. An option would be to assume the construction of a synthetic indicator of regional competitiveness that combines the complex time-interaction of a number of regional indicators (see Poot 2000; Batey and Friedrich 2000a). Hewings et al. (1996) focused on the kind of relationships (competition or complementarity) that exist for regions within a context where it is only possible for regions to achieve growth in the relative distribution of regional output at the sacrifice of other regions. Their method is based on a loglinear relative dynamics to explore the regional interaction in multiregional growth. These authors work with relative regional shares, like Márquez et al. (2003), who introduce some modifications over the basic specification of Hewings et al. (1996). The new assumptions generate an empirical model that allows detection of the underlying regional economic interconnections in the process of multi-regional dynamic growth. The use of relative shares is abandoned by Márquez and Hewings (2003), who uncover the patterns of geographical competition in the Peninsular Spanish Regional system. Márquez et al. (2006) proposed a dynamic space-time empirical model that provides the competition structure within a Spanish multiregional economic system over both the short and long term. Cointegration and error-correction modeling techniques were used. In these recent approaches, the main underlying hypothesis is the existence of regional growth spillovers, where the dynamics of one regional economy influences growth of neighboring regional economies (Cheshire and Carbonaro 1996).

All the aforementioned papers share considerable attention being directed to the study of space-time regional interactions, connecting with the current interest about how interregional externalities affect regional growth (Egger and Pfaffermayr 2006). Externalities are introduced into the arguments of both endogenous growth theory (new growth theory) and of the new economic geography (see Feser 1998). Usually, the concept of externality is used to describe a great variety of situations (Fujita and Thisse 2002), although there is consensus about its crucial role in the formation of economic agglomeration (Marshall 1890); the analysis of the geographic agglomeration of economic activity has been stimulated by Krugman (1991). In addition, the geographic concentration of economic activities within a country could be significantly influenced by the volume and nature of trade (Haaparanta 1998). It is well known that trade is a major source of explanation of economic growth (Fujita and Thisse 2002) and the evaluation of the impact of the existence of trade relationships on growth has been analyzed in the context of economic integration. The main results state that economic activity in integrating economies tends to be increasingly agglomerated. For example, and addressed under an international perspective, Krugman and Venables (1995) show how greater integration of the global economy through trade can induce regional agglomerations of activities that promote inequality. These authors improve the study of the dynamics of regional disparities, illustrating the existence of a bell-shaped relationship between economic integration and spatial inequality. To the extent that this process occurs at the regional level, we would need empirically-based models that could be used to measure the effects of regional trade pattern on regional economic performance (Combes et al. 2008). Interregional trade, understood as the exchange of goods and services over a regional system, would configure long-run equilibrium settlement patterns of trade (Fujita et al. 1999).

The New Economic Geography adopts imperfect competition and increasing returns, highlighting the influence of country and market size, and predicting the existence of growing economic disparities across a multiregional system. Trade and factor mobility will be the main determinants fostering a growing divergence across regions. The empirical evidence shows that regional economies rely heavily on trade for their economic wellbeing. Since this paper addresses the effects of regional trade patterns, some discussion is needed of the existence of trade connections and their affect on the spatial distribution of economic activity. It may do so more through competition (and thus, less production) and through the size of the market shares (more regional production), although as noted earlier, there is increasing evidence of increases in complementarity between regions (generated by the fragmentation of production) while, at the same time, regions are competing with each other to locate or retain production within specific value chains.

While the analysis of the evolution of concentration and specialization patterns has been analyzed in different studies at a sectoral level (for empirical evidence, see, for example, Imbs and Wacziarg 2003; Aiginger and Davies 2004; or Kancs 2007) there are few contributions about how to empirically assess the multifaceted impacts of the existence of trade relationships at an interregional level. Hence, we will concentrate our attention on the existence of regional trade linkages as the main driver of the dynamic external determinants of the relative size of a regional economy within a regional economic system. Regions share characteristics with small open economies, exhibiting a high dependence on trade; however, in the case of regions, trade comprises both domestic and international. For regions in the more developed economies, interregional trade is usually a much larger component of total trade, hence the need to focus on this dimension. Thus, a clear geographical dimension is present in regional economic growth through regional spillovers related to the existence of trade linkages. Regional trade spillovers exert positive or negative effects on other regions by means of different transmission channels where trade linkages occupy a relevant place (Capello 2009b).

Growth spillovers from countries have also been attracting increasing attention. The relevance of space and trade for understanding cross-country patterns of economic growth and convergence has been emphasized by, among others, Quah (1997). Arora and Vamvakidis (2005) show that a country's growth is strongly correlated with trading partner growth, even after controlling for common global and regional trends. It is important to emphasize that these authors affirm that the relative importance of a country's trading partners tends to be stable over time. In all probability, the relative importance of regions' trading partners tends may not change much over time. This stability is underlined from another perspective by the gravity relationship: it states that economic interactions between two regions are proportional to the size of these regions and inversely related to the distance between them. Different studies have shown a robust empirical regularity of bilateral trade flows between pairs of regions; these trade flows are well explained by the size of their gross domestic products. In a similar vein, previous studies report that the overall world trade network structure is fairly stationary through time (see for example, Serrano and Boguña 2003; Garlaschelli and Loffredo 2005; and Fagiolo 2010). For the Spanish regional system, Llano (2004) and Pérez et al. (2008) show evidence about how regional trade linkages act as transmission chains for interregional spillovers in terms of growth, employment and productivity. The stability of the regional trade pattern for each Spanish region throughout the period 1995–2005 indicates the presence of important interregional inter-sectoral linkages between them (Llano et al. 2010).

These regularities at both international and interregional levels could be used as the point of departure for the consideration of trade linkages among regional economies. What is perhaps both new and interesting is a test for the extent to which the existence of trade linkages could have an unfavorable or favorable impact on the regional changes of relative productive capacity. Following Kali and Reyes (2007), rather than focusing on trade levels of individual regions, we will consider the pattern of linkages that tie together regions within a regional system. Previous work has overlooked these regional trade linkages, although they would appear to be quite relevant and that they could lead to new ways of understanding the relationship between trade and regional growth. How do trade linkages affect the production share of a region within a regional system? In this chapter, this question is explored empirically. Under this context, albeit in an indirect way, the relationship between growth in regional productive capacity and regional trade connections will be emphasized. It will be hypothesized that the existence of trade relationships among regional economies is a relevant feature to help explain the evolution of regional shares. Besides, the impulse responses are calculated to provide insights about the effects of shocks in relative productive capacity in different locations on a region's share. The analysis of the stability of relative regional economic sizes to shocks in the size of its trade-related regional neighborhood may be fruitful.[2]

The New Economic Geography emphasizes the role of the so-called *second nature geography* (distance to consumer markets and distance to input suppliers) as a way of explaining differences in income levels among regions or countries. In this sense, Davis and Weinstein (1999) found that economic geography is very

[2] The stability of city sizes to substantial shocks were examined by different works (see, for example, Davis and Weinstein 2002 or Brakman et al. 2004), being the general conclusion that they remain stable (Combes et al. 2008). Nevertheless, the analysis of the stability of regional sizes to shocks within the system is a pending task.

important in determining the structure of regional production. Our path dependent hypothesis on interregional trade could find support on the reasons argued by these authors: the existence of low costs for trade between regions of a country would lead to stronger regional effects derived from second nature geography; besides, the greater mobility of factors across regions will tend to reinforce these economic geography effects. This way, interregional trade would be the transmission channel through which regional economies interact.

But economic geography is not the unique mechanism through regional trade patterns are stable. For example, Roy and Thill (2004) state that some historical trade patterns are based on cultural affinities and perceptions of quality and reliability of delivery.

3 Data and Stylized Facts

The aim of this section is to trace the evolution of the regional economies in Spain from 1955 to the present. The main message that emerges of this analysis is the already known "Spanish regional puzzle" (Garrido-Yserte and Mancha-Navarro 2010); convergence is observed in relative terms (for income per capita and particularly for productivity), but divergence and higher concentration is observed in absolute terms (for production, employment or population).[3]

Spain is a decentralized state of the European Union comprising 17 regions (the so-called Autonomous Communities) and Ceuta and Melilla, two Spanish cities of North Africa. The database that is used contains data for Spain and its regions over the period 1955–2009, and was constructed from data of the *Instituto Nacional de Estadística* (INE), the *Instituto Valenciano de Investigaciones Económicas* (IVIE) and the BD.MORES, REGDAT and C-INTEREG databases (de Bustos et al. 2008; de la Fuente 2010; Llano et al. 2010).

It is well known that the regional performance in Spain between 1955 and 2009 can be considered as highly significant in terms of gross value added (GVA) per capita. During this period, the cumulative average rates of GVA per capita growth for the Spanish regions were very important, with less-developed regions experiencing notable improvements (Castile and Leon, Castile-La Mancha, Extremadura and Galicia), although advances were fairly slight in certain cases (Andalusia and Murcia). At the same time, the most-developed regions in 1955 (Balearic Islands, Catalonia, Madrid and Basque Country) have registered a relative decline. Both facts imply that there has been a 'catching-up behavior' of income per capita within the Spanish regional system (Table 1).

This result can be corroborated by analyzing the relationship between the initial level and the growth of GVA per capita observed between 1955 and 2009 in the

[3] The above-mentioned work uses data only from 1986 to 2007, but the main conclusions remain in our research.

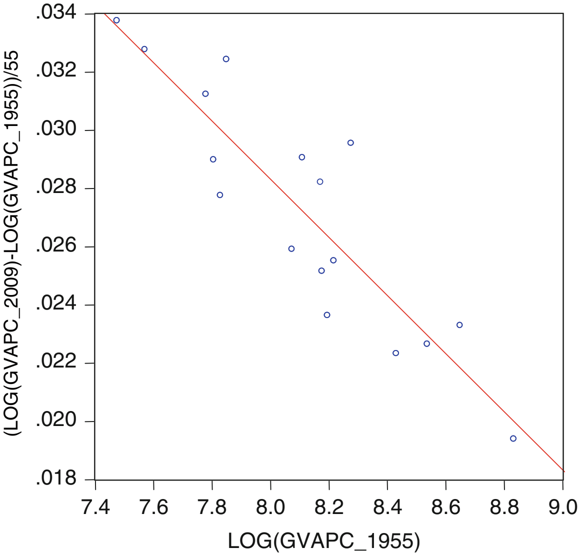


Fig. 1 Beta-convergence in GVA per capita among the Spanish regions

Spanish regions. As shown for the whole period in Fig. 1, there is an inverse relationship between these variables (β-*convergence*). In statistical terms, a negative and statistically significant estimation of the β parameter in the regression:

$$(1/T)\log(GVA_{i,2009}/GVA_{i,1955}) = \alpha + \beta \log GVA_{i,1955} + e_i$$

would indicate the presence of a beta-convergence process. This process would evidence the existence of convergence in terms of both income level and growth rate. In the estimates that were made, a negative and highly significant β parameter was obtained ($\beta = -0.009989$, $P - value = 0.0000$), pointing to an annual reduction of regional disparities in GVA per capita of nearly 1 %.

The evolution of the regional inequality level over time can also be analyzed using the concept of σ-*convergence*, associated with a reduction of the dispersion of the GVA per capita in the regions under consideration. This dispersion is measured in this case through the cross-section index:

$$\sigma_t = \left[\frac{1}{N}\sum_i \left(\log GVA_{it}/GVA_{SP,t}\right)^2\right]^{1/2}$$

where GVA_{SP} is the GVA per capita of Spain. Results are presented in Fig. 2, showing a rapid process of sigma-convergence from 1955 to the end of the 1970s and a fairly slight process of convergence from 1980 onwards.

Behind the observed changes in regional GVA per capita, there are changes in other variables, specifically in productivity and employment figures. Theil's inequality index was used to explore the role of labor productivity and employment

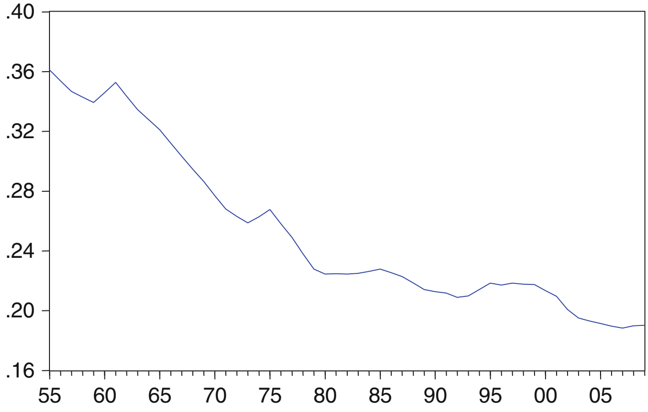


Fig. 2 Sigma-convergence in GVA per capita among the Spanish regions

per inhabitant in the explanation of the evolution of regional differences of GVA per capita. The Theil index can be decomposed in each moment *t* as:

$$\sum_i \frac{GVA_i}{GVA_{SP}} \log\left(\frac{GVA_i/N_i}{GVA_{SP}/N_{SP}}\right) = \sum_i \frac{GVA_i}{GVA_{SP}} \log\left(\frac{GVA_i/E_i}{GVA_{SP}/E_{SP}}\right) + \sum_i \frac{GVA_i}{GVA_{SP}} \log\left(\frac{E_i/N_i}{E_{SP}/N_{SP}}\right)$$

where N represents the population and E the number of workers in the region i or in Spain. As observed in Fig. 3, the Spanish regional inequality in GVA per capita can mostly be explained through differences in productivity from 1955 to the end of the 1980s, and through differences in the employment-population ratio from the 1990 onwards. Throughout the entire period, a significant convergence in labor productivity levels can be observed. Also, a divergence process in relative employment is observed from 1980 onwards. Both facts explain the stagnation path of the income inequality since the beginning of the 1980s, as the Theil index tends to stabilize in this period.

Despite the convergence process observed between the Spanish regions over the whole five decades analyzed, the distance between the most developed and the less developed regions has become more pronounced. As shown in Fig. 4, if a distribution dynamics approach[4] is used to illustrate the extent and evolution of per capita income differences across Spanish regions, the probability mass of GVA per capita was much more concentrated around the mean in 1955 than in 1980, which can be interpreted as a sign of divergence in regional income distribution. Moreover, the figure shows a bimodal shape (Quah 1996) for the distribution in 2009, one peak corresponding to low-income and the other to high-income regions.

[4] This approach views the convergence process as a question about the evolution of the cross-section distribution of income, focusing the attention into the entire distribution and not in particular regions.

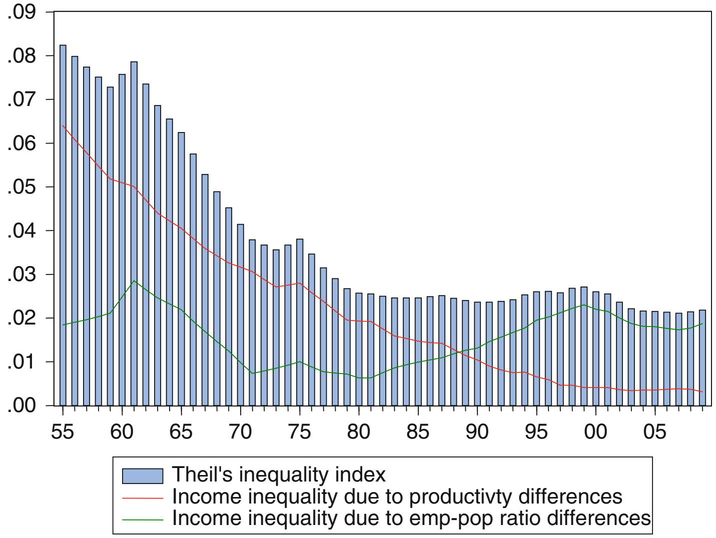


Fig. 3 Theil's inequality index for GVA per capita in the Spanish regional system

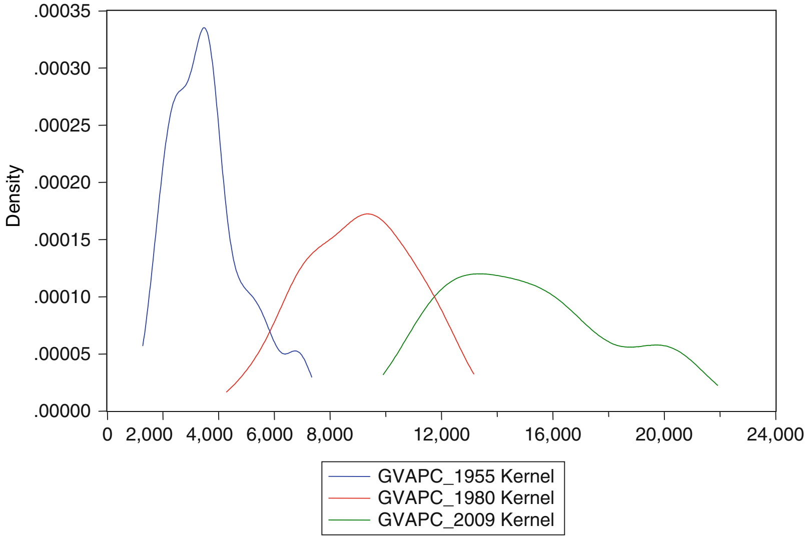


Fig. 4 Distributions of GVA per capita in Spain for 1955, 1980 and 2009 (Note: The plots are densities calculated non-parametrically using a Gaussian kernel with optimal bandwidth chosen as suggested in Silverman 1986)

Turning attention to the process of spatial concentration of production in the Spanish economy, Fig. 5 represents the time evolution of regional shares of GVA over the period 1955–2009. Spatial agglomerations of production in Spanish

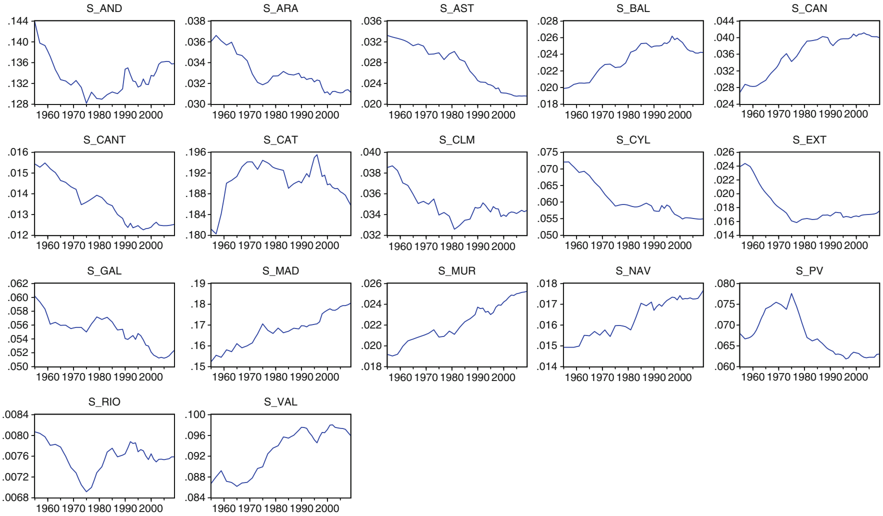


Fig. 5 Evolution of regional shares of GVA in Spain, 1955–2009 (Note: Regional abbreviations: Andalusia (AND), Aragón (ARA), Asturias (AST), Balearic Islands (BAL), Canary Islands (CAN), Cantabria (CANT), Catalonia (CAT), Castile-La Mancha (CLM), Castile and León (CYL), Extremadura (EXT), Galicia (GAL), Madrid (MAD), Murcia (MUR), Navarre (NAV), Basque Country (PV), La Rioja (RIO), and Valencian Community (VAL))

regions are reflected in this figure: Madrid, which has a geographical area corresponding to 1.6 % of the country's total, produced 18 % of the total GAV of Spain in 2009; conversely, Extremadura, with 8.2 % of the total area accounted for 1.7 % of the Spanish GAV in 2009. The main question addressed in this paper concerns which geographical dimensions are causing the trajectories observed for the shares (Figure 5).

Over the time period considered, there has been a clear process of concentration of activity, as shown in Fig. 6 through a spatial concentration index and an entropy diversity index. Both indices show that the concentration of production in some regions has increased over time, illustrating the decline of regional production diversity in Spain (Andalusia, Catalonia, the Valencian Community and Madrid accounted for almost 60 % of the total GVA in 2009), which was more accentuated from 1985 onwards.

4 The Empirical Model

In this section, by means of a proposed dynamic space-time approach (the spatial vector autoregressive –SpVAR- specifications), a competitive perspective of regional growth is adopted through the examination of the temporal evolution of the *regional shares*. The aim of the empirical application presented here is not to

Table 1 Evolution of regional income (GVA) per capita (Spain = 100)

	1955	1980	2009
Andalusia	72.50	75.28	76.06
Aragon	96.01	102.62	108.18
Asturias	102.66	98.93	93.33
Balearic Islands	132.34	134.56	103.00
Canary Islands	92.62	106.45	87.79
Cantabria	106.92	101.32	99.16
Castile and León	74.01	86.53	100.42
Castile-La Mancha	55.99	75.70	77.37
Catalonia	147.25	120.97	116.65
Valencian Comm.	104.62	96.23	87.48
Extremadura	50.86	58.10	74.27
Galicia	69.02	77.31	87.67
Madrid	197.82	133.60	131.04
Murcia	70.85	83.27	79.53
Navarre	113.37	117.95	131.26
Basque Country	164.65	120.20	135.11
La Rioja	84.05	108.79	109.84

provide estimates of future regional GVA levels or GVA growth rates, but to explain the spatial distribution of the Spanish GVA, attempting to detect the forces driving the regional shares over time.

The focus of our econometric analysis is not to estimate a fully-fledged economic model (for an example, see Márquez et al. 2010); instead, the analysis provides some insights into the relevance of spillover effects in the spatial distribution of the Spanish output. Thus, in the empirical application, a basic bivariate SpVAR specification is used, mainly aiming at assessing the impact on the relative production of a region (measured as its share) of a shock to the relative output in neighboring areas.

To uncover possible local differences in the parameters of the statistical processes, spatially heterogeneous specifications are considered. Specifically, the SpVAR model for each region is assumed to be of the form[5]:

$$\Delta Z_{it} = \Gamma_{0i} + \Gamma_{1i}\Delta Z_{i,t-1} + \ldots + \Gamma_{pi}\Delta Z_{i,t-p} + \Phi_{0i}t + \Phi_{1i}\Delta X_{i,t} + U_{it}$$

$$t = 1, 2, \ldots, T,\ i = 1, 2, \ldots, N$$

where Δ is the first-difference operator ($\Delta z_t = z_t - z_{t-1}$); $Z_{it} = (s_{it}, s^*_{it})$, the (local) variable s_{it} denoting the share of the production of the region i in the national

[5] We assume that the original variables contain temporal unit roots and are therefore nonstationary. In this case, our SpVAR specification is estimated in first differences, and not levels, to avoid spurious regressions (Phillips and Moon 1999). To check for non-stationarity, unit roots tests were performed both at the individual and panel levels. Both types of tests fail to reject the null hypothesis of non-stationarity for all the variables. For the sake of brevity, we omit the details of this and other intermediate outputs. Complete results can be request from the authors.

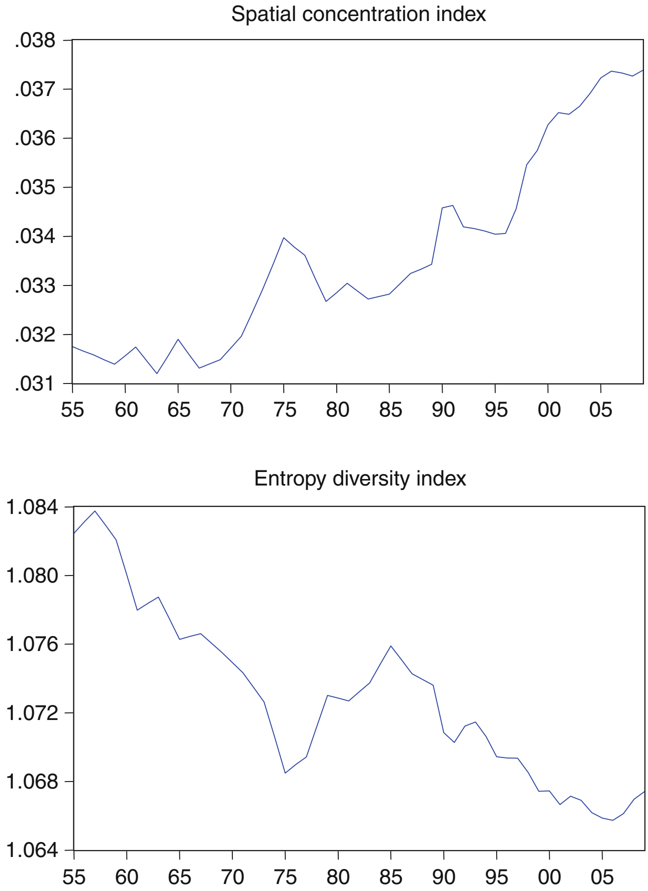


Fig. 6 Spatial concentration of production in the Spanish regions (Note: First plot represent the spatial concentration index, calculated as $\left(\frac{N}{N-1}\right)\sum_{i=1}^{N}\left(s_i-\frac{1}{N}\right)^2$; second plot represent the entropy diversity index, calculated as $-\sum_{i=1}^{N} s_i \log s_i$)

economy at time period t, $s_{it} = GVA_{it}\big/GVA_{SP,t}$, and the asterisked (external) variable s_{it}^* referring to the spatial lag of s_{it}, defined as $s_{it}^* = \sum_{j\neq i} w_{ij}s_{ij}$, where w_{ij} are row-summed spatial weights with $\sum_j w_{ij} = 1$; $X_{it} = \left(x_{1,it}, x_{2,it}, \ldots, x_{K,it}\right)$ denotes a vector of K observed covariates that are hypothesized to affect the regional shares; Γ_{ji} ($j = 0,1,\ldots, p$) and Φ_{ji} ($j = 0,1$) are matrices of coefficients to be estimated; and U_{it} is a bivariate vector of non-autocorrelated disturbances with mean zero and a nonsingular covariance matrix, Σ_i.

As the performance of a regional economy as a whole in terms of its relative size will be influenced by external and internal conditions, the empirical specification accounts for both external and internal factors. Consequently, although the specification provided in this paper is not directly derived from a specific model, it could be framed within different general theoretical contexts. Thus, first, our empirical approach could be related to the gravity model of trade. Although the gravity-based benchmark is used for the analysis of the volume of trade, the foundations of gravity could be applied to other types of flows such as capital and knowledge flows. In the current context, the underlying premise would be that the relative productive capacity of a region within a regional system is influenced by gravity forces. According to Newton's theory of gravitation, two regions are attracted to each other in proportion to the product of their sizes and in inverse proportion to the distances separating them. In the context of international trade (Tinbergen 1962), the sizes of the countries were given by their GDPs, while flows were measured by the imports and exports of countries.[6]

Secondly, and from another perspective, in the empirical specification, lagged regional shares of production are introduced (see Sect. 4); these shares could be used as a proxy of the home market effect underlined in the new economic geography (see Krugman 1980, and Helpman and Krugman 1985 for the theoretical basis and Davis and Weinstein 1999, 2003, and Baldwin et al. 2003 for empirical estimations). In the same way, the weighted share of the regions that have trade connections with the region of reference can be seen as a proxy of the real market potential (Redding and Venables 2004). Real market potential has been used recently in different empirical studies (see, for example, López-Rodríguez and Faiña 2007). In Redding and Venables (2004), a distinction was made between separate terms with different impacts: the domestic and foreign components of the market potential. In the present context, the market potential of a regional economy is separated into a term accounting only for the intra-regional demand and another term considering the (trade cost-weighted) demand from all other regions in the Spanish Regional System.

With respect to the vector of conditioning variables, X, two types of variables were introduced: local and national (Table 2). The local variables represent important underlying factors explaining the observed differences in competitiveness between the Spanish regions or the observed inequality in regional income levels within Spain (see the contributions in Cuadrado-Roura 2010) and are similar to those used in the specification of econometric models of regional growth in the most recent literature (see for example, the MASST model of Capello et al. 2008).[7] The national variable was used to make some allowance for the impact of common macro shocks on SpVAR estimates. All the local covariates variables were introduced in lagged form to prevent simultaneity problems that would arise if current values of the conditioning variables were included.

In the application, a maximum of two lags ($p = 2$) is considered for each regional SpVAR model, with the optimal lag being determined by the standard

[6] Under different functional form assumptions (basically, icebergs costs and CES preferences; see, for example, Harrigan 2003 or Combes et al. 2008), it could be possible to generate a regional gravity equation where the share of regional income for a region is a function of an income-weighted average of the rest of regions. The production side of this gravity equation was provided by the monopolistic competition model (Helpman and Krugman 1985).

[7] The temporal limitations of the Spanish regional databases prevented us from using information about other important variables, as human capital or infrastructure indicators.

Table 2 The conditioning variables in the SpVAR models

Local indicators	National indicator
Lagged per capita income	Current national income
Lagged population	
Lagged employment rate	
Lagged share of primary production	

statistical information criteria. In all cases, AIC and BIC statistics point to a first-order SpVAR(1,1) model for each individual region.

Finally, in reference to the choice of the spatial weighting scheme, the spatial weights w_{ij} were defined in order to capture the *economic* importance of region j for the *ith* region's economy. Following Pesaran et al. (2004) and Dees et al. (2007), in this application the geographical patterns of trade provided the main source of information for this purpose, because trade is expected to have an important effect on regional growth and bilateral trade also is expected to be one of the most important sources of inter-regional business cycle linkages.

Specifically, fixed trade weights based on the average regional trade flows computed over the 2002–2007 period were used to specify the channels of transmission of local shocks across the Spanish regional system. Table 3 presents these trade weights for the 17 Spanish regions (trade shares are displayed in row-normalized form by region, such that each row sums to one), which were computed as shares of region j in the total trade (exports plus imports) of region i, measured in millions of euros. This matrix shows the degree of integration between the different regions, and highlights the key role played by six regions: Andalusia, Castile and León, Catalonia, Valencian Community, Madrid and the Basque Country. These regions are the more integrated with the rest of Spanish regional economies.

After the description of the econometric models to be estimated, they will be used to address the following question: are there significant spillover effects within the Spanish regional system? Through the estimation of the proposed model in Sect. 4, some of the empirical links between regional economic performance and the forces driving the regional economic performance along the time will be revealed; in the process, it will be possible to uncover the spatial robustness of regional sizes to shocks in the trade-neighborhood.

Figure 7 presents the estimated impulse responses of regional production shares to a shock in the share of trade-neighboring regions. A numerical synthesis is shown in Table 4. The estimates of this table provide the base year 0 (short-run) and the accumulated year 25 (long-run) percentage changes in shares for a one-percentage-point change in the share of neighboring regions.[8] One or two asterisks marking an estimate in the table indicates that the corresponding confidence interval does not include zero.

[8] To facilitate the interpretation of the impulse responses, the endogenous variables of the SpVAR models (first differences of local and external shares) have been multiplied by 100, so that the accumulated impulse responses provide the percentage change in the level of the respective variable.

Table 3 Spatial weights (w_{ij}) based on the average regional trade flows

	AND	ARA	AST	BAL	CAN	CANT	CYL	CLM	CAT	VAL	EXT	GAL	MAD	MUR	NAV	PV	RIO
AND		0.03	0.03	0.02	0.05	0.01	0.06	0.10	0.16	0.11	0.08	0.04	0.17	0.08	0.01	0.04	0.01
ARA	0.05		0.01	0.00	0.01	0.01	0.05	0.04	0.41	0.10	0.00	0.03	0.08	0.01	0.07	0.10	0.03
AST	0.10	0.03		0.00	0.01	0.06	0.16	0.03	0.08	0.06	0.01	0.20	0.10	0.01	0.02	0.12	0.01
BAL	0.15	0.00	0.01		0.05	0.00	0.00	0.00	0.45	0.22	0.00	0.04	0.05	0.02	0.00	0.01	0.00
CAN	0.20	0.02	0.01	0.03		0.01	0.04	0.02	0.24	0.09	0.00	0.08	0.18	0.01	0.01	0.06	0.00
CANT	0.04	0.03	0.08	0.00	0.02		0.17	0.04	0.15	0.04	0.01	0.05	0.08	0.01	0.03	0.23	0.02
CYL	0.07	0.04	0.05	0.00	0.01	0.04		0.06	0.13	0.06	0.03	0.11	0.17	0.01	0.03	0.17	0.02
CLM	0.15	0.03	0.01	0.00	0.01	0.01	0.07		0.10	0.16	0.04	0.03	0.29	0.05	0.01	0.03	0.01
CAT	0.10	0.16	0.02	0.04	0.04	0.02	0.07	0.05		0.16	0.01	0.04	0.13	0.03	0.04	0.08	0.01
VAL	0.11	0.06	0.02	0.03	0.02	0.01	0.05	0.10	0.25		0.01	0.03	0.13	0.13	0.02	0.04	0.01
EXT	0.41	0.02	0.01	0.00	0.01	0.01	0.11	0.13	0.05	0.04		0.02	0.15	0.01	0.01	0.03	0.01
GAL	0.08	0.03	0.12	0.01	0.04	0.02	0.19	0.04	0.14	0.06	0.01		0.14	0.02	0.02	0.11	0.01
MAD	0.13	0.04	0.02	0.01	0.03	0.01	0.11	0.16	0.17	0.10	0.02	0.05		0.02	0.02	0.07	0.01
MUR	0.21	0.02	0.01	0.01	0.01	0.01	0.03	0.09	0.12	0.35	0.01	0.02	0.08		0.01	0.02	0.01
NAV	0.03	0.13	0.02	0.00	0.01	0.02	0.07	0.03	0.19	0.05	0.00	0.02	0.07	0.01		0.29	0.06
PV	0.05	0.08	0.04	0.00	0.02	0.06	0.17	0.02	0.15	0.06	0.01	0.06	0.11	0.01	0.12		0.04
RIO	0.06	0.11	0.03	0.00	0.01	0.02	0.10	0.03	0.14	0.08	0.01	0.02	0.06	0.02	0.13	0.19	

Note: (1) Trade weights are computed as shares of exports and imports over the 2002–2007 period, displayed in rows by region. (2) Regional abbreviations: Andalusia (AND), Aragón (ARA), Asturias (AST), Balearic Islands (BAL), Canary Islands (CAN), Cantabria (CANT), Castile and León (CYL), Castile-La Mancha (CLM), Catalonia (CAT), Valencian Community (VAL) Extremadura (EXT), Galicia (GAL), Madrid (MAD), Murcia (MUR), Navarre (NAV), Basque Country (PV), and La Rioja (RIO)

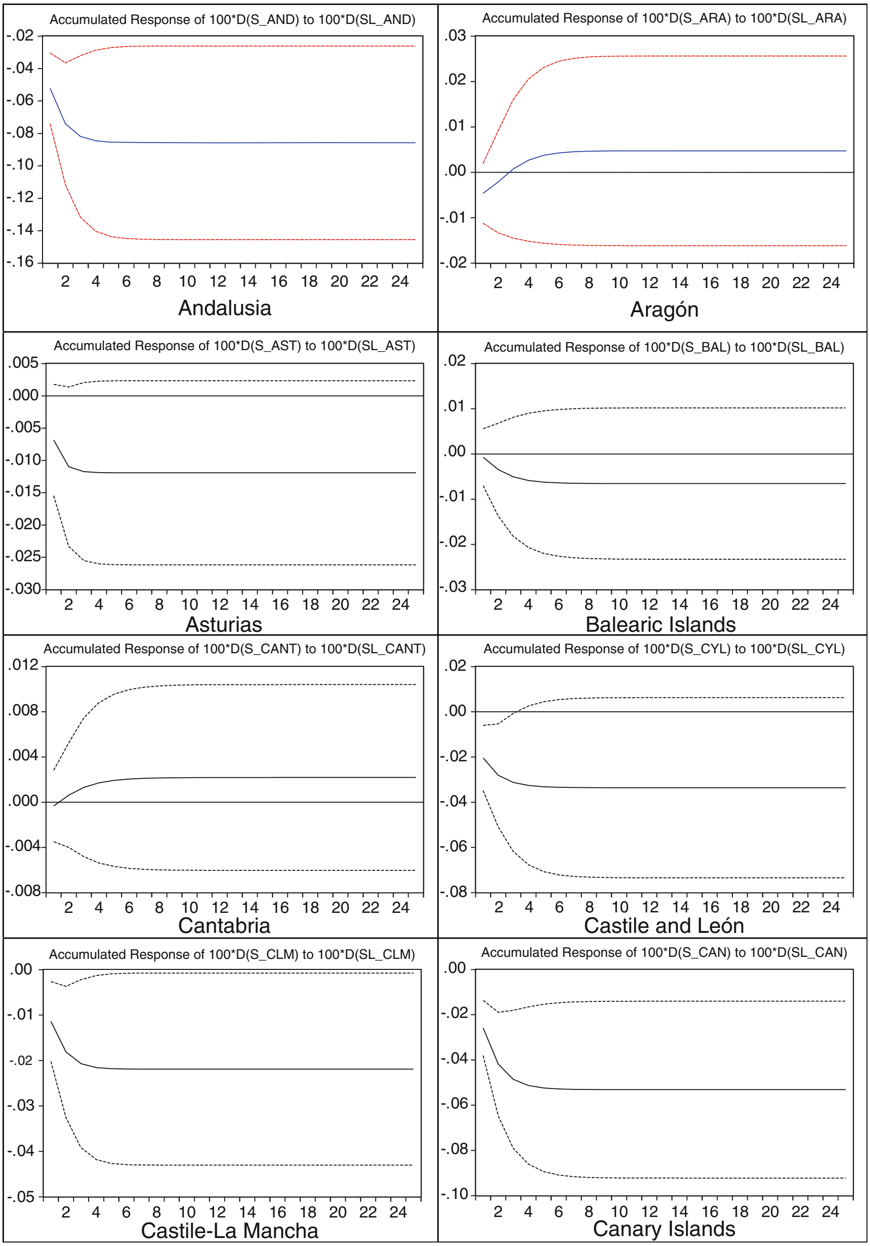


Fig. 7 (continued)

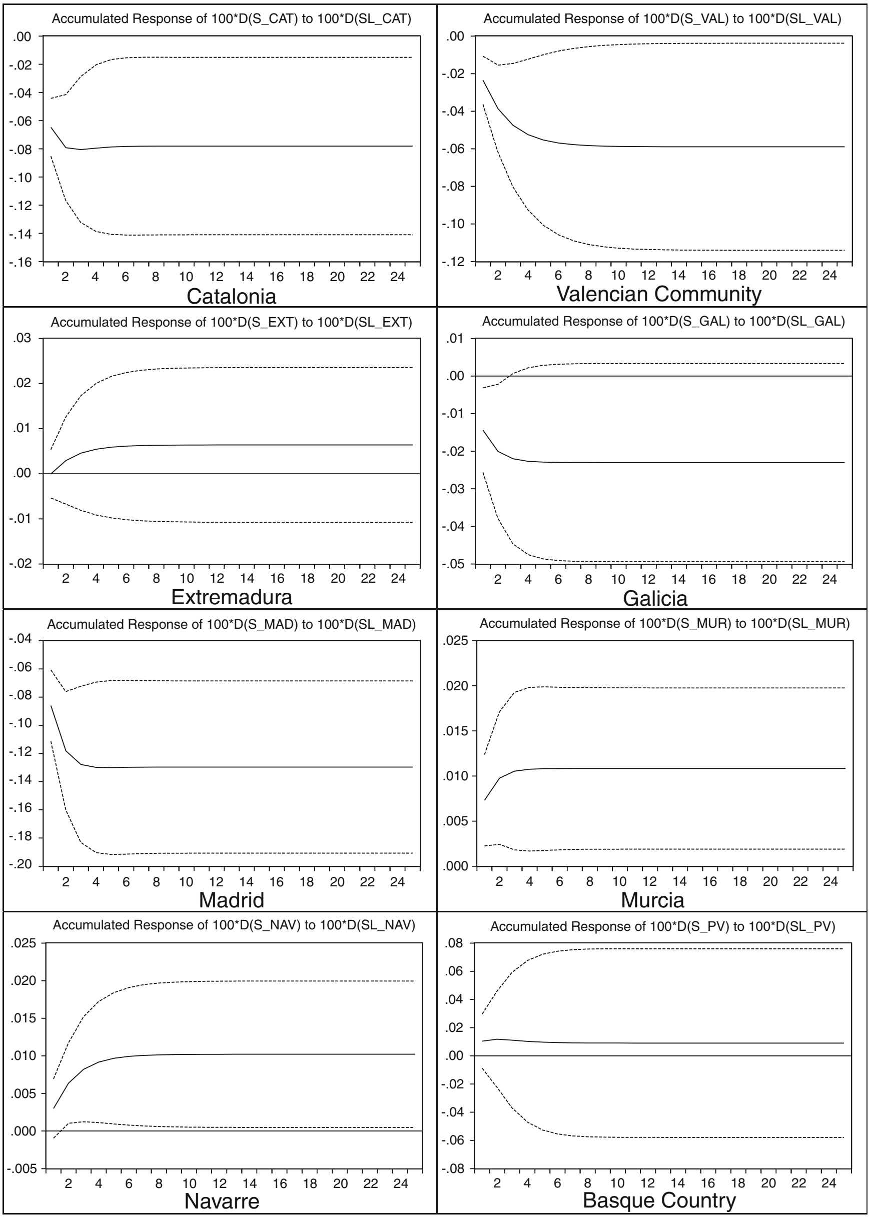


Fig. 7 (continued)

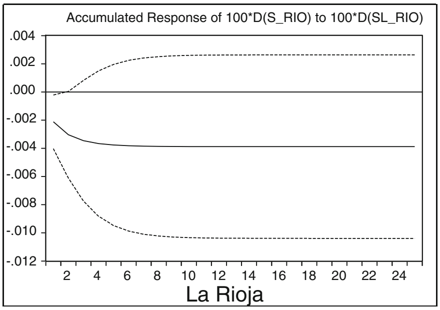


Fig. 7 Estimated impulse responses of production shares to a shock in the share in trade-neighboring areas (Note: Graphs show the accumulated responses to generalized one standard deviation innovations. *Dotted lines* represent ± 2 standard error confidence bands)

Table 4 Trade-related neighbors' spillover elasticities

	Short-run elasticities	Long-run elasticities
Andalusia	−3.65**	−4.30**
Aragón	−0.14*	0.09
Asturias	−0.63*	−0.86*
Balearic Islands	−0.02	−0.15
Canary Islands	−1.21**	−2.23**
Cantabria	−0.02	−0.04
Castile and León	−1.00**	−0.92*
Castile-La Mancha	−0.42**	−0.58**
Catalonia	−5.24**	−4.51**
Valencian Comm.	−1.36**	−2.09**
Extremadura	0.00	0.18
Galicia	−0.99**	−1.05*
Madrid	−6.11**	−6.72**
Murcia	0.36**	0.38**
Navarre	0.13*	0.20**
Basque Country	0.85*	0.48
La Rioja	−0.14**	−0.12*

Note: Superscripts * (**) denotes that the corresponding 68 % (95 %) percentile confidence interval does not include zero. The confidence intervals for individual regions were computed by adding and subtracting one or two standard errors (Sims and Zha 1999)

The analysis of data reveals that, in general, all the regions have negative elasticities both in the short-run and long-run (see Fig. 8). Murcia and Navarre are regions that present significant positive elasticities both in the short-run and long-run: an increase in the relative share of their respective trade-neighboring regions would cause positive accumulated responses in Murcia and Navarre. On the other hand, Balearic Islands, Cantabria and Extremadura are regions that present no significant elasticities both in the short-run and long-run; these regions do not show

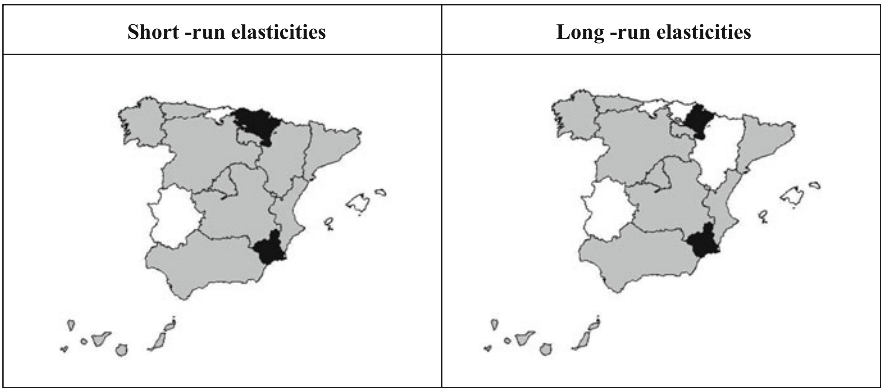


Fig. 8 Trade-related neighbors' spillover elasticities (Note: *White color* indicates a non-significant elasticity, *gray color* indicates a significant negative effect, *black color* indicates a significant positive effect)

significant responses to shocks generated by their respective trade-neighboring regions. Hence, it is possible to conclude that, to the extent that a shock in the trade-neighboring regions occurs, a relative spatial inequality emerges.

The existence of regional competition is not a new result. For example, Márquez and Hewings (2003) showed for the case of the Spanish peninsular regions that, in general, the regions that are gaining increments of share over time are producing negative influences on their geographical neighbors. Similarly, the study of the intraregional dynamics of the Spanish regional system by Márquez et al. (2006) suggests that regional imbalances in the Spanish regional system are strongly influenced by the macro-effects that are operating within this system. The authors found the general view that, in terms of the whole Spanish system, peripheral regions are not increasing their relative productive capacities; in other words, peripheral regions reveal negative economy-wide effects.

What is new from the current findings is the detection of negative effects related to the existence of stable trade connections between regions. Thus, regional competition emerges and the dynamics of regional disparities related to trade will provide new insights using this proposed approach. In particular, the results may suggest the presence of some long term fixed effects that may signify some form of path dependence that may help explain some of the findings from the convergence analyses that have been presented for Spain.

5 Summary and Conclusions

Regional competition is the result of a dynamic process in which the analysis of space-time interactions among regions can offer valuable information about regional economic performance. In this chapter, a series of approaches is presented that reveal the complexity of the problem and the different insights that can be gained by adopting alternative methodologies. Following the parameterization of the dynamic behavior of the Spanish regional system, this chapter identifies the response of shocks to the relative size of regions.

A spatial vector autoregressive (SpVAR) model for the Spanish regions for the period 1955–2009 is used. The SpVAR model considers spatial as well temporal lags of the variables. The estimated SpVAR is used to calculate impulse responses that provide insights about the effects of shocks in relative productive capacity in different locations. The empirical results suggest that relative regional sizes have different, significant impacts on the relative productive capacity for the 17 Spanish regions interpreted through the existing trade linkages. There is evidence that, in general, regional economies are in competition with respect to their trade-neighboring regions (regions that have trade relationships with the reference region). In general, the results point to a negative relationship between regional size and the growth of the size of its corresponding trade-neighborhood.

With growing economic integration across regions, one would expect economic growth in a region to be influenced by growth in other regions. The research here suggests that trading partners do indeed matter significantly for growth. Regional competition is operating within the Spanish regional economic system; the detected responses imply, in general, negative growth spillovers from the trade-neighborhood. A logical next step for this type of analysis would be the development of regional business cycle models (for a review, see Hewings and Vera 2011) to capture the interregional dynamics with higher frequency data (e.g. quarterly or monthly), with these models incorporating some of the insights from the spatial vector autoregressive models adopted in this chapter.

Acknowledgements The authors acknowledge and appreciate the funding received from the Ministry of Science and Innovation of Spain through the project ECO2009-12506

Appendix

This appendix includes a brief discussion about multivariate spatial vector autoregressive –SpVAR- models, containing both the general formulation of such models and an explanation on the main properties of this type of econometric specifications.

Spatial VARs are a special type of vector autoregressions (Sims 1980), which include spatial as well as temporal lags of the state variables. Contrary to standard VARs, that do not allows the joint modeling of dynamic spatio-temporal interdependencies within a group of connected local economies (regions or states, metropolitan areas or local districts), SpVAR models permit that endogenous variables can exhibit co-movements over time and also over space.

Let $Y_{it} = (y_{1,it}, y_{2,it}, \ldots, y_{G,it})$ and $X_{it} = (x_{1,it}, x_{2,it}, \ldots, x_{K,it})$ denote two $G \times 1$ and $K \times 1$ vectors of stationary endogenous and exogenous variables, respectively, recorded on the ith region ($i = 1,2,\ldots,N$) at time t ($t = 1,2,\ldots,T$). A reduced-form SpVAR specification of order p for these data can be written as

$$Z_{it} = \Gamma_{0i} + \Gamma_{1i}Z_{i,t-1} + \ldots + \Gamma_{pi}Z_{i,t-p} + \Phi_{0i}t + \Phi_{1i}X_{i,t} + U_{it}$$

where $Z_{it} = (Y_{it}, Y_{it}^*)$; $\Gamma_{ji}\,(j = 0,1,\ldots,p)$ and $\Phi_{ji}\,(j = 0,1)$ are matrices of coefficients to be estimated; U_{it} is a $2G \times 1$ vector of non-autocorrelated reduced-form disturbances with mean zero and a nonsingular covariance matrix, Σ_i; and $Y_{it}^* = (I_G \otimes W)Y_t$, with $Y_t = (Y_{1t}, Y_{2t}, \ldots, Y_{Nt})$ and W being a row-standardized $N \times N$ connectivity matrix with elements w_{ij} fixed over time satisfying $w_{ii} = 0$ and $\sum_{j=1}^{N} w_{ij} = 1$. The components of the spatial lagged dependent vector Y_{it}^* can be written as $y_{g,it}^* = \sum_{j=1}^{N} w_{ij}y_{g,jt}$ and would be the weighted average of y_g in region i for regions s where $w_{is} \neq 0$.

It can be seen that spatially heterogeneous model dynamics is allowed because parameters are assumed to vary unrestrictedly at the level of the individual regions. Also, it can be observed that contemporaneous relations across the states variables are not modeled explicitly but captured by the elements of the covariance matrix Σ_i.

Written in disaggregated form, the SpVAR model for region i takes the following form:

$$\begin{cases} Y_{it} = \Gamma_{0i}^1 + \Gamma_{1i}^1 Y_{i,t-1} + \Gamma_{2i}^1 Y_{i,t-1}^* + \ldots + \Phi_{0i}^1 t + \Phi_{1i}^1 X_{it} + U_{it}^1 \\ Y_{it}^* = \Gamma_{0i}^2 + \Gamma_{1i}^2 Y_{i,t-1} + \Gamma_{2i}^2 Y_{i,t-1}^* + \ldots + \Phi_{0i}^2 t + \Phi_{1i}^2 X_{it} + U_{it}^2 \end{cases}$$

This expression implies that a spatial VAR can be seen as a spatial-extended VAR model for the vector Y_{it}. For each endogenous variable y_m of this vector ($m = 1,2,\ldots,G$), the SpVAR with N regions takes the form of the following system of equations (similar specification exists for the external variables y_m^*):

$$\begin{cases} y_{m,1t} = \varphi_{01,m}^1 + \sum_{g=1}^{G} \varphi_{11,m,g}^1 y_{g,1,t-1} + \sum_{g=1}^{G} \varphi_{21,m,g}^1 y_{g,1,t-1}^* + \ldots + \phi_{01,m}^1 t \\ \qquad + \sum_{k=1}^{K} \phi_{11,m,k}^1 x_{k,1t} + u_{m,1t}^1 \\ y_{m,2t} = \varphi_{02,m}^1 + \sum_{g=1}^{G} \varphi_{12,m,g}^1 y_{g,2,t-1} + \sum_{g=1}^{G} \varphi_{22,m,g}^1 y_{g,2,t-1}^* + \ldots + \phi_{02,m}^1 t \\ \qquad + \sum_{k=1}^{K} \phi_{12,m,k}^1 x_{k,2t} + u_{m,2t}^1 \\ \qquad \vdots \\ y_{m,Nt} = \varphi_{0N,m}^1 + \sum_{g=1}^{G} \varphi_{1N,m,g}^1 y_{g,N,t-1} + \sum_{g=1}^{G} \varphi_{2N,m,g}^1 y_{g,N,t-1}^* + \ldots + \phi_{0N,m}^1 t \\ \qquad + \sum_{k=1}^{K} \phi_{1N,m,k}^1 x_{k,Nt} + u_{m,Nt}^1 \end{cases}$$

These reduced-form equations include the deterministic and exogenous variables, a set of temporally lagged variables (as in the traditional VARs), and a set of new temporally lagged spatial-lag variables.

SpVARs can be used to simulate the spatio-temporal dynamic effects of exogenous shocks within the system. Now, the impulse response analysis is more general than in the traditional VARs: an exogenous shock that occurs in a given region (or a group of them) at a point time can affect the economic conditions of other regions in the next periods. Therefore, shocks can propagate over time as well as across space, permitting the existence of spatial spillover effects: an exogenous shock in a region can spill over to the locations considered as neighbours ($w_{is} \neq 0$).

References

Aiginger K, Davies SW (2004) Industrial specialization and geographic concentration: two sides of the same coin? Not for the European Union. J Appl Econ 7:231–248

Armstrong H, Taylor J (2000) Regional economics and policy. Blackwell, Oxford

Arora V, Vamvakidis A (2005) Economic spillovers. Exploring the impact trading partners have on each other's growth. Finance Develop 42:48–50

Barro RJ (1991) Economic Growth in a Cross-section of Countries. Quarterly Journal of Economics 106(2):407–443

Baldwin RE, Forslid R, Martin P, Ottaviano GIP, Robert-Nicoud F (2003) Economic geography and public policy. Princeton University Press, Princeton

Barro R, Sala-i-Martin X (1991) Convergence across states and regions. Brook Pap Econ Act 2:107–182

Batey P, Friedrich PP (2000a) Aspects of regional competition. In: Batey PWJ, Friedrich P (eds) Regional competition. Springer, Berlin/Heidelberg/New York, pp 3–33

Batey PWJ, Friedrich P (eds) (2000b) Regional competition. Springer, Heidelberg

Brakman S, Garretsen H, Schramm M (2004) The strategic bombing of German cities during WWII and its impact on city growth. J Econ Geogr 4:201–217

Capello R (2009a) Space, growth and development. In: Capello R, Nijkamp P (eds) Handbook of regional growth and development theories. Edward Elgar, Cheltenham, pp 33–52

Capello R (2009b) Spatial spillovers and regional growth: a cognitive approach. Eur Plann Stud 17:639–658

Capello R, Camagni R, Chizzolini B, Fratesi U (2008) Modelling regional scenarios for the enlarged Europe. Springer, Berlin/Heidelberg

Cheshire P, Carbonaro G (1996) Urban economic growth in Europe. Urban Stud 33:1111–1128

Cheshire P, Magrini S (2009) Growing urban GDP or attracting people? Different causes, different consequences. In: Karlsson C, Andersson AE, Stough RR, Cheshire PC (eds) New directions in regional economic development, advances in spatial science. Springer, Berlin, pp 291–315

Combes PP, Mayer T, Thisse JF (2008) Economic geography: the integration of regions and nations. Princeton University Press, New Jersey

Cuadrado-Roura JR (2001) Regional convergence in the European Union from hypothesis to the actual trends. Annal Reg Sci 35:333–356

Cuadrado-Roura JR (ed) (2010) Regional policy, economic growth and convergence. Springer, Berlin/Heidelberg

Davis DR, Weinstein D (1999) Economic geography and regional production structure: an empirical investigation. Eur Econ Rev 43:379–407

Davis DR, Weinstein D (2002) Bones, bombs, and break points: the geography of economic activity. Am Econ Rev 92:1269–1289

Davis DR, Weinstein D (2003) Market access, economic geography and comparative advantage: an empirical assessment. J Int Eco 59:1–23

De Bustos A, Cutanda A, Díaz A, Escribá FJ, Murgui MJ, Sanz MJ (2008) La BD.MORES en base 2000: Nuevas estimaciones y variables, D-2008–2002, Dirección General de Presupuestos, Ministerio de Economía y Hacienda

De la Fuente A (2010) Series anuales de algunos agregados económicos y demográficos regionales, 1955–2009, D-2010–2005, Dirección General de Presupuestos, Ministerio de Economía y Hacienda

Dees S, Di Mauro F, Pesaran MH, Smith LV (2007) Exploring the international linkages of the euro area: a global VAR analysis. J Appl Econometr 22:1–38

Egger P, Pfaffermayr M (2006) Spatial convergence. Pap Reg Sci 85:199–215

Fagiolo G (2010) The international-trade network: gravity equations and topological properties. J Econ Int Coordin 5:1–25

Feser EJ (1998) Enterprise, external economies and economic development. J Plann Liter 12:283–302

Friedrich P (1987) Regional competition under stagnation. In: Friedrich P, van Rompuy P (eds) Fiscal decentralization. Baden-Baden, Nomos, pp 84–133

Fujita M, Thisse JF (2002) Economics of agglomeration: cities, industrial location, and regional growth. Cambridge University Press, Cambridge

Fujita M, Krugman P, Venables AJ (1999) The spatial economy. The MIT Press, Cambridge

Garlaschelli D, Loffredo M (2005) Structure and evolution of the world trade network. Physica A 355:138–144

Garrido-Yserte R, Mancha-Navarro T (2010) The Spanish regional puzzle: convergence, divergence and structural change. In: Regional policy, economic growth and convergence. Springer, Berlin/Heidelberg

Glaeser EL, Scheinkman JA, Shleifer A (1995) Economic growth in a cross-section of cities. J Monet Econ 36:117–143

Haaparanta P (1998) Regional concentration, trade, and welfare. Reg Sci Urban Econ 28:445–463

Harrigan J (2003) Specialization and the volume of trade: do the data obey the laws? In: Kwan Choi E, Harrigan J (eds) Handbook of international trade. Blackwell Publishing, Malden, Massachussetts, pp 85–118

Helpman E, Krugman PR (1985) Market structure and foreign trade. MIT Press, Cambridge, MA

Hewings JD, Vera R (2011) "The Economic Impact of the Recession on Metropolitan Areas". Unpublished paper, Regional Economics Applications Laboratory, University of Illinois at Urbana-Champaign, USA

Hewings GJD (2008) On some conundra in regional science. Annal Reg Sci 42:251–265

Hewings G, Sonis M, Cuello F, Mansouri F (1996) The role of regional interaction in regional growth competition and complementarity in the U.S. regional system. Aust J Reg Stud 2:133–149

Imbs J, Wacziarg R (2003) Stages of diversification. Am Econ Rev 93:63–86

Kali R, Reyes J (2007) The architecture of globalization: a network approach to international economic integration. J Int Bus Stud 38:595–620

Kancs D (2007) Does economic integration affect the structure of industries? Empirical evidence from the CEE. Econ Syst Res 19:73–97

Krugman P (1991) Geography and trade. MIT Press, Cambridge, MA

Krugman PR (1980) Scale economies, product differentiation, and the pattern of trade. Am Econ Rev 70:950–959

Krugman PR, Venables AJ (1995) Globalization and the inequality of nations. Quart J Econ 110:857–880

Llano C (2004) Economía espacial y sectorial: el comercio interregional en el contexto de un modelo Multirregional para la economía española. Instituto de Estudios Fiscales. Ministerio de Economía y Hacienda

Llano C, Esteban A, Pérez J, Pulido A (2010) Opening the interregional trade “Black Box”: the C-intereg database for the Spanish economy (1995–2005). Int Reg Sci Rev 33:302–337
López-Bazo E, Esther Vayá E, Mora AJ, Suriñach J (1999) Regional economic dynamics and convergence in the European Union. Annal Reg Sci 33:343–370
López-Rodríguez J, Faiña A (2007) Regional wage disparities in the European Union: what role for market access. Investig Reg 11:5–23
Márquez MA, Ramajo J, Hewings GJD (2006) Dynamic effects within a regional system: an empirical approach. Environ Plann A 38:711–732
Márquez MA, Ramajo J, Hewings GJD (2003) Regional interconnections and growth dynamics: the Spanish case. Aust J Reg Stud 9:5–28
Márquez MA, Ramajo J, Hewings GJD (2010) A spatio-temporal econometric model of regional growth in Spain. J Geogr Syst 12:207–226
Márquez MA, Hewings GJD (2003) Geographical competition between regional economies: the case of Spain. Annal Reg Sci 37:559–580
Marshall A (1890) Principles of economics. Macmillan, London
Pérez J, Dones M, Llano C (2008) An interregional impact analysis of the EU structural funds in Spain (1995–1999). Pap Reg Sci 88:509–529
Pesaran MH, Schuermann T, Weiner SM (2004) Modelling regional interdependencies using a global error-correcting macroeconometric model. J Bus Econ Stat 22:129–162
Phillips PCB, Moon H (1999) Linear regression limit theory for nonstationary panel data. Econometrica 67:1011–1057
Poot J (2000) Reflection on local and economy-wide effects of territorial competition. In: Batey PWJ, Friedrich P (eds) Regional competition. Springer, Berlin/Heidelberg/New York, pp 205–230
Quah DT (1997) Empirics for growth and distribution: stratification, polarization, and convergence clubs. J Econ Growth 2:27–59
Quah DT (1996) Empirics for economic growth and convergence. Eur Econ Rev 40:1353–1375
Ramajo J, Márquez MA, Hewings GJD, Salinas MM (2008) Spatial heterogeneity and interregional spillovers in the European Union: do cohesion policies encourage convergence across regions? Eur EconRev 52:551–567
Ramos R, Suriñach J, Artís M (2010) Human capital spillovers, productivity and regional convergence in Spain. Pap Reg Sci 89:435–447
Redding S, Venables AJ (2004) Economic geography and international inequality. J Int Econ 62:53–82
Richardson HW (1973) Regional Growth Theory. New York: John Wiley & Sons
Roy JR, Thill J-C (2004) Spatial interaction modelling. Pap Reg Sci 83:339–361
Sala-i-Martin X (1996) The classical approach to convergence analysis. Econ J 106:1019–1036
Silverman B W (1986) Density Estimation for Statistics and Data Analysis. London: Chapman & Hall
Serrano A, Boguña M (2003) Topology of the world trade web. Phys Rev E 68:015101(R)
Sims CA (1980) Macroeconomics and reality. Econometrica 48:1–48
Sims CA, Zha T (1999) Error bands for impulse responses. Econometrica 67:1113–1155
Tinbergen J (1962) Shaping the world economy: suggestions for an international economic policy. Twentieth Century Fund, New York

The Long Run Interplay Between Trade Policy and the Location of Economic Activity in Brazil Revisited

L.P. do C. Ferraz, E.A. Haddad, and M.C. Terra

Abstract This paper addresses the question of whether further trade liberalization in Brazil may exert some influence on its already heterogeneous economic landscape. This is a rather relevant issue for the Brazilian economy, still one of the closest economies in the world, with huge economic disparities among its regions. The central question was motivated by an influent (and also controversial) theoretical insight brought about by Krugman and Elizondo (J Develop Econ 49: 137–150, 1996). Trough the specification of an interregional CGE model for the Brazilian economy, this paper argues that the adoption of (horizontal) liberal trade policies in Brazil, beyond traditional gains from trade, can also contribute to ameliorate regional inequality in the country, a result quite in line with Krugman and Elizondo's predictions. In this sense, it makes the case for trade liberalization in Brazil as an additional (horizontal) public policy apparently effective in fighting regional inequality trough traditional market forces.

Keywords Trade liberalization • Spatial general equilibrium • CGE models • Latin America • New Economic Geography

---

JEL: R13, F17, D58

L.P. do C. Ferraz
EESP/FGV, Brazil

E.A. Haddad (✉)
Department of Economics, University of São Paulo, Brazil
e-mail: ehaddad@usp.br

M.C. Terra
THEMA/Université de Cergy-Pontoise, France, and EPGE/FGV, Brazil

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_16,

1 Introduction

Despite policy maker's intense and persistent interest in the location of economic activity, economist's interest in the subject has surprisingly waxed and waned over the last century. However, over the last two decades it seems there has been a renaissance of the field in mainstream economics, starting with the seminal work of Paul Krugman (1991a), which has been considered the backbone of the so called "New Economic Geography" (NEG for short).

Perhaps the main contribution of the so called "Core-Periphery" model (Krugman 1991a) to economic theory was to open the black box of agglomeration economies. According to Krugman's Core-Periphery model (CP for short), the potential for circular causality – or backward and forward linkages – is created when the market-access effect and the cost-of-living effect are simultaneously combined with labor mobility.

The new insights provided by the NEG also helped to clarify some complexes urbanization patterns, as the way it can vary considerably from one region to another inside many countries. For instance, it has been demonstrated (Krugman 1993; Fujita and Krugman 1995; Fujita and Mori 1995) that primate cities can emerge in a homogeneous geographical space purely endogenously through agglomeration forces generated by the interaction between increasing returns, transport costs and factor mobility.

A better understanding of the urbanization patterns in developing countries is particularly relevant, since the growth of very large agglomerations such as Mexico City and São Paulo is widely regarded as a social problem. Other cities like Montevideo, Buenos Aires and Santiago, just to mention a few, concentrate over 40 % of its country's population. Associated with metropolitan growth is an obviously related concern about regional inequality.

An influent attempt to model the formation of large agglomerations in the developing world, where the key features of the new economic geography are present, can be found in Krugman and Elizondo (1996). The authors suggest that the rise of Third World metropolises may be an unintended by-product of import-substitution policies carried out in the past, and those metropolises will tend to shrink as developing countries liberalize. Accordingly, for an inward-looking process of industrialization, such as the ones carried out in Mexico and some other countries in Latin America, backward and forward linkages (agglomeration forces, mainly scale economies) may overcome dispersion forces such as high rents and commuting costs, wages, congestion and pollution. In that case, the rise of giant metropolis may be a stable equilibrium. However, when the economy opens up to foreign trade, the balance between agglomeration and dispersion forces may be reversed. Both backward and forward linkages are weakened, since domestic firms can now serve the foreign markets and domestic consumers are able to import products from abroad. Meanwhile, the strength of the dispersion forces remains just as large as before. Hence, the only stable equilibrium is with dispersed consumption and production.

The work by Krugman and Elizondo inspired a new strand of papers in the field, some of them confirming their basic conclusions, some others contradicting it. Perhaps the two most relevant critics to Krugman and Elizondo's results came from Henderson (1996) and Isserman (1996). For Henderson the impact of trade on national space may be not so clear-cut as suggested by Krugman and Elizondo's three location model. Instead, it will depend on the precise geography of the country. According to him, most countries, even small ones, have dozens if not hundreds of cities. Therefore, "a more general framework is needed if urban concentration must be addressed in all its complex dimensions".

For Isserman (1996), even if free-trade policies result in a greater proportion of domestic production being sold abroad, there remains a significant part of its economic base that is still dependent on the primate population concentration. For instance, as a result of trade liberalization, primate city functions such as finance, government, trade, and communication, might expand and reinforce its dominant position. Also, for Krugman's dispersion effect to take action, primate city manufacturers must be able to compete in the international markets and residents must be able to afford the goods produced by foreign producers. To the extent that local producers are unable to compete in the world market and local consumers do not have good access to foreign products, Krugman's posited trade effect will not be strong enough for the primate city production to disperse to other sites. Therefore, "the model may yield erroneous conclusions regarding the overall effects of free trade on population concentration".

In the same spirit of Isserman's critics, Alonso-Villar (2001) suggests that megacities may not be only the consequence of protective trade policies but also of the technological backwardness of a country relatively to international markets. This is what the author calls "the competition effect". Once it is taken into consideration, it might well be the case that under free-trade policies, domestic firms are unable to compete with more technological advanced firms in international markets. In that case, free-trade policies will not lead to dispersion, since domestic firms will keep its dependence on primate city markets.

Despite the challenging theoretical issues raised by Krugman's NEG, little empirical work has been done in order to corroborate its main theoretical insights. The work by Ades and Glaeser (1995) provide some mild empirical support to Krugman's hypothesis on the relationship between trade openness and urban concentration. According to their results, an increase in 10 % in the share of trade in GDP leads to a reduction of 6 % on a country's primate city size. Also, a 1 % increase in the share of GDP spent on government transportation and communications reduces a country's primate city size by 10 %. The first evidence supports Krugman and Elizondo's hypothesis on the relationship between import-substitution policies and urban concentration. The second one gives support to Krugman (1991a): high internal transportation costs create an incentive for urban concentration in space. The work by Sanguinetti and Martincus (2005) also provide some empirical support for Krugman and Elizondo (1996): according to the authors, free-trade policies in Argentina were responsible for industry relocation away from Buenos Aires in the 1990s. More recently, Haddad et al. (2009) also find some empirical support for the positive correlation between free-trade policies and industry relocation in Colombia, specifically to the case of the trade-liberalization process that occurred in the 1990s. Moreover, the dispersion effects stemming from the adoption of liberal trade policies can be triggered by a change in relative prices due to the imposition of spatially asymmetric environmental standards, as suggested by Altomonte and Bosco (2011) in the case of European energy intensive sectors, such as the cement industry.

The debate over trade liberalization in Brazil has hardly ever tackled the issue of spatiality. Instead, it has often been concerned with possible sectoral implications and/or general economy-wide effects of liberal trade policies, treating the regions as isolated and aspatial entities (Flores 1997; Campos-Filho 1998; Gonzaga et al. 1999).

A better understanding of the economic forces behind urban concentration and regional inequality is especially important in the case of Brazil, an economy where just the state of São Paulo concentrates over 40 % of the country's manufacturing activity. Due to its continental dimensions and a remarkable history of import-substitution policies, it might well be the case that high transportation costs associated with the adoption of protective trade polices contributed to reinforce pre-existing regional inequalities in Brazil. By the same token, assuming Krugman and Elizondo's predictions are correct, the adoption of Washington consensus policies in the 1990s should have led to relocation of economic activity away from the richer south and southeast regions towards less developed regions such as north and northeast regions. In this regard, both the works by Haddad and Hewings (1999) and Haddad and Azzoni (2001) would suggest that trade liberalization policies in the 1990s may have led to the enhancement of pre-existing interregional inequalities in Brazil from a macro spatial perspective. Using a general equilibrium macro-regional model, with a data base from the year 1985, the authors point out that liberal import tariff policies in Brazil are more prone to benefit the most developed region in the country (which they call "center-south", comprising the south, southeast and center-west regions), leaving behind the less developed regions (north and northeast regions).

Despite the significant import tariff reduction carried out in the beginning of the 1990s, it seems little has changed in terms of regional inequality in Brazil. The real causes behind this apparently rigidity in regional inequality are still an open question. However, as pointed out by Ferraz and Haddad (2009), the evolution of import penetration in Brazil has been also historically disappointing, notwithstanding the huge cut in import tariffs occurred in the 1990s.

With its continental dimensions, poor infrastructure, the existence of huge income disparities among its 27 states and a solid industrial base mostly concentrated in the country's richest regions, the Brazilian economy seems to fit quite well in the paradigm of the NEG models. However, any simplified theoretical version of this rather heterogeneous multi-region economy would certainly loose many of its most relevant aspects and would probably lead to erroneous conclusions in terms of the role played by the key NEG forces behind urban concentration and regional inequality.

- Methodological Aspects

The model used in this paper is a natural adaptation of the B-MARIA model (Haddad and Hewings (2005)) to include the possibility of differentiated trade policy shocks in each of the 27 Brazilian states (see Appendix 1). It is noteworthy that the original structure of the class of B-MARIA models adapted to a real economy fully contemplates Isserman (1996) and Henderson (1996) critics on the usual simplifications of theoretical NEG models applied to multi-regional systems analysis. Since economies of scale, factor mobility and regional distances from internal and external markets are accommodated, simulation results can shed some light on the balance between agglomeration and dispersion forces, and its contribution to urban concentration and regional inequality, once liberal trade policies are carried out.

The contribution of this paper is twofold: by applying a interstate CGE model to the case of the Brazilian economy, where the main NEG features are present, it investigates the long run effects of multilateral trade liberalization policies on the spatial location of economic activity and growth. Secondly, it investigates the long run interplay between trade policy, scale economies (at the industry level) and regional inequality. As already mentioned, the long run general equilibrium effects of trade liberalization policies on Brazilian interregional growth were firstly investigated in the works of Haddad and Hewings (2001) and Haddad and Azzoni (2001). However, both studies divide the Brazilian economy in just 3 huge macro regions, were state level results are tackled using top-down ad-hoc techniques. Moreover, the authors are focused on the spatial effects of a unilateral 25 % import tariff cut (at the border level), disregarding possible differentiated spatial effects stemming from multilateral trade liberalization policies. Lastly, possible economies of scale at the industry level as well as interstate comparative-advantages (differentiated factors proportions and technologies at the state level) are also disregarded.

- Organization of the paper

The remaining of this paper is organized as follows: Section 2 explains the basic structure of the regional CGE model and spells out how it incorporates the main features of the new economic geography. Section 3 describes the basic experiment and discusses its results. Section 4 is devoted to clarify the role played by agglomeration forces and liberal trade policies on regional inequality in Brazil, testing on Krugman (1991) and Krugman and Elizondo (1996). Section 5 concludes.

2 Modeling Issues: General Features

The CGE model recognizes the economies of 27 Brazilian states. Agents' behavior is modeled at the regional level, so that variations in the structure of regional economies can be accommodated. Results are based on a bottom-up approach, i.e. national results are obtained from the aggregation of regional results. The model indentifies eight sectors in each state producing eight commodities, one representative household in each state, regional governments and one Federal government, and a single foreign consumer who trades with each state. Government finances are considered, as well as regional labor markets. The model was structurally calibrated for 2002.

Two previous important extensions were made in the microeconomic structure of the seminal B-MARIA model. The first one attempted to develop a more flexible functional form for the manufacturing sector production function in each of the 27 Brazilian states to incorporate non-constant returns to scale, a fundamental assumption for the recent theoretical developments of the NEG. The second major extension in the structure of B-MARIA was the formal inclusion of transportation margins to account for the real costs of moving goods from one region to another. More details can be found in Haddad and Hewings (2005).

The analysis in this paper focuses a long run "time" frame. For this closure, capital and labor are mobile across regions and industries. Moreover, capital and investment are generally assumed to grow at the same rate. Long run assumptions on the structure of labor markets entail fixed regional unemployment rates (at their initial equilibrium levels) and fixed nominal wage differentials. Labor supply is endogenously increased in regions experiencing employment expansion, according to population interregional migration and exogenous regional labor participation rates. Labor is attracted to more competitive industries in more favored geographical areas. By the same token, capital is oriented towards more attractive industries. This movement keeps rates of return at their initial levels. Household consumption follows household disposable income, and government consumption, at both regional and federal levels, is assumed to move with regional household consumption and national household consumption, respectively.

2.1 Modeling Export Shocks by Regions

The simulations took into account the differences among each set of international trade partners connected to each state in the country. Therefore export side shocks are not homogeneous over the 27 Brazilian states, since they accommodate for differences in foreign policy barriers faced by each regional exporter in the country, according to each regional set of export destinations. This specification is clearly more realistic compared to the traditional one, where a linear export shock is applied equally to all states in the country, as if they all faced the same set of trade partners.

Total exports in each country state were divided in two basic aggregates, agricultures and manufacturing, where foreign tariffs paid by Brazilian exporters were easily available. For each country state, it was considered the 42 most relevant destinations for exports. Information on Brazilian regional exports (FOB prices) in 2002 is available in the public federal site Aliceweb. Information on foreign tariffs

Table 1 Trade liberalization and long-run aggregate results (in percentage-change)

	Trade policy (25 %)	Trade policy (50 %)	Trade policy (75 %)	Trade policy (100 %)
Activity level				
Agriculture	0.0521	0.1395	0.2578	0.4671
Manufacturing	0.2840	0.6361	0.9369	1.1634
Utilities	0.1487	0.2938	0.4548	0.4192
Construction	0.3367	0.6816	1.0401	1.2796
Trade	0.2016	0.3929	0.6209	0.6651
Financial institutions	0.3509	0.8065	1.1619	1.4423
Public administration	0.1547	0.3980	0.5706	0.6958
Transportation and other services	0.1916	0.3797	0.6059	0.5949
Prices				
Investment price index	−1.3828	−2.5006	−3.2713	−5.1088
Consumer price index	−1.1666	−2.0680	−2.6562	−4.2139
Exports price index	−1.0555	−1.9250	−2.3923	−4.2880
Imports price index	−1.5770	−2.9124	−3.8934	−6.0284
GDP price index, expenditure side	−1.1662	−2.0274	−2.5986	−4.1312
National terms of trade	0.5289	1.0171	1.5638	1.8517

(agriculture and manufacturing) was taken from Nereus[1] data base (FEA-USP). With this information it was possible to estimate the average foreign tariff faced by each region in the country.

With the information on average regional foreign tariffs, those values were transformed into foreign power tariffs. Export shocks could then be easily calculated through comparative reductions on the initial set of foreign power tariff values.

3 Describing the Basic Experiment

The first basic exercise subjects the economy to a sequence of huge and increasing liberal trade policy shocks: 25 %, 50 %, 75 % and 100 %, where domestic and foreign import tariffs are simultaneously reduced.

Tables 1 and 2 present the general results for some of the most important macro variables in the economy. As expected, stronger foreign competition translates into lower domestic prices and a better allocation of resources in the long run, promoting generalized efficiency gains and higher welfare levels to consumers. Lower capital costs stimulate long run investment decisions and capital accumulation growth, benefiting both construction and manufacturing sectors in particular. Investment growth entails higher import volume levels. Improvements in terms of trade

[1] Núcleo de Economia Regional e Urbana da USP (www.usp.br/nereus).

Table 2 Trade liberalization and long-run aggregate results (in percentage-change)

	Trade policy (25 %)	Trade policy (50 %)	Trade policy (75 %)	Trade policy (100 %)
Primary factors				
Aggregate payments to capital	−0.9585	−1.6406	−1.9632	−3.5324
Aggregate payments to labor	−0.5097	−0.5746	−0.4458	−1.2332
Aggregate capital stock, rental weights	0.3552	0.7426	1.1300	1.4580
Aggregate employment, wage bill weights	−0.0128	0.0031	0.0134	0.1806
Aggregate demand				
Real household consumption	0.2273	0.4351	0.6858	0.7343
Aggregate real investment	0.4406	0.9103	1.3723	1.7202
Aggregate real Reg. gov. demand	−0.0968	0.4648	0.3173	0.9786
Aggregate real Fed. gov. demand	0.2268	0.4351	0.6863	0.7343
Export volume	2.4702	5.0152	7.4234	10.7046
Import volume	2.5082	5.0779	7.5906	10.6441
Aggregate indicators				
Eq. variation-total (change in $)	8538.29	16404.89	24545.96	36947.46
Real GDP	0.2663	0.5260	0.7941	0.9105

translate into additional welfare gains to consumers. Moreover, a higher relative price of labor over capital in the long run promotes a labor substituting type GDP growth.

Spatial effects from trade liberalization are reported in Table 3. Most of the regions in the country seem to benefit from increasingly lower trade barriers, with just a few exceptions. For instance, some of the richer states in the country such as Minas Gerais and all the states belonging to the south region are clearly harmed by enhanced trade liberalization. On the other hand, GRP evolution for São Paulo, the richest state in the country, follows closely national performance, i.e., higher levels of trade liberalization translate into higher GDP growth rates.

Taking weighted average GRP growth rates, however, preliminary spatial results in Table 3 suggest domestic economic activity to move outside the two most industrialized regions in the country[2] (South and Southeast, so called “developed macro-region”) towards less developed regions (North, Northeast and Center-West, so called “undeveloped macro-region”) as illustrated in Table 4.

Though the “macro tendency” in Table 4 suggests some sort of dispersion of economic activity towards less developed regions as the economy liberalizes, a more careful analysis is needed in order to reach reliable conclusions. Beyond GRP evolution, the analysis in the next section will focus on long run factor mobility (capital and labor) as well.

[2] The so called developed macro_region (south and southeast) responded to over 78 % of Brazilian GDP in 2002. Taking only the state of São Paulo, the figure was 39 %. The national figure breaks down as follows: North region (4.72 %); Northeast region (11.38 %); Center-West region (6.18 %); Southeast region (60.24 %); South region (17.47 %).

Table 3 Trade liberalization and long run regional growth (%)

Region	State	Trade policy (25 %)	Trade policy (50 %)	Trade policy (75 %)	Trade policy (100 %)
North	Acre	−0.0279	−0.1262	−0.1127	−0.3416
	Amapá	0.2044	0.4305	0.8573	0.8515
	Amazonas	0.9886	1.6504	2.1596	1.6641
	Pará	2.4870	4.5596	3.2505	14.4176
	Rondônia	0.1041	0.0371	0.0977	−0.1911
	Roraima	0.1216	0.2110	0.3163	0.5112
	Tocantis	1.1500	2.4903	3.4761	5.1435
	Alagoas	1.1405	2.0996	2.9982	4.1766
	Bahia	−0.3652	−0.7894	−0.7804	−1.6243
	Ceará	0.4440	0.8552	1.4300	1.6639
	Maranhão	1.2606	2.3201	3.2103	4.4979
Northeast	Paraíba	0.6125	1.0552	1.5525	1.8260
	Pernambuco	0.9614	1.6639	2.3754	3.0542
	Piauí	0.4046	0.6945	0.8470	1.4609
	Rio G. do Norte	0.3779	0.5201	0.8922	0.6259
	Sergipe	−0.6323	−1.3352	−1.8203	−2.7024
	Espírito Santo	3.3225	7.0453	11.472	16.4409
Southeast	Minas Gerais	−0.5297	−1.0715	−1.2983	−1.8837
	Rio de Janeiro	0.2442	0.3282	0.4375	0.3679
	São Paulo	0.8635	1.7677	2.5056	3.0262
	Paraná	−0.6630	−1.4412	−2.1786	−3.0560
South	Santa Catarina	−1.2156	−2.3422	−3.2510	−4.5195
	Rio G. do Sul	−2.0360	−3.5637	−4.5457	−6.0294
	Distrito Federal	0.2881	0.3539	0.7596	0.5215
Center-West	Goiás	0.9820	2.3915	3.4908	4.4734
	Mato Grosso	0.9953	1.6230	2.3299	2.5860
	Mato G. do Sul	1.1881	1.9081	2.8743	2.9389

Note: Regional growth is measured as the percentage change in GRP (Gross Regional Product)

3.1 Trade Liberalization and Factor Mobility

For a Johansen CGE type model, the greater the magnitude of any shock applied, the higher the uncertainty over simulated results. Therefore, one possible way to evaluate the robustness of the "macro-tendency" showed in Table 4 is to replicate the first basic experiment using an increasing sequence of milder trade liberalization shocks and to check if qualitative results are maintained.

The second modeling exercise will subject the economy to the following trade liberalization shocks: 5 %, 10 %, 15 % and 20 %. Figures for (macro) regional growth are reported in Table 5 and seem to corroborate (at least qualitatively) previous results presented in Table 4. Therefore, the unraveled "macro-tendency", i.e. dispersion of economic activity as the economy gradually opens up to foreign trade, seems to be robust, whatever the magnitude of the liberalization shock applied to the domestic economy.

Table 4 Trade liberalization and (macro) regional growth (%)

Regions	Trade policy (25 %)	Trade policy (50 %)	Trade policy (75 %)	Trade policy (100 %)
North (1)	1.2998	2.2792	2.2194	4.8309
Northeast (2)	0.2687	0.4070	0.7487	0.6754
Center-West (3)	0.7853	1.4483	2.2040	2.4594
Undeveloped macro region (1 + 2 + 3)	**0.6305**	**1.0926**	**1.4641**	**2.0510**
South (4)	−1.4044	−2.5984	−3.4837	−4.7208
Southeast (5)	0.6230	1.2452	1.8303	2.2005
Developed macro region (4 + 5)	**0.1672**	**0.3811**	**0.6356**	**0.6445**

Table 5 Trade liberalization and (macro) regional growth (%) (milder shocks)

Regions	Trade policy (5 %)	Trade policy (10 %)	Trade policy (15 %)	Trade policy (20 %)
North (1)	0.2710	0.4785	0.8829	1.0834
Northeast (2)	0.0303	0.0040	0.2543	0.2275
Center-West (3)	0.1465	0.2581	0.5220	0.6319
Undeveloped macro region (1 + 2 + 3)	**0.1135**	**0.1751**	**0.4618**	**0.5211**
South (4)	−0.3129	−0.6158	−0.8583	−1.1398
Southeast (5)	0.1482	0.3259	0.3256	0.4946
Developed macro region (4 + 5)	**0.0445**	**0.1142**	**0.0594**	**0.1271**

- Long run Labor and Capital mobility

As international trade flows are intensified trough gradual reductions in trade barriers, a new set of relative prices comes up and renewed market clearing conditions must be met. It turns out that changing domestic relative prices cause factors of production to relocate. Sectors and regions with increased returns to capital and/or labor are benefited.

Long run rates of returns determine capital mobility across sectors and regions. Regional industries where current rates of return[3] to capital increase faster than the national average are expected to accumulate capital at a higher rate than regional capital stock growth. By the same token, for regional industries with lower-than-national average increase in their current rates of return, capital is expected to accumulate at a smaller pace compared to regional capital stock growth. The long run closure assumes that capital interregional mobility is able to keep national average rate of return at its initial equilibrium level.

According to Krugman and Elizondo results, gradual trade liberalization should lead to de-concentration of economic activity towards less developed regions in the

[3] Current rates of returns are defined by the ratio of the rental values of a unit of capital and its cost.

Table 6 Trade liberalization and (macro) regional capital stock growth (%)

Regions	Trade policy (25 %)	Trade policy (50 %)	Trade policy (75 %)	Trade policy (100 %)
North (1)	1.3840	2.6450	2.9930	6.5900
Northeast (2)	0.4170	0.7110	1.1750	1.3160
Center-West (3)	0.5180	1.5610	2.0090	2.8310
Undeveloped macro region (1 + 2 + 3)	**0.6553**	**1.3249**	**1.7597**	**2.8213**
South (4)	−1.0220	−1.8880	−2.4970	−3.4300
Southeast (5)	0.7310	1.4740	2.2170	2.7830
Developed macro region (4 + 5)	**0.2735**	**0.5958**	**0.9861**	**1.1610**

Table 7 Trade liberalization and (macro) regional labor supply growth

Region	Trade policy (25 %)	Trade policy (50 %)	Trade policy (75 %)	Trade policy (100 %)
North (1)	0.7697	1.2147	1.7163	3.1355
Northeast (2)	0.1819	0.2542	0.4731	0.3690
Center-West (3)	0.8636	1.8318	2.6137	3.2657
Undeveloped macro region (1 + 2 + 3)	**0.3922**	**0.6744**	**1.0334**	**1.3074**
South (4)	−1.8229	−3.3597	−4.5276	−6.0589
Southeast (5)	0.2536	0.5192	0.8169	0.8395
Developed macro region (4 + 5)	**−0.2821**	**−0.4815**	**−0.5619**	**−0.9401**

country. When the economy opens up to foreign trade, agglomeration forces are weakened since firms are able to serve foreign markets and consumers have relatively better access to cheaper products from abroad, wherever they decide to locate. Therefore, there is a marginal tendency for primary factors to move outside the most developed regions in the country toward less developed ones, where facilitated access to foreign markets may give rise to higher rates of return to capital and labor.

Long run results for factor mobility also seem to corroborate Krugman and Elizondo's predictions. Results reported in Tables 6 and 7 suggest that, as far as the domestic economy liberalizes, both capital and labor migrate preferentially to less developed regions in the country, accelerating local regional growth and ameliorating pre-existing interregional inequalities.

Capital supply movements follow relocation of economic activity towards less developed regions in the country (Table 6). Labor force moves accordingly, causing labor supply to shrink is the most developed regions in the country and to grow in the less developed ones (Table 7).

It is noteworthy that both (macro) regions in the country become more capital intensive as the economy liberalizes, with the more developed region showing the highest substitution rate between labor and capital (see Tables 6 and 7). This is an

Table 8 Trade liberalization and (macro) regional growth

Regions	Trade policy (25 %)	Trade policy (50 %)	Trade policy (75 %)	Trade policy (100 %)
Undeveloped macro region (1 + 2 + 3)	0.6305	1.0926	1.4641	2.0510
Developed macro region (4 + 5)	**0.1672**	**0.3811**	**0.6356**	**0.6445**
Developed macro region (without SP)	−**0.4939**	−**0.9246**	−**1.1085**	−**1.5501**
São Paulo	0.8635	1.7677	2.5056	3.0262

Note: North region (1); northeast region (2); center-west region (3); south region (4); southeast region (5)

Table 9 Trade liberalization and (macro) regional growth (%)

Regions	Trade policy (25 %)	Trade policy (50 %)	Trade policy (75 %)	Trade policy (100 %)
North (1)	1.2998	2.2792	2.2194	4.8309
Northeast (2)	0.2687	0.4070	0.7487	0.6754
Undeveloped macro region (1 + 2)	0.5711	0.9561	1.1800	1.8942
Developed macro region (4 + 5)	0.1672	0.3811	0.6356	0.6445
Center-South (3 + 4 + 5)	0.2128	0.4597	0.7512	0.7782
Center-South (without SP)	−0.3216	−0.6051	−0.6625	−1.0102

Note: Center-west region (3); south region (4); southeast region (5)

expected result, since the more developed region is also the focal point of manufacturing activity in the country.

- Do dispersion effects spread out uniformly all over the richer macro region?

As results presented in Tables 1, 2, 3, 4, 5, and 6 would suggest, the relative worse performance of the so called “developed macro-region” is far from uniform among its compounding states. First of all, its poor performance is greatly influenced by the negative results presented (uniformly) by the south region. Secondly, the relative better performance of the southeast region is clearly influenced by the results in the state of São Paulo – the focal point of manufacturing activity in the country.

Results in Table 8 reproduce the growth performance of the so called “developed macro-region” (previously shown in Table 4) but now with the exclusion of the state of São Paulo from its set of compounding states.

The growth performance of the “developed macro region” seems to be fully dependent on São Paulo’s results (Table 8). Once this state is excluded from its set of compounding regions, growth rates in the “developed macro region” turn out to be negative.

Would dispersion effects (the so called “macro-tendency”) be maintained once aggregated results are broken down or macro regions are redefined? Table 9 shows the results for the “developed macro-region” as defined in Haddad and Azzoni (2001) to be the aggregation of the South, Southeast and Center-West regions (so called

Center-South). Also, the "undeveloped macro region" is redefined to be the aggregation of the North and Northeast regions (the poorest regions in the country).

When the "developed macro region" is defined to include the center-west, its growth performance is clearly improved. Comparing the growth performance of the so called center-south to the growth performance of the northeast region, it is not so clear-cut that the poorer region is relatively better as the economy opens up to trade (Table 9). However, it seems clear again that the good performance of the "improved" center-south is fully dependent on the results for São Paulo. Once São Paulo is excluded from the set of its compounding members (last line in Table 9), growth performance in the center-south clearly deteriorates as the economy liberalizes. No doubt, the center-south (without SP) fully underperforms the northeast region, corroborating dispersion effects from trade liberalization.

How can theoretical predictions (*á la* Krugman and Elizondo) be reconciled with such a heterogeneous dispersion pattern of economic activity detected in the richest macro-region of Brazil? One possible explanation has to do with huge structural differences among states in the south and southeast regions. São Paulo is a clear focal point of agglomeration economies in the country, concentrating over 40 % of national output (2002). With lower capital costs (due to trade liberalization) and significant additional economies of scale to be explored through renewed investment decisions and the existence of a huge number of local consumers, firms in São Paulo can more easily adapt themselves to enhanced foreign competition in comparison to less competitive industries in the south and southeast regions in the country, where dispersion effects towards less developed regions (lower market-crowding effects) seems to be more evident. Therefore, progressive trade liberalization in Brazil is likely to cause economic diversion towards less developed regions in the country – but not at the expense of economic activity in São Paulo, a state clearly able to reap the benefits from a more open economy, due to its exploitation of additional economies of scale at the local level.

4 The Long Run Interplay Between Economies of Scale, Trade Barriers and Regional Inequality

In the short-run analysis described in Ferraz and Haddad (2009), the authors simulated different values for scale economies in the manufacturing sector in the state of São Paulo. They considered this region to be the focal point of agglomeration economies in the country, concentrating over 40 % of Brazilian manufacturing activity in 2007. In order to isolate the effects of agglomeration forces, constant returns to scale were assumed in every sector in every state. The only exception was the manufacturing sector in the state of São Paulo, for which an interval in the increasing returns to scale curve was considered, ranging from high increasing returns ($\mu = 0.5$) to high decreasing returns to scale ($\mu = 1.5$), i.e., $\mu \in [0.5, 1.5]$ in the manufacturing sector. A set of simulations was then run for different values of μ in this interval.

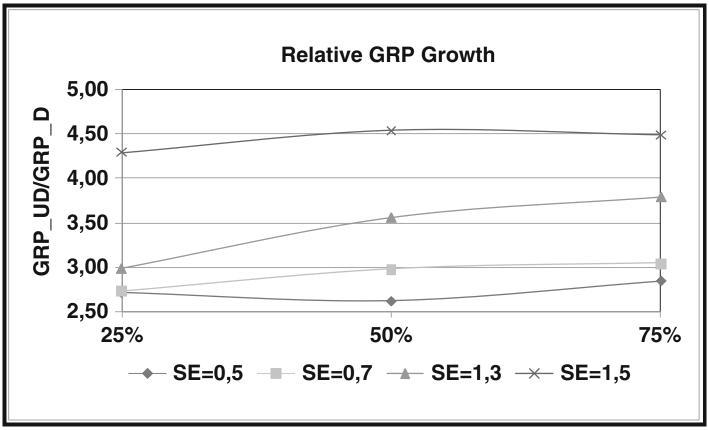


Fig. 1 Scale economies, trade liberalization and regional inequality

Ferraz and Haddad took the usual CGE assumption that producers could not reevaluate their investment decisions in the short-run. Therefore, for a trade liberalization shock under significant interregional transportation costs and fixed investment effort, firms in São Paulo could not fully reap the benefits from lower capital costs in the first place, undermining the exploitation of additional scale economies. It turned out that increased competition from trade liberalization hindered São Paulo's manufacturing activity, especially, and apparently unexpectedly, for higher degrees of increasing returns to scale at the firm level.

The long run analysis developed in this paper fits better to the purpose of the exercise previously devised by Ferraz and Haddad (2009). Since primary factors are now allowed to move, there must be a tension between scale economies and trade liberalization in the long run (*a la* Krugman and Elizondo 1996), in the same spirit as the one involving centripetal and centrifugal forces in the NEG literature. If that is the case for the Brazilian economy, for a given level of openness to trade, the more manufacturing firms in São Paulo can exploit scale economies (higher degrees of scale economies at the firm level), the lower should be the ratio between the undeveloped and the developed region output growth, meaning the developed region grows relatively faster. On the other hand, for a given degree of scale economies, the more the Brazilian economy is opened up to foreign trade, the higher should be the ratio between the undeveloped and developed region output growth, meaning the undeveloped region grows relatively faster (see Fig. 1).

Figure 1 shows the interplay between scale economies, trade liberalization and regional inequality in Brazil. Now that investment decisions can be reevaluated and primary factors are allowed to move, the more developed region (including São Paulo) in the country can fully exploit additional scale economies. For values of SE (scale economies) lower than 1 (increasing returns to scale), the relative interregional growth rate is lower compared to their counterparts for SE values greater than one (decreasing returns to scale), whatever the magnitude of the trade shock.

This evidence corroborates the role of scale economies as an agglomeration force, contributing to the maintenance of pre-existing regional inequalities, as the economy opens up to foreign trade. On the other hand, the greater the intensity of the trade shock, the higher the relative regional growth rate tends to be, despite the level of scale economies. This evidence now illustrates the role of trade liberalization as a dispersion force, contributing to ameliorate pre-existing regional inequalities as the country liberalizes.

The curves in Fig. 1 also give some idea on the net effects resulting from the tension between scale economies and trade liberalization. The relative inter-regional growth ratio is greater than one for all the scenarios under consideration, suggesting dispersion forces to outweigh agglomeration forces even for the highest level of scale economies (SE = 0.5) combined with the mildest trade liberalization shock (25 %). This result suggests net effects from trade liberalization to lead unambiguously to regional activity de-concentration, in the same spirit of Krugman and Elizondo's predictions.

Dispersion effects seem to be more evident for lower levels of increasing returns to scale and more relevant trade liberalization shocks. A question remains to be settled: is the best scenario to fight regional inequality trough trade policy also the best one to stimulate Brazilian GDP growth? As simulation results have shown, the fact that the most developed region in the country can exploit additional scale economies seems to undermine the potential for regional de-concentration, as the economy opens up to trade. However, since scale economies benefits the richest region in the country, it might well be the case that it contributes significantly to GDP growth, even though it can harm regional activity de-concentration. (see Figs. 2 and 3).

Results in Figs. 2 and 3 suggest trade liberalization policies to have a positive impact on GDP growth (for all levels of scale economies), following greater GRP growth rates in the most industrialized region in the country. Also, Southeast regional growth rates are higher for higher levels of increasing returns to scale (SE = 0.5 and SE = 0.7), suggesting extra gains from trade whenever local firms are able to exploit additional scale economies.

Regarding the interplay between scale economies, trade liberalization, regional inequality and growth, results in Figs. 1, 2 and 3 can be summarized as follows:

1. Multilateral trade liberalization policies stimulate higher GDP growth rates and contribute to a more spatially homogeneous interregional growth, whatever the level of scale economies in São Paulo's manufacturing sector;
2. National gains from trade can be higher when São Paulo's manufacturing sector is able to exploit high levels of scale economies. On the other hand, scale economies contributes to exacerbate pre-existing regional inequalities;
3. The best scenario to achieve higher GDP growth rates under trade liberalization does not coincide with the best one to fight regional inequality, suggesting aggregate gains from trade may be higher the more efficient is the manufacturing sector in São Paulo relatively to the less developed region in the country. A clear regional equity-efficiency trade-off emerges.

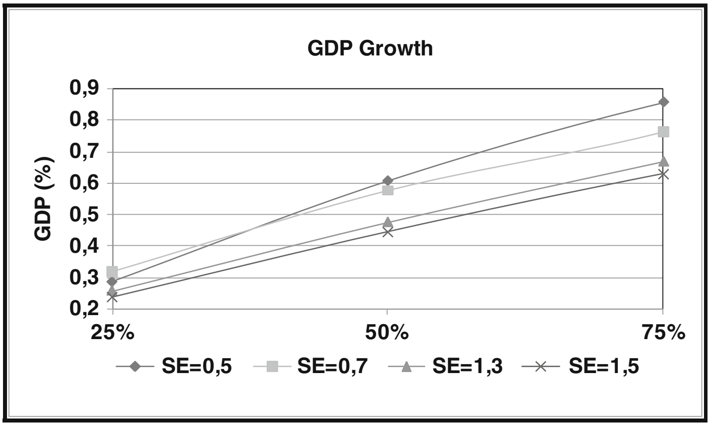


Fig. 2 Scale economies, trade liberalization and GDP growth

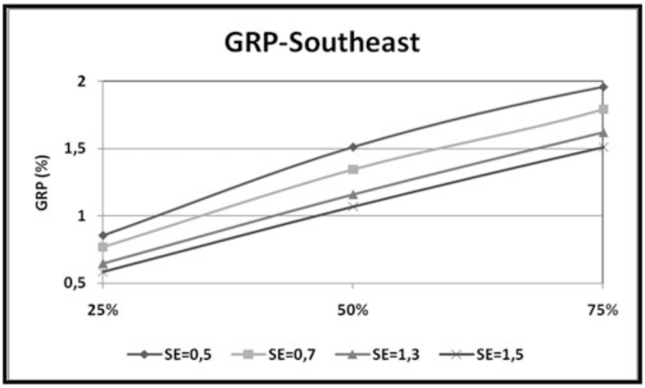


Fig. 3 Scale economies, trade liberalization and southeast GRP growth

5 Final Remarks

The existence of huge urban concentrations associated with high levels of regional income inequality seems to be a typical third world phenomenon. Recent literature on economic development has called attention to the role of physical capital accumulation policies – instead of human capital accumulation policies – in exacerbating pre-existing regional income inequality in developing countries. This paper focus on another possible – tough related – explanation for the existence of huge regional income disparities in the third world: the widespread adoption of importing substitution policies in the past, a point made earlier by Krugman and Elizondo (1996).

Results in this paper seem to corroborate Krugman and Elizondo’s predictions, contradicting previous findings described by Haddad and Azzoni (2001). However, the CGE framework developed by those authors does not model the Brazilian economy at the state level. Instead, the economy is broken down in three macro regions and is subjected to a *unilateral* import tariff shock. State level results are forecasted using top-down ad-hoc techniques. The authors came to the conclusion that unilateral trade liberalization may enhance pre-existing regional inequalities, since the center-south region (an aggregation of the south, southeast and center-west) performs much better in comparison to the poorest regions in the country (north and northeast). To the extent that a multilateral trade liberalization shock cause the same spatial effects as a unilateral one (which seems quite unreasonable), results reported in Haddad and Azzoni may be basically driven by the outstanding performance of the state of São Paulo, as already discussed in this paper, where agent behavior is modeled at the state level.

Results in this paper suggest that multilateral trade liberalization policies in Brazil may potentially contribute to income dispersion, benefiting relatively more the poorest regions in the country, a result that seems to be robust to possible variations in the way the macro regions are defined. Dispersion effects do not seem to come to the expense of economic activity in São Paulo, the focal point of agglomeration economies in Brazil. Growth in the more industrialized developed macro region tends to be highly labor-substituting, following progressive reduction in capital costs. On the other hand, labor substitution in the less developed macro region (labor intensive region) tends to be at more moderate rates.

The best scenario to stimulate output growth does not coincide with the best one to fight regional inequality. With long run factor mobility, São Paulo's manufacturing sector can benefit from the exploitation of additional external scale economies. This is right the opposite result compared to the short run incursions described elsewhere in Ferraz and Haddad (2009), where investment levels were fixed and firms could not reap the benefits of additional gains in efficiency through scale economies. Higher levels of scale economies in São Paulo's industries are then good for regional (southeast) and national growth, but may hinder the full potential to fight regional inequality trough trade policy.

In general, simulation results suggest that trade openness leads unambiguously to income diversion and national output growth, no matter the level of scale economies in São Paulo's manufacturing sector. This result seems to be an important one, particularly if Brazilian authorities are concerned with market oriented growth policies that do not lead to the exacerbation of regional disparities.

Appendix A. Appendix 1 (Overview of the B-MARIA Framework: Underling Assumptions and Philosophy)

We departed from the B-MARIA-27 model, described in details elsewhere (Haddad and Hewings 2005). Its theoretical structure stems from the MONASH-MRF Model (Peter et al. 1996), which represents one interregional framework in the ORANI suite of CGE models of the Australian economy. The interstate version of B-MARIA, used in this research, contains over 600,000 equations, and it is designed for forecasting and policy analysis. Agents' behavior is modeled at the regional level, accommodating variations in the structure of regional economies. The model recognizes the economies of 27 Brazilian states. Results are based on a bottom-up approach – national results are obtained from the aggregation of regional results. The model identifies 8 sectors in each state producing 8 commodities, one representative household in each state, regional governments and one Federal government, and a single foreign consumer who trades with each state. Special groups of equations define government finances, accumulation relations, and regional labor markets. The model qualifies as a Johansen-type model.

A.1. CGE Core Module

The basic structure of the CGE core module comprises three main blocks of equations determining demand and supply relations, and market clearing conditions. Nested production functions and household demand functions are employed; for production, firms are assumed to use fixed proportion combinations of intermediate inputs and primary factors in the first level while, in the second level, substitution is possible between domestically produced and imported intermediate inputs, on the one hand, and between capital, labor and land, on the other. At the third level, bundles of domestically produced inputs are formed as combinations of inputs from different regional sources. The modeling procedure uses a constant elasticity of substitution (CES) specification in the lower levels to combine goods from different sources. Changes in the production functions of the manufacturing sector[4] in each one of the 27 Brazilian states were implemented in order to incorporate (external) non-constant returns to scale, a fundamental assumption for the analysis of integrated interregional systems.

The treatment of the household demand structure is based on a nested CES/linear expenditure system (LES) preference function. The structure of household demand follows a nesting pattern that enables different elasticities of substitution to be used. At the bottom level, substitution occurs across different domestic sources of supply. Utility derived from the consumption of domestic composite goods is maximized. In the subsequent upper-level, substitution occurs between domestic composite and imported goods.

Equations for other final demand for commodities include the specification of export demand and government demand. Exports face downward sloping demand curves, indicating a negative relationship with their prices in the world market.

[4] Only the manufacturing activities were contemplated with this change due to data availability for estimation of the relevant parameters.

Government entrepreneurial behavior is dictated by the same cost minimization assumptions adopted by the private sector.

The model is calibrated taking into account the specific transportation structure cost of each commodity flow, providing spatial price differentiation. Such structure is physically constrained by the available transportation network, modeled in a geo-coded transportation module.[5]

A.2. Behavioral Parameters

Experience has suggested that interregional substitution is the key mechanism that drives model's spatial results. In general, interregional linkages play an important role in the functioning of interregional CGE models. These linkages are driven by trade relations (commodity flows), and factor mobility (capital and labor migration). In the first case, of direct interest in our exercise, interregional trade flows should be incorporated in the model. Interregional input–output databases are required to calibrate the model, and regional trade elasticities play a crucial role in the adjustment process.

One data-related problem that modelers frequently face is the lack of such trade elasticities at the regional level. In this sense, an extra effort was undertaken to estimate model-consistent regional trade elasticities for Brazil, to be used in the B-MARIA-27 model. Estimates are presented in Table A.9.

Other key behavioral parameters were properly estimated; these include econometric estimates for scale economies; econometric estimates for export demand elasticities; as well as the econometric estimates for regional trade elasticities. Another key set of parameters, related to international trade elasticities, was borrowed from a recent study developed at IPEA (www.ipea.gov.br), for manufacturing goods, and from model-consistent estimates in for agricultural and services goods.

A.3. Closure

A long-run equilibrium closure is used in which capital is mobile across regions and industries. Capital and investment are generally assumed to grow at the same rate. The aggregate employment is determined by population growth, labor force participation rates, and the natural rate of unemployment. The distribution of the labor force across regions and sectors is fully determined endogenously. Labor is attracted to more competitive industries in more favored geographical areas,

[5] See Haddad and Hewing (2005), for more details.

Table A.9 Trade elasticities

	International	Regional
Agriculture	0.343	1.570
Manufacturing	1.278	2.079
Utilities	0.011	1.159
Construction	0.002	0.002
Trade	0.694	0.001
Financial institutions	0.137	1.385
Public administration	0.070	0.001
Transportation and other services	1.465	0.001

keeping regional wage differentials constant. While in the same way, capital is oriented towards more attractive industries. This movement keeps rates of return at their initial levels.

References

Ades AF, Glaeser EL (1995) Trade and circuses: explaining urban giants. Quart J Econ 110 (1):195–227

Alonso-Villar O (2001) Large metropolises in the third world: an explanation. Urban Stud 38 (8):1359–1371

Altomonte C, Bosco MG (2011) Environmental standards, spatial location and employment: the EU cement sector case. In: Riccardo C, Marco P (eds) Geography, institutions and regional economic performance. Springer, England, Chapter 12

Campos-Filho L (1998) Unilateral liberalisation and Mercosul: implications for resource allocation. Rev Brasileira de Econ 52(4):601–636

Ferraz L, Haddad E (2009) On the effects of scale economies and import barriers on Brazilian trade performance and growth: an interstate CGE analysis. Stud Reg Sci 39(1):53–65

Flores RG Jr (1997) The gains from Mercosul: a general equilibrium, imperfect competition evaluation. J Policy Model 19(1):1–18

Fujita M, Krugman P (1995) When is the economy monocentric?: von Thunen and Chamberlin unified. Reg Sci Urban Econ 25:505–528

Gonzaga GM, Terra MCT, Cavalcante J (1999) O Impacto do Mercosul sobre o Emprego setorial no Brasil, mimeo

Haddad EA, Azzoni CR (2001) Trade and location: geographical shifts in the Brazilian economic structure. In: Guilhoto JJM, Hewings GJD (eds) Structure and structural change in the Brazilian economy. Aldershot, Ashgate

Haddad EA, Hewings G (2001) Trade and regional development: international and interregional competitiveness in Brazil. In Johansson B, Karlsson Ch, Stough RR (eds) Theories of endogenous regional growth: lessons for regional policy. Theories of endogenous regional growth: lessons for regional policy. Springer, Berlin, pp 181–208

Haddad EA, Hewings GJD (2005) Market imperfections in a spatial economy: some experimental results. Q Rev Econ Finance 45:476–496

Haddad EA, Bonet J, Hewings GJD, Perobelli FS (2009) Spatial aspects of trade liberalization in Colombia: a general equilibrium approach. Pap Reg Sci 88:699–732. doi:10.1111/j.1435-5957.2009.00268.x

Henderson JV (1996) Ways to think about urban concentration: neoclassical urban systems versus the new economic geography. Int Reg Sci Rev 19(1&2):31–36

Isserman AM (1996) "It's obvious, it's wrong, and anyway they said it years ago"? Paul Krugman on large cities. Int Reg Sci Rev 1996(1&2):37–48

Krugman P (1991) Increasing returns and economic geography. J Polit Econ 99:483–499

Krugman P (1991a) Geography and trade. MIT Press, Cambridge

Krugman P (1991b) Increasing returns and economic geography. J Politic Econ 99:483–499

Krugman P (1993) First nature, second nature, and metropolitan location. J Reg Sci 34:129–144

Krugman P, Elizondo LR (1996) Trade policy and third world metropolis. J Develop Econ 49:137–150

Peter MW, Horridge M, Meagher GA, Naqvi F, e Parmenter BR (1996) The theoretical structure of MONASH-MRF. Preliminary Working Paper No. OP-85, IMPACT Project, Monash University, Clayton

Sanguinetti P, Martincus CV (2005) Does trade liberalization favor spatial de-concentration of industry? Investigación y Publicaciones, Universidad Torcuato di Tella

Environmental Standards, Delocalization and Employment: The Case of the EU Cement Industry

Maria Giovanna Bosco and Carlo Altomonte

Abstract CO_2 emission reduction is on the political agenda of all developed countries. In this paper we try to assess the impact on one energy – intensive industry, cement, of a likely rise in the cost of pollution permits generated by the new EU cap-and-trade system on emissions (EU ETS). Exploiting the characteristics of the industry (high unit transport costs and homogeneous product), we present a theoretical model of market segmentation for the overall EU regional area. Based on this, we then provide an estimate of the economic effects for the industry of an increase in the production costs due to the full enforcement (auctioning) of the new European ETS system, i.e. in the case of its implementation without free allowances or carbon leakage provisions. We proceed in two steps: first, we estimate the extent to which the increase in environmental costs leads to a substitution of locally produced cement with imports from non EU countries, not subject to the ETS; second, we estimate the impact of this substitution effect on the employment of the industry. We find that in case of full auctioning of allowances the overall impact on employment would be negative, leading to a loss of some 25 % of jobs in the industry. The latter strongly justifies the use of carbon leakage provisions in the upcoming European regulation.

Keywords Environmental standards • Trade • Delocalization

JEL Classification: F23, R12

The authors are grateful for the very useful comments received during seminar presentations at the Marshall School of Business – University of Southern California, the University of San Francisco, DEIR – Sassari University and ISLA – Bocconi University. The authors retain the sole responsibility for their views, errors and omissions.

M.G. Bosco (✉)
ISLA, Bocconi University, Via Guglielmo Rontgen, 1, 20136 Milan, Italy
e-mail: mariagiovanna.bosco@unibocconi.it

C. Altomonte
IAM Bocconi University and SDA Bocconi, Milan, Italy

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_17,

1 Introduction: The New EU ETS System

The reduction in carbon dioxide (CO_2) emission is one of the pillars of the Kyoto protocol. And yet, although European countries have been pioneers in the effort of cutting emissions by the introduction of a cap-and trade mechanism, we do not yet have a proper tradable carbon credit system at the global level.[1] In this sense, we face a typical prisoner's dilemma: cooperation would make everybody better off, but the different degrees of economic development reached by the players create a lot of incentives to behave opportunistically. Developing countries in fact, eager to reach the level of economic development and GDP per capita of developed countries, are reluctant to face newly imposed environmental constraints in this convergence process (as much as their developed counterparts did in the past). Developed countries, on the other hand, do not want to be left alone in bearing the cost of such a system. Therefore, setting a per capita emission level on a global basis is feasible only in the long run, or only to the extent that appropriate compensation mechanisms, with extra costs for the developed countries and waivers for developing ones, can be negotiated.

However, understanding that the cost of waiting might only worsen the risks associated to climate change, the European Union has embarked into a pioneering, and insofar unilateral, binding effort to significantly reduce emissions in the post-Kyoto context. In particular, the EU and its member States have committed[2] to an independent quantified economy-wide emission reduction target of a 20 % emission reduction by 2020, compared with 1990 levels, via a specific European Emission Trading System (EU ETS).[3]

In this article, we thus try to assess the potential implications for a particular energy-intensive industry, cement, of a strict implementation of the new EU ETS, as stemming from the current legal framework.

The key element of the new EU ETS is related to the reform of the emission allowance trading system. The new rules provide that for the third phase of the ETS (2013–2020) progressively lower quotas of total CO_2 emissions will be allowed (cap), with full auctioning of allowances for the individual firms operating in the

[1] The agreements reached on December 2010 in Cancun, Mexico, at the 2010 United Nations Climate Change Conference include a number of different commitments at the country-level for reducing emissions, but insofar no formal legally binding international agreement, as envisaged in the original Kyoto protocol, has been set up.

[2] Directive 2009/29/EC of the European Parliament and of the Council of 23 April 2009, amending directive 2003/87/EC.

[3] Emission trading is one of the three international market mechanisms foreseen by the Kyoto Protocol on climate change in order to meet the GHG reduction objectives. Emission trading is cost-effective because it guarantees the achievement of an emission target with the lowest overall abatement costs. In the absence of transaction costs the initial allocation of emission permits, while having potentially large distributional effects, does not change the final outcome, i.e. the emission reduction taking place in the most efficient way. The allocation of emission rights (such as grandfathering and auctioning), determines the economic burden undergone by the participants.

power sector (trade), while a transitional system will be put in place for other sectors. The latter implies that, always within the new CO_2 quota, emissions for industrial producers will be initially allocated for free, based on strict technical benchmarks defined according to the 10 % most efficient installations: firms meeting the benchmarks (and thus efficient emission-wise) are in principle able to receive allowances for free. Those emitters not meeting the benchmarks will instead have to either lower their emissions or purchase additional allowances.

In order to distinguish the different cases, the European Commission issued a document (EC 2008) assessing the impact, for various EU energy-intensive sectors, of the implementation of the EU ETS directive. The baseline assumption, which we will also use in this paper, is that under the new system unit production costs would increase by adding the 'full' price of the CO_2 allowance, which can be estimated trading at 20€ per ton. In case of low international competition and in the presence of highly concentrated industries with exit barriers, such an increase in costs would likely be passed through to customers as increased final prices, and thus final consumers would ultimately bear the costs of tighter environmental standards. Things would instead get more complicated for those sectors exposed to international competition: in this case, firms could delocalize or fragment internationally the value chain and consumers could turn to external producers, so that their demand elasticity could be affected sharply. The latter would generate 'carbon leakage',[4] with the result that the total amount of emissions will not be reduced at the global level, while the EU economy might suffer from the delocalization.

To solve this problem, the EU ETS system has been designed in such a way that the evolution over time in the free allocation of allowances actually depends on the exposure of a given industry to carbon leakage.[5] According to the EU ETS directive, those sectors not deemed to be exposed to a significant risk of carbon

[4] Carbon leakage would occur in case customers located in Europe start to source their supply from producers located outside from the EU. These producers, lacking a global agreement on carbon emissions, do not have to bear the costs imposed by the EU ETS rules, and thus become cheaper notwithstanding transport costs. These producers could also be suppliers formerly located in Europe that have delocalised their production to non-EU locations (direct delocalization), or that keep their existing capacity within the EU, but undertake their capacity expansion outside from the EU countries (indirect delocalisation).

[5] According to the EU legislation (Article 10a of the revised Directive), a sector or sub-sector is "deemed to be exposed to a significant risk of carbon leakage if: the extent to which the sum of direct and indirect additional costs induced by the implementation of this directive would lead to a substantial increase of production cost, calculated as a proportion of the Gross Value Added, of at least 5 %; and the Non-EU Trade intensity defined as the ratio between total of value of exports to non EU + value of imports from non-EU and the total market size for the Community (annual turnover plus total imports) is above 10 %. A sector or sub-sector is also deemed to be exposed to a significant risk of carbon leakage: if the sum of direct and indirect additional costs induced by the implementation of this directive would lead to a particularly high increase of production cost, calculated as a proportion of the Gross Value Added, of at least 30 %; or if the Non-EU Trade intensity defined as the ratio between total of value of exports to non EU + value of imports from non-EU and the total market size for the Community (annual turnover plus total imports) is above 30 %."

leakage will see an initial reduction of the free allocation of allowances of 80 % with respect to the emissions quotas to be allocated in 2013. This amount will gradually decrease to 30 % in 2020, with a view to reaching no free allocations in 2027. Instead, in case a sector or sub-sector has been put on the carbon leakage list, these progressive reductions do not apply, with free allocation remaining constant throughout the period 2013–2020.

And yet, this exception has generated a lot of debate, especially in light of a possible revision over time of these carbon leakage provisions. For these reasons, in this paper we try to put forward a methodology able to assess the possible costs in terms of employment of a 'full' implementation of the ETS directive, i.e. without carbon leakage/free allowance provisions.

More specifically, we will estimate the potential job loss that would occur in the cement industry in Europe if local production were substituted by imports, elaborating a method to estimate the so-called Armington (1969) elasticity for the cement industry.[6] In this case, production could be relocated in other Mediterranean countries such as Egypt or Morocco, and cement could be shipped to continental Europe, as the shipping costs would be lower than the increased price of European cement.

The intuition behind our empirical work is supported by a theoretical model for the potential delocalization process that relies on the new economic geography approach and deals with the EU regional area as proxied by a circumference, along which producers distribute according to consumers' demand. It proves that one of the consequences of increased environmental standards costs would be that of driving off the EU many plants/producers and increasing internal transport costs, an effect that would add to the drop in labor.

As a corollary, our work provides an original measure for the Armington elasticity for the cement industry. As underlined in EC (2007), reliable estimates of prices and Armington elasticities do not seem to exist for the markets analyzed in their study, suggesting the need for alternative indicators to measure exposure to (international) competition. As such, the methodology could be used as a benchmark in light of future revisions of the EU ETS system.

In terms of structure of our work, Sect. 2 describes the cement industry, its main characteristics, and its exposure to the new emission regulations. Section 3 discusses a review of the relevant literature used in this paper, while Sect. 4 presents our theoretical model of market segmentation. Based on this, Sect. 5 develops our empirical strategy, with an estimate of the economic effects for the cement industry discussed in Sect. 6. Section 7 concludes.

[6] Armington (1969) introduced the assumption that final products which are traded internationally are differentiated based on the location of production. The Armington elasticity specifies the degree of substitution in demand between similar products produced in different countries. The higher the value of the Armington elasticity, the closer is the degree of substitution.

2 Facts and Figures from the Cement Sector

The cement sector has some peculiar features that make it attractive to economists. First, although produced from natural raw materials which vary from plant to plant, cement can be considered a relatively homogeneous product as there are only a few classes of cement, with each class and products interchangeable across producers. Therefore, the price is the most important sales parameter next to customer service; quality premiums exist but are insofar rather limited.[7]

Second, cement is a capital-intensive industry with huge entry barriers. The cost of cement plants is usually above € 150 M per million tons of annual capacity, with correspondingly high costs for modifications. The cost of a new cement plant is equivalent to around 3 years of turnover, which ranks the cement industry among the most capital intensive sectors. Long time periods are therefore needed before investments can be recovered and plant modifications have to be carefully planned and must take account of the long-term nature of the industry.

Third, the industry is extremely energy-intensive: each ton of cement[8] produced requires 60–130 kg of fuel oil or its equivalent, depending on the cement variety and the process used, and about 105 KWh of electricity. As a result, measured by the share of energy purchases in turnover, cement and brick manufacture is almost twice as energy intensive as the next most energy-dependent sector, pulp and paper (EC 2007).

Fourth, land transportation costs are significant. Cement is the main input for ready –mix concrete, and ready – mix concrete does have one remarkable characteristic: other than manufactured ice, perhaps no other manufacturing industry faces greater transport barriers (Syverson 2008). As such, in general it is not profitable to haul cement beyond 200 or at most 300 km by road. Bulk shipping has changed that, however, and it is now cheaper to cross the Atlantic Ocean with 35,000 t of cargo than to truck it for 300 km. As a result of this transport costs' structure, in large countries the market is segmented into regional areas, with the exception of a few long-distance transfers (where, for example, sea terminal facilities exist).

Last but not least, about 99 % of the demand for cement stems from the construction sector. The fates of the two industries are linked, as the cement sector is a monopsony, where suppliers completely follow the ups and downs of the building industry. And in fact, cement production has been growing steadily in the past decade, pulled by the construction sector in both developed and developing countries. The turnover of the EU-25 cement, lime and plaster industry was almost

---

[7] Although a homogeneous product, new technologies currently developed by some of the top companies in the industry are increasingly providing "greener" cements, less intensive in energy use and/or with specific technological properties.

[8] Clinker, the main intermediate of cement, is actually responsible for 100 % of cement's direct emissions. Around 60 % of emissions come from the de-carbonation process of limestone, the basic material for cement. The rest is related to the energy necessary to heat up the de-carbonated limestone in order to produce clinker.

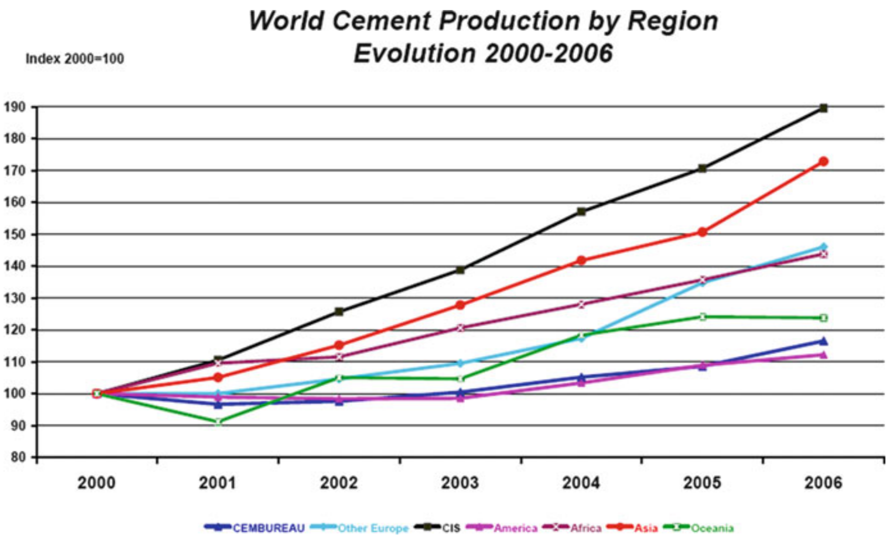


Fig. 1 World cement increasing production (Source: Cembureau 2008)

€22 billion in 2004 and the value added was €9 billion. The workforce of the cement industry in the EU-15 amounts almost to 60,000. Figure 1 shows the growing world production in cement while Fig. 2a, b illustrate the trend in EU imports and exports.

The above discussed characteristics are likely to affect the industry in terms of impact of the new EU ETS regulation mainly via two channels. On the one hand, the energy intensity of the industry is such that the EU average net emission ratio is 858 kg CO_2/ton of clinker (see BCG 2008 for more detailed figures), and only a slight reduction is forecasted for the future, with the use of the best available technologies. The latter implies that if 1 t of raw output (clinker) is being sold on the wholesale international markets at x € per ton, a gross estimate of the increase in costs, assuming a full pass-through of the cost of CO_2 emissions and a (historic) average price of emissions set at 25 €/t, would see a jump of roughly x + 20 € per ton of output (EC 2008), that is an increase of around +60 % from current clinker prices.

The second characteristic is that the industry, notwithstanding its capital intensity and high road transport costs, is nevertheless subject to a risk of carbon leakage. The latter stems from the great reduction in transport costs by sea induced by the emergence of cement producers in developing countries, in particularly China, as shown by the huge increase in EU-27 imports reported in Fig. 3:

Table 1 below summarizes some interesting figures about current cement prices versus transport costs:

As a result of these combined effects (high impact of full auctioning on the cement price, and growing availability of relatively cheap imported cement from overseas), the cement industry is de facto at risk of carbon leakage, with delocalization of production outside from the EU. The latter would be true especially for

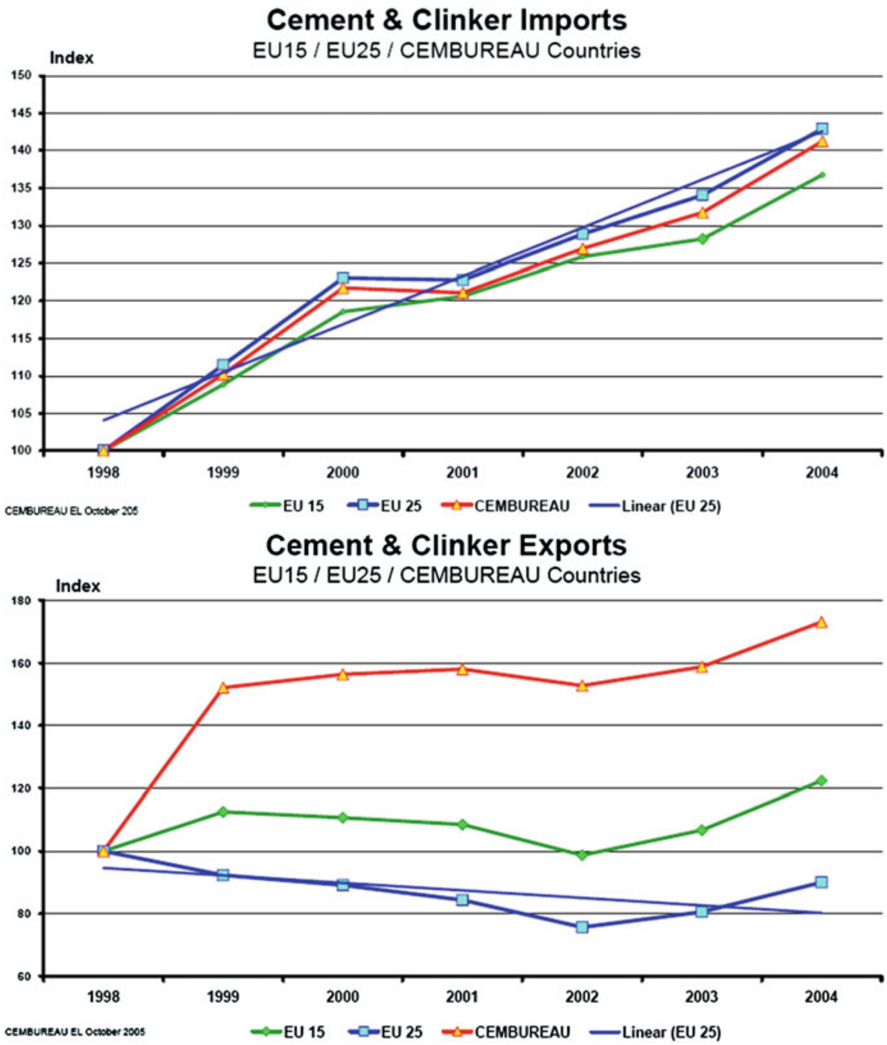


Fig. 2 (**a**) Imports in the EU- Cembureau countries (Source: Cembureau 2008). (**b**) Exports in the EU- Cembureau countries (Source: Cembureau 2008)

those European countries located around the Mediterranean basin, such as Greece, Italy, southern France and Spain, where the possibility of substitution of locally produced cement with the one imported from third countries is highest (BCG 2008; EC, McKinsey and Ecofys 2006). Incidentally, these are also the markets where both supply and demand of the industry are among the highest in Europe. For these reasons, the EU ETS directive has currently included cement among the industries at risk of carbon leakage, thus granting it free allowances based on benchmarks until 2020.

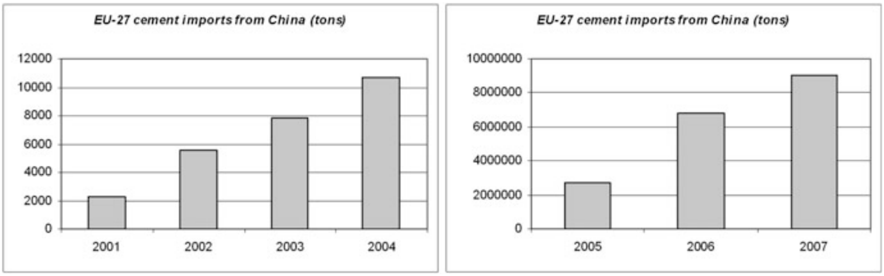


Fig. 3 EU-27 cement imports from China (Source: Comext 2008)

Table 1 Example of shipping costs and market prices

Shipping costs	Cement prices	Final European price
From China (and Thailand): 30 euro/t	Europe: 60/90 euro/ton	From China/Thailand: 50/90 euro/t
From Egypt: 16 euro/t	China: 20/40 euro/t	
From Algeria: 20 euro/t		

Source: Italcementi

But what about the alternative scenario, i.e. what would have happened had a full auctioning system been applied to the EU cement industry, as some of the initial proposals were suggesting? In what follows we try to set up a methodology able to estimate the employment effects of these regulatory changes, applying it to the cement industry. As such, however, the methodology can be replicated for any industry in Europe, provided data are available, and could thus be used as a benchmark in light of future revisions of the EU ETS system.

3 Related Literature and Studies

For our purposes, we consider three streams of literature. The first area of studies concerns the impact on employment of delocalization (offshoring).

3.1 Potential Job Loss Due to Offshoring

The potential jobs losses that could stem in industrialized countries from the relocation of jobs abroad, typically to lower cost locations, is a central tenet of the globalization debate. In general, it is widely acknowledged that offshoring decreases labor demand in the considered industry, because firms substitute domestic workers with intermediate inputs imported from abroad: ceteris paribus, this worsens employment opportunities and reduces average wages in that industry

(Feenstra and Hanson 2003). Levy (2005) also underlines how offshoring – just like trade – can lead to wealth creation for shareholders, but not necessarily for countries and employees, with many displaced workers having difficulties 'trading up' to higher skilled jobs.

For the European case, Aubert and Sillard (2005) study for French firms the effect on employment due to offshoring, finding that between 1995 and 2001, 13,500 jobs are lost each year on average due to offshoring in the manufacturing sector.

A study by Ferraz, Haddad and Terra (2010) analyzes the effects of trade liberalization on a quite closed economy as Brazil by a CGE model. They find that trade liberalization may potentially contribute to income dispersion because of relocation of economic activity, though in the Brazilian case this is good news since the poorest regions in the country may benefit from this relocation, at the expense of the larger and richer urban areas as São Paulo.

Recently, the debate on offshoring has focused on service jobs, as the world economy has gradually shifted towards services as the main share of GDP. A cornerstone in the debate about trading tasks against trading jobs is Blinder's mainly qualitative article about the potential risks for American jobs due to outsourcing services from India, China and other emerging economics (Blinder 2005). Jensen and Kletzer (2005) evaluated the potential job loss in the service sector in the US, while Kletzer, (2000, 2001) evaluated the loss in the manufacturing US firms after outsourcing.

In our case, the potentially displaceable jobs belong to an industry – cement – whose production could be offshored abroad as production becomes too costly in the EU, while production in peripheral areas is still competitive. The standard empirical approach for evaluating the effects for the manufacturing sector uses Input–output tables and the prominent literature starting with Feenstra and Hanson (1996, 1999) and more recently Amiti and Wei (2004, 2007), whose methodology will be discussed in details in Sect. 5. The relevant result from these studies is that the job displacement depends on the level of aggregation at which the effect is evaluated. When very detailed industries are analyzed, some negative effects from offshoring appears, while with aggregated data, especially in a context of flexible labor markets, the effect disappears, due to some relocation of the displaced jobs across industries.

3.2 The Cement Industry

The second area of literature relevant for our work concerns characteristics of the cement industry (Syverson 2004, 2008; Szabó et al. 2006). As already stated, some features of the cement industry make it attractive to economists, such that the industry, previously neglected, has recently been studied with respect to productivity issues and market segmentation, relevant because of high transport costs. Moreover, as an energy-intensive sector it also enters the environmental debate in the EU as one of the sectors most potentially affected by the new UE policy in terms of CO_2 emissions. To that extent, Szabó et al. (2006) in a study on the EU emission trading system find that the introduction of a carbon value tax reduces the cement demand in the EU, that the cement sector would be a net buyer of permits at the EU-wide permit price, set at 28 €/t $CO_{2.}$ Most interestingly, they also provide a simulation about the positive link between permits price and carbon leakage: at the permit price of 40 €/t for CO_2 the EU15 will produce 5Mt of cement less than in the BAU (Business As Usual) case (around 3.5 % of its total production), while the other regions will increase their production, however to a less extent. Recent estimates set the fall in cement production (and exports) in the EU at about 30 % with a price of €45/t CO_2 and the increase in imports at about 25–30 %, with consequent carbon leakage.

3.3 Market Structure, Preparatory and Industry Studies

The third stream of literature is represented by articles trying to theoretically assess the impact of climate change challenges on firms' behavior, and by studies carried independently or by/for the European Commission herself, aimed at assessing the impact – in terms of costs – of the EU ETS. In a forward looking study Mani and Wheeler (1997) show that if the level of abatement expenditure per unit of output is considered, five sectors emerge as "dirty industries": Iron and Steel, Non-Ferrous Metals (such as aluminum), Industrial Chemicals, Pulp and Paper, and non Metallic Mineral Products (such as cement). A theoretical model by Babiker (2005) adopts a general equilibrium framework to analyze interactions between firms producing energy intensive products and finds that significant relocation of energy-intensive industries away from the OECD may occur, depending on the type of market structure, with leakage rates as high as 130 %, in which case GHG control policies in the industrialized countries actually lead to higher global emissions due to carbon leakage.

In the article by Kolk and Pinske (2008), the strategic reaction of a firm facing the challenge of climate change is examined, especially under the possibility of developing "green" firm-specific advantages (FSAs) that typically would give the firm some competitive margin over rivals, by means of innovation and technological change. Although they do not deal with the cement sector, but with the automotive and oil industries, they implicitly introduce the issue of carbon leakage by recognizing that climate change creates a geographically dispersed and moving target: while it may form a threat in one location, it can be an opportunity in another. And firms can re-use and re-adapt their FSAs into territories that did not subscribe the Kyoto protocol, gaining competitive margins at home and abroad. A recent study by Sanna – Randaccio and Sestini (2009) analyzes the impact of asymmetries in environmental policy on the international location strategy of firms, when countries differ in terms of market size, and barriers to trade and foreign direct investment (FDI) have been removed. Taking as a representative case the cement sector, they prove how the higher the markets' asymmetry, the higher the probability for firms to relocate abroad via FDI. As a consequence, in order to avoid carbon leakage, differentiated environmental standards should be adopted by markets of different size.

The preparatory studies for the EU ETS directive (EC 2007; ETUC et al. 2007; JRC and IPTS 2007; EU, Mckinsey and Ecofys 2006) envisaged different scenarios under the ETS scheme compared with a BAU situation. Although they underline the global positive effect for the EU in terms of reduced CO_2 emissions and reduced costs from accomplishing the Kyoto goals, they also indicate which sectors will encounter larger costs during the processes, and clearly cement is among these. A Bruegel Policy Brief (Delgado 2007) also highlights how the EU is not carbon competitive if compared to the rest of the world, and how in the absence of fair and unbiased carbon pricing schemes worldwide, there is a real risk that businesses will resort to regulatory arbitrage which will entail a shift in where emissions take place – but no reduction in global emissions.

In a study by OECD (2005) a numerical partial equilibrium model (originally developed by the Commission's Institute for Prospective Technological Studies, IPTS) was used to simulate the impact of climate policies on the cement sector. The CO_2 tax was set at a level of €15/t of CO_2. Rather than using an Armington elasticity, in which one parameter is used to capture the various "rigidities" affecting international trade, the model incorporates transportation costs, capacity utilization rates and profits for every producing country, and the average cement price for every consuming country. In this simulation, the fall in production is found to be relatively modest in Europe, set at less than 1 % below baseline in 2020.

However, in a recent and comprehensive study undertaken by the EU in terms of price impact of the new EU ETS [DG Ecfin Economic Paper N. 298,] the cement sector is included in the minority of roughly 50 sub-sectors and 100 production processes, that "would require significantly higher price increases to recover their additional costs triggered by the carbon constraint". (Economic Paper N. 298, p. 78). While comprehensive and detailed, this study however does not explicitly address the market consequences under the new emission costs, therefore lacking an estimate of the economic impact for the cement industry, to which we now turn.

4 A Theoretical Framework for the Effects of Delocalization on Internal EU Transport Costs

We present here a theoretical model aimed at assessing the increase in internal transport costs in Europe, should the local cement production in the EU be offshored. The intuition stems from the fact that, moving from a more or less homogeneously segmented production market as the current one (according to the 200/300 km. rule), to a situation in which production arrives from abroad at the EU borders, it would be more expensive to serve 'central' land-locked locations. As such, the model represents the justification for the second part of our empirical estimation, where we proxy the effect of an increase in the emission price as being analogous to a reduction in the distance (transport costs) of imports from third countries.

In the model, in particular, we study the movement of one unit of production from a number *n* of internal European production sites geographically segmented (as it is the case of the cement industry) to a number *m* of external locations placed at the border of the EU (e.g. harbors), serving the same amount of consumers.

The assumptions and data of the model are as follows.

1. Europe's surface is proxied by a circle. The total EU27 surface is $S_{EU} = 4{,}200{,}000\ km^2$; thus the average distance from centre (radius) $R_{EU} = \sqrt{\frac{S_{EU}}{\pi}} = 1{,}156.25$ km.
2. Let *n* be the number of cement production units, this number being the current equilibrium configuration of the EU industry. Suppose that each production unit serves consumers located in a disk in which all production concentrates in the center of the disk, and consumers are randomly distributed throughout the rest of the area. Figure 4 below depicts the situation, with $n = 4$ producers. Note that as *n* increases, the sum of all the individual areas covered by each producer as under hypothesis 2) closely approximates the area of the entire circle.[9] As a result, with *n* relatively large we would have that the average area served by each production units is S_{EU}/n.

If producers are located as in Fig. 4, the average internal distance between a producer and a consumer is given by

$$d_n = \int_0^R z f(z) dz$$

where R denotes the radius of the disk of area S_{EU}/n, and f(z) is the density of consumers at any given distance *z* to the center. Since $R = \sqrt{\frac{S_{EU}}{\pi n}}$ we then have $R = 1{,}156.25/n$.

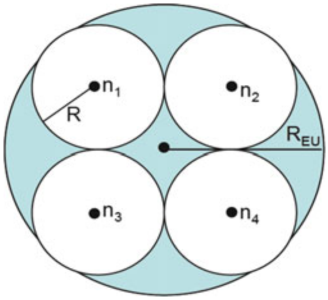
Considering, as it is standard in the location literature, a uniform distribution of consumers $f(z) = 2z/R^2$ (e.g. Head and Mayer 2000 for the EU case), by solving the integral we get that the average internal distance between a producer and a consumer for every production plant is given by:

$$d_n = \frac{2}{3}R = \frac{2}{3}(1,156.25/n) = \frac{770.8}{n}$$

For the case of the cement industry, if Q is the total amount of tons of cement produced in Europe by the *n* plants, the total number of km travelled by cement will

[9] The latter is true under the assumption that the new producers will also segment the market, which is however a feature typical of the cement industries, as discussed in Syverson (2004).

Fig. 4 Internal production of goods within a circle of homogenously distributed consumers



be $D = d_n * Q/C$, where C is the average capacity of a truck. D is then the total amount of km that can be used to calculate the total amount of emissions of CO_2 associated to the transport of domestically produced cement within the EU. Once the simple parameter θ of kg of CO_2 emitted per kilometer by an average truck is known, we thus have that the total emissions are $\theta = \vartheta D$, in kg.

Suppose now that we substitute local production with imports from production plants located outside Europe. Excluding the transport cost from the (non-EU) production site to the EU, i.e. focusing only on internal transport cost, we can assume that cement enters Europe from three locations E*m*, where *m* is: (1) North-West, e.g. Rotterdam; (2) North-East, e.g. Eastern European routes; (3) South, e.g. Italy or France shores. For simplicity, these locations are equidistant along the EU circumference, and sit at the centre of the circumference (the geographical border) of the market they serve, as depicted in Fig. 5.

In equilibrium, each location E*m* will thus serve consumers distributed over a surface area of $S_{EU}/3$ for a maximum span of R_{EU}, above which it is convenient to serve consumers from the other location. Given the same density function of consumers, what is then the average distance of the consumers served by each producer located in E*m* over its share of the circle?

The problem is not trivial since, as each E*m* is placed at the border (and not at the centre) of its market share, we cannot use the well-known formulas for the circular sector, but we would need to solve an elliptic integral. We can however proxy the actual market segmentation of each location with three circular areas, each of radius R_{EU} and centered around each of the three locations E*m*, as shown in Fig. 6 below. The intersection of each dotted circle (of radius R_{EU}) with the main one (also of radius R_{EU}) generates a well-known geometric figure known as "lens", which is a good proxy of the market we want to capture (we discuss the degree of approximation later)

The advantage of modeling location as in Fig. 6 is that when, as it is in our case, the centres of the two (symmetric) circles are placed at a distance equal to their radius (as shown for example in the case of location E3 in Fig. 6), the area of the lens has a precise value given by[10]

[10] The latter case is known to mathematicians as the "fish bladder".

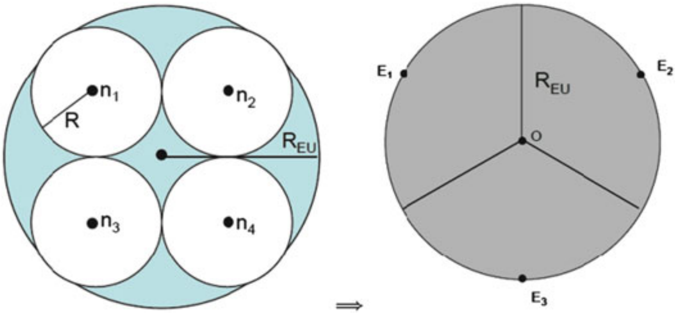
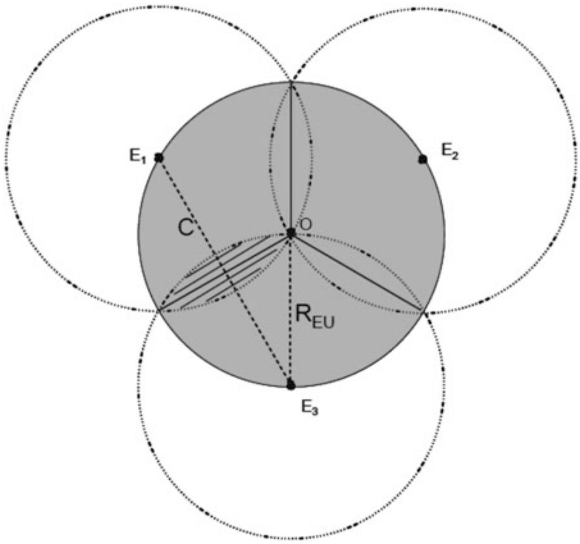


Fig. 5 From internal to border production

Fig. 6 Calculus of average distance faced by each producer E*m*



$$A = \frac{1}{6}(4\pi - 3\sqrt{3})R^2 = 1.22837\ldots R^2$$

Now, focus on the Southernmost circle in Fig. 6 and imagine two scenarios: (a) consumers are homogenously distributed over a circular area of radius R_{EU}, in which case the average distance of consumers that can be served by a (central) location E_3 is $d_A = 2/3$ R, as we have already derived; (b) consumers are homogeneously distributed only along the radius, in which case the average distance is clearly dB = ½ R. It follows that when consumers are distributed within the lens of the circle, the average distance has to be an intermediate value between its upper (dA) and lower (dB) bound. In other words, we have to add to the lower bound dB a value $\alpha\frac{1}{6}R$, where α is the share of the area of the lens with respect to the entire area of the circle.[11]

It is straightforward to see that $\alpha = \frac{A}{S_{EU}} = \frac{1.22837R^2}{\pi R^2} = 0.391$.

As a result, under the latter approximation, each "border" producer will serve its share of consumers located at an average distance of $d_E = d_B + \alpha\frac{1}{6}R = 0.5R + 0.065R = 0.565R$.

In the EU case, where $R = R_{EU} = 1{,}156.25$ km, the latter value is 653.5 km, which is the average distance that each producer E*m* sitting at the central border of its market (as in Fig. 6) has to travel to serve its share of consumers.

The total average distance to serve all EU consumers (covering a surface S_{EU}) when production is located in three sites at the EU border is thus $D_E = 653{,}5$ km * 3.

Note that this is again an upper bound measure, since the three producers actually generate a certain overlap in the segmentation of the market, as shown by the striped areas shown in Fig. 6 (technically this is another lens obtained by the intersection of the two circles of radius REU). The linear distance separating each producer E*m* is a chord $C = 2R_{EU} \sin \alpha$, where $\alpha = 120°$. It follows that $C = 2{,}002.625$ km, which put the center of the radical axis of the lens at a (linear) distance of C/2 from each producer, i.e. 1,001.3 km. With respect to our radius $R_{EU} = 1{,}156.25$ km we thus obtain that our estimate of distance d_E travelled from a "border" location has an upper bias factor of around 15 %.

We are now ready to calculate the increase φ in total distance travelled by one production unit in the EU when we move from a location with *n* internal producers to a location with three "border" producers E*m*. The number is obtained by simply taking the ratio between the total average distance to serve all EU consumers when production is located in three sites at the EU border D_E (minus the 15 % error of our approximation), with respect to the average distance d_n between an internal producer and a consumer.

The computation yields $\varphi = \frac{D_E(1-0.15)}{d_n} = 2.16d_n$.

In other words, if we substitute one internal producer (n = 1) located at the centre of the EU with three external producers located at the border, the total amount of kilometers necessary to serve the same consumers will more than double, more precisely it will increase by a factor of 2.16. Clearly, the higher the number of internal producers that are going to be substituted (i.e. the closer production is to consumers before the delocalization shock), the higher will be the total distance necessary to serve them once production is delocalized.

Note that it is not necessary to suppose direct delocalization for this example to work. For example, imagine that demand of cement in Europe increases in such a way that existing producers could increase capacity as if 10 new producers could

[11] To see why, note that if α = 1 (consumers distribute over the entire circle), then our distance will be the upper bound. $d_B + \alpha\frac{1}{6}R = \frac{1}{2}R + \frac{1}{6}R = \frac{2}{3}R = d_A$

If α = 0 (consumers concentrate along the radius) then we remain with our lower bound dB.

enter the EU, but prefer to serve the extra capacity with production sourced from outside the EU (indirect delocalization), due to limits on regulations. Then, since the new demand is going to be served from outside locations at the EU border, it is as if 10 producers have been displaced, in which case the total transport costs (and the ensuing emissions) with respect to the non-restriction scenario will increase by more than 20 times.

5 Empirical Methodology

How far the asymmetric cost impact depicted above could actually trigger a significant reduction in the turnover and jobs in the EU cement sector depends on the answers to the following questions: is the cost and, thus, price increase significant enough so that clients start engaging in substitution activities, i.e. looking for alternative products or alternative suppliers?

The key to this question lies in the knowledge of price elasticity of demand originating from consumers. In the specific, price elasticity for cement, demanded almost completely by the construction sector. As we have seen in the review of the literature it is reasonable to believe that under the availability of lower costs alternatives (shipping by sea) at least for southern Europe coastal areas the demand for cement becomes more elastic.

In terms of our theoretical model, the presence of low cost alternatives, either in terms of direct imports without delocalization or with offshoring, could shift producers E*m* along the borders of the areas: now consumers have more choice. The empirical analysis focuses on a indirect consequence of the theoretical model that could be added as a corollary: what happens to jobs?

Given the observations above on distances and road transport costs, the question is of uttermost importance for southern Europe coastal areas, as shipping costs by sea are particularly attractive.

Our approach is therefore aimed at estimating the indirect impact on jobs in the cement sector given the increase in environmental costs, by expressing labor demand in the cement industry as depending, given all other factors, also on environmental costs. Let us start by defining:

$$\frac{\Delta L}{\Delta EC} \tag{1}$$

Where L is employment in cement sector and EC are the environmental costs – the carbon dioxide permit costs – that the cement sector should bear from 2012 onwards. The ratio above can be split further into two components:

$$\frac{\Delta L}{\Delta EC} = \frac{\Delta L}{\Delta OSM} * \frac{\Delta OSM}{\Delta EC} \tag{2}$$

where OSM is the share of outsourced (imported) material inputs, a proxy for the relative convenience of substituting local inputs with imported inputs (or, in other words, the weight of offshored intermediate phases of the value chain). The first part of Eq. 2 can be evaluated by turning to the methodology proposed by Amiti and Wei (2004, 2007) and their use of Input – Output tables for various industrial sectors. In particular, what Amiti and Wei (2004) estimate derives from a common empirical specification for labor demand:

$$\ln L_{it} = \alpha_0 + \alpha_1 \ln w_{it} + \beta \ln \omega_{it} + \gamma \ln y_{it}$$

where w is the wage rate, ω is a vector of other input prices, and y is the level of output. In general, an increase in the wage is expected to have a negative effect on employment demand, whereas an increase in the price of other inputs would lead firms to substitute away from the more expensive inputs toward labor. Of course, an increase in output would lead to higher employment.

The question arises as to which input prices to use for outsourcing. We use here the quantity of imported intermediate inputs, that is clinker imports in some EU countries, as an inverse proxy of quantities of imported inputs: in other words, the higher the quantity of imported inputs, the higher the outsourcing intensity. As an alternative, the original work by Amiti and Wei opts for the use of prices of imported inputs. We could therefore use the value of imported clinker imports divided by the quantity, so to get a price effect instead of a quantity effect. Taking first differences of the above equation, we estimate:

$$\Delta \ln L_{it} = \alpha_0 + \alpha_1 \Delta \ln w_{it} + \alpha_2 \Delta \ln OSM_{it} + \beta \Delta \ln \omega_{it} + \gamma \Delta \ln y_{it} + \delta D_t + \varepsilon_{it} \quad (3)$$

where the impact on employment L at time t in country i due to offshoring is given by the coefficient α_2, w are wages, OSM is an index for offshored and imported materials, ω is a vector of other input prices (as capital and electricity), and y is the level of output.

Since the clinker import content of final output cement in the Eurostat Input – Output Table is not available, we turn to the aggregate of imported clinker import, knowing that about 100 % of imported clinker is demanded by the cement sector. We can then expect a decrease in employment L if we observe an increase in outsourcing intensity, that would be equivalent to a decrease in imported inputs prices. In general, an increase in the wage level is expected to have a negative effect on employment demand, whereas an increase in the price of other inputs would lead firms to substitute away from the more expensive inputs toward labor. (We may, notwithstanding, suppose that in capital intensive sector as the cement industry no such a large substitution is possible). Of course, an increase in output would lead to higher employment. We also include a time trend D to control for any unobserved effect common across all countries (think of the rise in oil price).

Moreover, taking the first difference should help controlling for the potential endogeneity of labour demand with output y as well as for any time-invariant country-specific effects such as market or technology differences.

Notice that with respect to Amiti and Wei, our estimation should provide us with a slightly different message. Indeed, in their original estimation they put in relation employment in sectors producing final output with respect to changes in imported inputs. Here instead, the comparison is for closer phases of production, since we want to evaluate how imports of clinker to Europe can displace jobs in the cement sector; imports of an almost – final output matched with jobs in the final output sector. At this level of analysis, unfortunately, we are not able to evaluate if some job creation takes place in other industries.

We expect that as long as prices of imported clinker decrease (quantities of imported clinker increase) with respect to locally (EU) produced cement, the size of the local cement sector decreases, or, in other words, the number of EU businesses engaging in cement production falls.

The estimated coefficient for α_2 provides an estimate for the first part of Eq. 2.

The second part of Eq. 2 is a bit more complicated to estimate, since it requests the availability of price elasticity for cement demand. With respect to the theoretical framework, it would measure the "force" moving cement manufacturers from core areas to the peripheral coastal areas.

Since we can observe the estimates for the rise in environmental costs presented in the studies for the European Commission, we only need to compare these costs with shipping costs from non-EU countries and see how clients' demand for cement (construction sector demand) would change as a result. This would provide us with an indirect estimate for the size of offshored inputs if we consider that cement demand can be proxied by the output in the construction sector.

An intuitive way to solve the issue is to turn to gravity models of international trade. The imports of cement from southern Mediterranean basin to southern Europe coastal countries depend on the value added in the construction sector in Europe, while gravity theory tells us that the closer are the countries, the stronger should be trade flows. Indeed, an increase in outsourced inputs and services in the construction sector was highlighted already in the 1980s, and a growing trade in construction products could back this hypothesis of increased outsourcing; on the other hand, the construction sector is acknowledged as an important source of backward linkages in that it pulls the demand of output of many other economic sectors (Pietroboni and Gregori 2003): as such, the construction sector generally has one of the highest output multipliers among all the sectors of a national economy.

Therefore, we could think of estimating the second part of Eq. 2 by a gravity equation:

$$\ln IMP_{ij} = \mu + v \ln(VA_i) + \pi \ln(VA_j) + \vartheta Dist_{ij} + \varepsilon_{ij} \tag{4}$$

Where *IMP* are imports of clinker-cement from a given set of southern Mediterranean countries and other partner countries to EU countries, *VA* is the value added in the construction sector, which represents the market power index for cement demand, and *Dist* is the distance between the couple of country pairs *i* and *j*. Notice that usually physical distances do not change, but economic distances (trade costs) do. As a consequence, the gravity equation can take into account the change in imports we could observe if economic distances became smaller as marginal costs for producing cement increased, due to the introduction of the EU ETS. Therefore we should evaluate how distance is reduced when the carbon cost increases, which is equivalent to say that the relative convenience of sea transport rises as locally produced cement price increases, and evaluate the change in imports, as this change takes place. This procedure can be thought of as an empirical method to estimate Armington's elasticity.

As we have seen, the Southern Mediterranean countries' proximity to EU importing countries should enter the equation as well: we add therefore a dummy M for each partner country j in the Mediterranean basin to capture the effect of historical and economic linkages in the area.

Estimating this equation should provide us with an estimated coefficient θ (with the opposite sign once interpreted as an increase in costs) that can be considered as a proxy for the second part of Eq. 2. In conclusion, the overall impact on L employment can be estimated as:

$$\ln\frac{\Delta L}{\Delta EC} = \hat{\alpha}_2 + \hat{\vartheta}$$

or

$$\frac{\Delta L}{\Delta EC} = e^{\hat{\alpha}_2 + \hat{\vartheta}}$$

6 Data and Empirical Analysis

Most data are from Eurostat. In particular, wages and employment data come from the Structural Business Statistics (SBS) database, which does not cover the whole EU population of enterprises but according to each state, a given dimensional selection of enterprises (usually small and medium sized firms). These detailed data were adjusted upward by considering the more aggregated corresponding figure for the industry as a whole. The I-O tables were found to be useless to our scope, since the detail does not cover the single clinker use by sector, forcing us to turn to Comext imports data. Data on inputs (clinker, electricity use, physical investment) were all found in Eurostat, with physical investment found under the SBS database. Output data were found in Prodcom. Data on physical (weighted) distance stem from the CEPII online database while sea distance between main ports can be computed on websites like http://www.searates.com/reference/portdistance or; http://www.distances.com/.

Data on exchange rates were found on the European Central Bank website and on the International Financial Statistics online. The time spam covers 1999–2006.

Table 2 Equation 3 estimates: changes in employment as import increase

Set of equations with clinker imports in *volume*

Dependent variable	Employment OLS, first differences	OLS, first differences, White corrected standard errors	OLS, ln-levels, lagged regressors	OLS, ln-levels, lagged regressors, White corrected standard errors	Panel, fixed effects, ln-levels	Panel, random effects, ln-levels
Wages	0.63 (0.13)*	0.63 (0.12)*	0.42 (0.09)*	0.42 (0.10)*	0.14 (0.19)	0.21 (0.11)**
Clinker imports	−0.03 (0.01)*	−0.03 (0.01)**	−0.08 (0.02)*	−0.08 (0.03)*	0.02 (0.01)	0.01 (0.01)
Physical capital	0.03 (0.02)	0.03 (0.04)	0.16 (0.10)	0.16 (0.11)	−0.01 (0.05)	−0.01 (0.04)
Electricity cost	−3.57 (1.59)*	−3.57 (2.42)	2.26 (3.26)	2.26 (1.71)	−7.57 (2.56)*	−6.77 (2.55)*
Output	0.08 (0.10)	0.08 (0.05)	0.30 (0.08)*	0.30 (0.06)*	0.13 (0.15)	0.29 (0.11)*
Constant	−0.04 (0.03)	−0.04 (0.01)*	2.18 (1.04)*	2.18 (0.42)*	6.05 (2.38)*	3.21 (1.43)*
Number of obs.	42	42	43	43	43	43
F test	5.64*	6.99*	18.76*	98.22*	3.28*	Wald 41.71*
R-squared	0.65	0.65	0.81	0.85	0.52	0.67

Set of equations with clinker imports in *value*

Dependent variable	Employment OLS, first differences	OLS, first differences, White corrected standard errors	OLS, ln-levels, lagged regressors	OLS, ln-levels, lagged regressors, White corrected standard errors	Panel, fixed effects, ln-levels	Panel, random effects, ln-levels
Wages	0.52 (0.12)*	0.52 (0.13)*	0.51 (0.09)*	0.51 (0.09)*	−0.08 (0.18)	0.26 (0.12)*
Clinker imports	−0.02 (0.01)*	−0.02 (0.01)	−0.09 (0.02)*	−0.09 (0.03)*	0.02 (0.01)**	0.02 (0.01)**
Physical capital	0.03 (0.02)	0.03 (0.03)	0.13 (0.07)**	0.13 (0.07)**	0.03 (0.03)	−0.03 (0.03)
Electricity cost	−4.12 (1.62)*	−4.12 (2.03)*	2.30 (3.01)	2.30 (1.97)	−4.50 (1.94)*	−4.39 (2.13)*
Output	−0.08 (0.10)	−0.08 (0.11)	0.21 (0.09)*	0.21 (0.07)*	0.02 (0.09)	0.06 (0.08)
Constant	0.02 (0.06)	0.02 (0.05)	4.38 (0.92)*	4.38 (0.56)*	8.75 (1.04)*	6.59 (0.85)*
Number of obs.	56	56	57	57	57	57

(continued)

Table 2 (continued)

Set of equations with clinker imports in *value*						
	Employment					
Dependent variable	OLS, first differences	OLS, first differences, White corrected standard errors	OLS, ln-levels, lagged regressors	OLS, ln-levels, lagged regressors, White corrected standard errors	Panel, fixed effects, ln-levels	Panel, random effects, ln-levels
F test	4.79*	3.48*	13.91*	25.11*	2.27	Wald 21.22*
R-squared	0.62	0.62	0.82	0.82	0.22	0.59

Standard errors in *parentheses*. * = significant at 5 %; ** = significant at 10 %. Time dummies included, almost never significant

Various methodologies were applied in the empirical assessment of Eq. 3, to test the robustness of results. We first estimate a log linear model in first differences by OLS and OLS with heteroskedasticity – consistent standard errors, then we also try a panel approach, with a GLS fixed and random effects model. Results from estimates are reported in Table 2.

We tried a double specification, with clinker imports expressed in both quantities and values. We also adapted the output variable, in terms of quantity and value accordingly.

The change in employment due to an increase in outsourced materials, in this case represented by clinker imports (remember that clinker represent more than 90 % in final cement output) is negative and significant, and robust to the various specifications. As expected, an increase in wages always increases employment and the expected positive relationship between employment and output is found in the specifications.

We included two additional input variables together with clinker imports, without distinction of origin: electricity consumption[12] (as the cement industry is an energy intensive sector, the cost of electricity is relevant) and investment in physical capital, an inverse proxy for the rental cost of capital. The cost of electricity has the expected negative sign most of the times and is statistical significant in about 50 % of the specifications. Instead, investment in physical capital is almost never significant.

We can conclude from these estimates that increased clinker imports displace jobs at a rate that ranges from 1 % to 9 %.[13] This is our interval estimation for the parameter α_2, that is, $\Delta L/\Delta OSM$ in Eq. 2.

[12] The cost of electricity was given with values smaller than one. Therefore, in order to avoid getting negative values when taking the logs, we added one to the series.

[13] The percentage change of a time series Y_t between $t - 1$ and t is approximately equal to $100^*\Delta\ln(Y_t)$; the smaller the percentage change, the more accurate is the approximation. Therefore, we can interpret the coefficient straight as percentages multiplying by 100.

Estimation of Eq. 4 turned out to be trickier. In adopting the gravity equation approach, we should consider that for the model to be complete, the economic dimension of the construction sector should be added for both countries, that is, in the reporting and the partner country.

We set up a database with the four main EU reporting-importing countries (France, Germany, Italy and Spain) and four partner countries (China, Egypt, Turkey and Japan).[14]

The results are reported in Table 3. The most relevant result from this exercise is that the key variable to our discussion, sea distance, is always significant at 5 % and with the correct negative sign. That is, a decrease in distance (relative transport costs) increases imports of cement. Note that, since bilateral import–export of EU countries and third countries is only available at a relatively aggregated level, we had to turn to cement imports rather than clinker to proxy our data. The size of the coefficient ranges from 2.6 to 3.7, and the importance of this finding will be discussed below.

The partners' economic dimension as interpreted by the value added in the construction sector, is never significant at 5 %, with a bit more performing estimates for the importing partner country. From the economic rationale point of view such an odd result can be expected, as we only are estimating an unilateral trade flow and for a single item. Actually, imposing that cements imports from China to Italy depend on the value added of the construction sector in China is quite demanding. On the other hand, there could be a growth effect captured by the value added in the construction sector in the exporters' markets due to the fact that countries as China, India and Turkey are growing quickly and new and efficient plants are being built, to face not only external demand but most of all internal demand. Far better results are in fact obtained when the model is estimated with the classical market power indicators of GDP per capita and GDP. These results lead us to a general conclusion: while the model is less reliable when considering market size as proxied only by the industry-specific value-added as of Eq. 3 (as gravity models have been developed for bilateral trade flows and GDP, not for single industry purposes) the role of distance is in general significant and negative.[15]

This result, although obtained by a procedure that does not take into express consideration the price of the imported good against the price of the locally produced good, as would be required by the Armington's formula, is notwithstanding consistent with the debated and heterogeneous findings about the cement sector, as collected by the EC (2007) (Table 4):

[14] According to AITEC (2006), in 2005 and 2006 the three largest world cement exporters were China, Thailand and Japan.

[15] The value of cement in the construction sector barely represent 5 % of total value added, therefore using valued added of the construction sector as explanatory variable could not explain that largest share of clinker imports.

Table 3 Equation 4 estimates

Set of equations with cement imports in *volume*

Dependent variable	Imports OLS	OLS, White s.e.	OLS	OLS, White s.e.	OLS	OLS, White s.e.	Fixed effects	Random effects
Value added construction country i	−1.97 (1.55)	−1.97 (1.61)	3.92 (2.01)*	3.92 (2.04)*			1.55 (1.49)	−1.97 (1.55)
Value added construction country j	0.19 (0.15)	0.19 (0.15)	0.17 (0.32)	0.17 (0.32)			0.05 (0.13)	0.19 (0.15)
Sea distance	−3.21 (0.49)*	−3.21 (0.49)*	−3.14 (0.94)*	−3.14 (0.88)*	−3.71 (0.98)*	−3.71 (0.93)*	−2.66 (0.44)*	−3.21 (0.49)*
Constant	80.81 (38.87)*	80.81 (39.84)*	43.15 (37.26)	43.15 (36.92)	129.20 (26.20)*	129.20 (25.76)*	−8.93 (37.64)	80.81 (38.87)*
GDP per capita country i			−11.01 (2.50)*	−11.01 (2.56)*	7.42 (4.65)**	7.42 (5.06)		
GDP per capita country j			0.08 (0.56)	0.08 (0.54)	−0.81 (0.43)**	−0.81 (0.49)**		
GDP country i					−7.01 (2.09)*	−7.01 (2.20)*		
GDP country j					1.40 (0.88)**	1.40 (0.90)**		
Number of obs.	100	100	100	100	107	107	100	100
F test	23.52*	23.93*	20.58*	41.61*	19.09*	51.1*	25.85*	Wald 70.57
R-squared	0.4237	0.4237	0.5226	0.5226	0.4859	0.4859	0.3859	0.4237

Set of equations with cement imports in *value*

Dependent variable	Imports OLS	OLS, White s.e.	OLS	OLS, White s.e.	OLS	OLS, White s.e.	Fixed effects	Random effects
Value added, construction, country i	−1.53 (1.27)	−1.53 (1.28)	3.03 (1.66)**	3.03 (1.62)**			1.59 (1.18)	−1.53 (1.27)
Value added, construction, country j	0.06 (0.12)	0.06 (0.13)	0.18 (0.26)	0.18 (0.27)			−0.08 (0.11)	0.06 (0.12)
Sea distance	−2.31 (0.40)*	−2.31 (0.39)*	−2.64 (0.78)*	−2.64 (0.76)*	−3.19 (0.85)*	−3.19 (0.79)*	−1.78 (0.35)*	−2.31 (0.40)*
Constant	69.77 (31.85)*	69.77 (31.78)*	42.56 (30.56)	42.56 (28.69)	117.48 (23.28)*	117.48 (22.03)*	−9.91 (29.79)	69.77 (31.85)*
GDP per capita country i			−8.92 (2.11)*	−8.92 (2.09)*	7.32 (3.85)*	7.32 (4.35)**		
GDP per capita country j			0.35 (0.46)	0.35 (0.43)	−0.51 (0.37)	−0.51 (0.41)		

(continued)

Table 3 (continued)

Set of equations with cement imports in *value*								
	Imports							
Dependent variable	OLS	OLS, White s.e.	OLS	OLS, White s.e.	OLS	OLS, White s.e.	Fixed effects	Random effects
GDP country i					−6.61 (1.75)*	−6.61 (1.96)*		
GDP country j					1.38 (0.75)**	1.38 (0.79)**		
Number of obs.	100	100	100	100	107	107	100	100
F test	20.89*	24.41*	18.48*	33.71*	16.58*	39.4*	24.39*	Wald 62.67
R-squared	0.395	0.395	0.4957	0.4957	0.4509	0.4509	0.3455	0.395

Standard errors in *parentheses*. * = significant at 5 %; ** = significant at 10 %. Time dummies included, almost never significant

In the latter study, the elasticity for the cement sector ranges from 1.09 to 2.4, so that the range of our estimated coefficients (2.6–3.7) falls in almost the same interval.

Let us now suppose that a theoretical reduction in external distances takes place, as hypothesized by the theoretical model with the emergence of offshoring activities, given by the increase in the permit cost, and suppose a mean value of elasticity between the two extreme results, 2.6 and 3.7, say 3.15: should distance theoretically decrease by one sea mile, clinker imports, *coeteris paribus*, should rise by 3.15. A rough estimate for the total effect ΔL/ΔEC should then consider the sum of the two effects. The interval for the global impact is computed as follows:

$$\frac{\Delta L}{\Delta EC} = e^{\hat{\alpha}_2 + \hat{\vartheta}} = e^{0,01+2.6} = 13.6; \frac{\Delta L}{\Delta EC} = e^{\hat{\alpha}_2 + \hat{\vartheta}} = e^{0,09+3.7} = 44.2$$

Taking then the mean values of our estimated parameters α_2 and θ we would obtain:

$$\frac{\Delta L}{\Delta EC} = e^{\hat{\alpha}_2 + \hat{\vartheta}} = e^{0,05+3.15} = 24.5$$

stating that a virtual decrease in distance, due to an increase in pollution permits costs, would decrease employment in the cement sector in France, Germany, Italy and Spain by 24.5 %. The overall impact is partly due to the imports' increase and partly due to the displacement in inputs used in cement production when more imported inputs are used. As our dataset for Eq. 4 is limited, since only four importing and four exporting countries are included, we can think of this estimate as a preliminary but representative result.

Table 4 Estimated Armington elasticities for the cement sector

	GEM-E3 (1)	GTAP (2)	Reinert et al. (3)	Other studies (4)
Cement (as part of nonmetallic)	2.4 (5)		1.09	2

Adapted from EC 2007

(1) Koschel and Schmidt (1998), (2) GTAP (2002), (3) Reinert and Roland-Holst (1992), (4) other studies: Cortes and Jean (1996), and Demailly and Quirion (2007). (5) Referred to as "other energy intensive industries" in the specifications of Armington elasticity values in the standard version of the GEM-E3 model

We should stress how, notwithstanding the experimental estimation (with a modified gravity equation) our approach is theoretically based. Syverson (2004) builds its model to explain the equilibrium presence of different productivity levels across cement plants upon the hypothesis of strong spatial barriers due to high transportation costs. Because of these barriers, spatial differentiation matters in explaining the presence of a varied panorama in terms of plant productivity, that is, there is low substitutability in the industry. Moreover, given these high barriers, demand is supposed to be completely anelastic. Therefore, a sharp reduction in trade barriers can motivate an increase in substitutability and therefore an increase in demand elasticity. This would be the case for cement dealers locating on the sea-side (near ports, for example) whose elasticity to clinker price could increase as they (and they only) have the sea-imports option, as illustrated in our lens-shaped theoretical model. This scenario could lead to a concentration of plants by cement producers in coastal peripheral areas where they could get lower cost clinker imports (causing carbon leakage as already observed) to be transformed in final output (cement) and from there, they could truck the finished product to customers.

This relocation would have a double employment impact on existing cement plants: on one hand, there would be a shift of workforce in the peripheral coastal plants; on the other, clinker production would be partially replaced by imports, causing less efficient plants to close. The outcome of this double trend could justify the sharp reduction in employment estimated by our model.

7 Conclusions

In this paper we have presented the state-of-the-art debate over the potential increase in costs for EU cement producers if a new EU ETS scheme of full auctioning was adopted without exceptions, and an empirical assessment on how this increase in costs could affect employment. The consequential increase in cement production costs for EU manufacturers would trigger a deep restructuring of the industry, with decrease in employment, relocation and offshoring/outsourcing of activities and higher price levels.

In particular, we have estimated an overall negative impact on the industry employment as triggered by an increase in environmental costs of about 24.5 %. The overall impact is partly due to the imports' increase from non-EU producers, and partly due to the displacement in inputs used in cement production when more imported inputs are used. In the theoretical section, we have also showed that the delocalization process would also lead to increased internal transport costs for cement in the EU, and thus higher emissions.

As a result of these, as well as similar calculations, the final provisions of the upcoming EU ETS directive, in which industries deemed exposed to the risk of carbon leakage are granted free allowances for the period 2013–2020, seems to be a balanced and correct outcome.

We hope that the methodology we have tried to construct in this paper could be used as a benchmark exercise for the cement or other industries, should the regulation be revised in the future.

References

AITEC (2006) Relazione Annuale, Italy

Amiti M, Wei S-J (2004) Fear of service outsourcing: is it justified? NBER working paper 10808

Amiti M, Wei S-J (2007) Does service offshoring lead to job losses? Evidence from the United States, mimeo

Armington PS (1969) A theory of demand for products distinguished by place of production. Int Monet Fund Staff Pap 16:159–76

Aubert P, Sillard P (2005) Offshoring and job losses in French industry. Working paper G 2005 7 03, National Institute for Statistics and Economic Studies

Babiker MH (2005) Climate change policy, market structure, and carbon leakage. J Int Econ 65 (2):421–445

Blinder AS (2005) Fear of offshoring. CEPS working paper no. 119, Dec 2005

Boston Consulting Group (2008) Assessment of the impact of the 2013–2020 ETS proposal on the European cement industry. http://www.cembureau.be/topics/climate-change/etd-carbon-leakage, Accessed November 2008

Bruegel Policy Brief (by J. Delgado) (2007) Why Europe is not carbon competitive. Issue 2007/05, Nov 2007

Cembureau (2008) CEMBUREAU's comments to the European Commission's Communication "Further guidance on allocation plans for the 2008 to 2012 trading period of the EU emissions trading scheme" (COM (2005) 703 final). www.cembureau.be, Accessed December 2008

Comext (2008) International Trade Statistics of the European Commission. http://epp.eurostat.ec.europa.eu/newxtweb/

Cortes O, Jean S (1996) Pays émergents, emploi déficient?, CEPII, Document de travail, n° 96–05, Paris

Demailly D, Quirion P (2007) Changing the allocation rules for EU greenhouse gas allowances: Impact on competitiveness, revenue distribution and economic efficiency. European Association of Environmental and Resource Economists 2007 Conference

Ferraz do CLP, Haddad EA, Terra MC (2010) The long run interplay between trade policy and the location of economic activity in Brazil revisited. Fundação Getulio Vargas, Mimeo

European Commission (2007) Imposing a unilateral carbon constraint on energy-intensive industries and its impact on their international competitiveness: data and analysis. European economy economic papers, no 298, Dec 2007, by Manfred Bergmann, Andreas Schmitz, Mark Hayden, Katri Kosonen

European Commission (2008) Impact assessment. Document accompanying the package of implementation measures for the EU's objectives on climate change and renewable energy for 2020. Proposals for DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Directive 2003/87/EC so as to improve and extend the EU greenhouse gas emission allowance trading system, SEC(2008) 85/3

European Commission, McKinsey and Company, Ecofys (2006) EU ETS REVIEW Report on international competitiveness

European Trade Union Confederation (ETUC), Instituto Sindical de Trabajo, Ambiente y Salud (ISTAS), Social Development Agency (SDA), Syndex, Wuppertal Institute (2007) Climate change and employment. Impact on employment in the European Union-25 of climate change and CO_2 emission reduction measures by 2030

Feenstra RC, Hanson GH (1996) Globalization, outsourcing, and wage inequality. Am Econ Rev LXXXVI:240–45

Feenstra RC, Hanson GH (1999) The impact of outsourcing and high technology capital on wages: estimates for the United States 1979–1990. Q J Econ 114(3):907–40

Feenstra RC, Hanson GH (2003) Global production sharing and rising inequality: a survey of trade and wages. In: Choi K, Harrigan J (eds) Handbook of international trade. Basil Blackwell, Malden, pp 146–185

GTAP (2002) GTAP 5 Data Package Documentation, Purdue University

Head K, Mayer T (2000) Non Europe: the magnitude and causes of market fragmentation in the EU. Weltwirtschaftlich Arch 136:285–314

Jensen JB, Kletzer LG (2005) Tradable services: understanding the scope and impact of services outsourcing. Institute for International Economics working paper WP 5–09

JRC and IPTS (2007) Global climate policy scenarios for 2030 and beyond: analysis of greenhouse gas emission reduction pathway scenarios with the POLES and GEM-E3 models. European Commission JRC

Kletzer LG (2000) Trade and job loss in US. Manufacturing, 1979–1994. In: Feenstra RC (ed) The impact of international trade on wages. University of Chicago Press, Chicago, pp 349–396

Kletzer LG (2001) Job loss from imports: measuring the costs. Peterson Institute, Washington, DC

Kolk A, Pinkse J (2008) A perspective on multinational enterprises and climate change: learning from “an inconvenient truth”? J Int Bus Stud 39(8):1359–1378

Koschel H, Schmidt T (1998) Modelling of Foreign Trade in Applied General Equilibrium Models: Theoretical Approaches and Sensivity Analysis with the GEM-E3 Model, ZEW Discussion Paper No. 98–08, Mannheim

Levy DL (2005) Offshoring in the new global political economy. J Manage Stud 42(3):685–693, May 2005

Mani M, Wheeler D (1997) In search of pollution havens? Dirty industry in the world economy, 1960–1995, mimeo. Paper presented at the OECD conference on FDI and environment, The Hague, 28–29 Jan 1999

OECD (2005) The competitiveness impact of CO_2 emission reduction in the cement sector. COM/ENV/EPOC/CTPA/CFA(2004)68/final, prepared by D. and P. Quirion of CIRED, France

Pietroboni R, Gregori T (2003) An input–output analysis of the construction sector in highly developed economies. Constr Manage Econ 21:319–327

Reinert K, Roland-Holst D (1992) Armington Elasticities for United States Manufacturing Sectors. Journal of Policy Modeling 14:631–39

Sanna-Randaccio F, Sestini R (2009) Foreign direct investment and environmental policy: have location factors been neglected? Paper presented at the ETSG 2009, Rome, Eleventh annual conference, 10–12 Sept 2009

Syverson C (2004) Market structure and productivity: a concrete example. J Pol Econ 112 (6):1181–1222

Syverson C (2008) Markets: ready-mixed concrete. J Econ Perspect 22(1):217–234, Winter 2008

Szabó L, Hidalgo I, Ciscar JC, Soria A (2006) CO_2 emission trading within the European Union and annex B countries: the cement industry case. Energy Pol 34:72–87

Immigrant Heterogeneity and Urban Development: A Conceptual Analysis

Jessie Bakens and Peter Nijkamp

Abstract In this chapter we examine the contribution of immigrant heterogeneity to the attractiveness of cities from both the production and the consumption side. Based on an extensive literature review, we hypothesize that the interaction of people from different cultural groups in cities will increase labour productivity in line with the concepts of Jacobs externalities. For the consumption side of the model – a far less researched issue – we hypothesize that urban cultural diversity increases the heterogeneity in the private goods provided, which will increase the utility of living in that area. We argue that future research should focus on the interaction of people from different cultures in the workplace in order to determine urban productivity externalities, and on immigrant-induced product heterogeneity in a city in order to determine immigrant-induced urban amenities. To answer these questions, the use of micro datasets is inevitable.

Keywords Cultural diversity • Urban amenities • Urban productivity

1 Cities and Immigrants

The increasing travel possibilities and lower real costs of transport, information, and communication due to ICT development, have not led to what has been called the ‘death-of-distance’ for economic activities. On the contrary, rather than becoming increasingly irrelevant, location choice seems to be more important than ever.

JEL code: R1, R3, J31, R23

Financial support from NORFACE research programme on Migration in Europe – Social, Economic, Cultural and Policy Dynamics is acknowledged. Furthermore, the authors wish to thank Thomas de Graaff and Peter Mulder for useful comments on earlier versions of this paper.

J. Bakens (✉) • P. Nijkamp
Department of Spatial Economics, VU University, Amsterdam, The Netherlands
e-mail: j.bakens@vu.nl

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_18,

Obviously, the agglomeration of economic activities induces high housing prices and negative externalities such as congestion and pollution (air, noise). So, why then do firms and workers continue to concentrate in urban areas?

In the field of economics, the explanation most often given for the apparent attractiveness of cities is that cities give rise to productivity externalities (see, for example, Glaeser et al. 1992). That is: firms profit from being close to other firms because of information spillovers (Jacobs externalities), the existence of a market for specialized intermediate inputs (Marshallian externalities), or increased efficiency and innovation induced by nearby competitors (Porter externalities). Higher productivity results in higher wages, which attracts workers in search of jobs, which in turn attracts firms wanting to benefit from a large pool of (skilled) workers.

An alternative explanation for the attractiveness of cities is that cities provide a higher number of consumer amenities, making them good places to live. We define amenities as all tangible and intangible consumer-related features of a location that makes it an attractive area in which to live. Researchers such as Glaeser et al. (2001), Glaeser (2010), and Graves (1980) point at (natural) amenity-led growth of urban regions. They stress the importance of climate (weather and natural scenery), but also the broad availability of consumer goods and the presence of restaurants or theatres in larger urban areas that attract people who value those amenities and derive a higher utility from living together in these areas.

As far as consumption is concerned, the death-of-distance hypothesis is proven false, too. Glaeser et al. (2001) find that, although the death-of-distance hypothesis may be true for manufactured goods that can be easily bought via the Internet and shipped all over the world, non-tradable private goods, such as restaurant and theatre visits, but also public goods like schools remain highly local. The increase in reverse commuting (or ‘counter-muting’) – living in the city centre and working in a more suburban area – is a good illustration of the impact urban amenities might have on location decisions.

So both production and amenity externalities can explain why firms or people (workers and consumers) want to locate in urban areas. Ultimately, the challenge is to understand the location choice of workers (consumers) and firms (producers) in their interrelatedness. We know that higher productivity increases wages, and a higher number of amenities increase house prices, since workers are willing to pay for places that provide more consumer amenities. This leads to spatial sorting because people with higher wages can afford to pay higher house prices. We also have some evidence (see, for example, Partridge 2010) that firms increasingly follow people, thereby further increasing productivity in cities. This agglomeration process increases prices, land rents, and wages simultaneously. Clearly, there is an important causality issue here on who follows who in location decisions (workers or firms), and this issue has been extensively investigated in the literature (see, for example, Boarnet 1994; Boarnet et al. 2005; Deitz 1998; Glaeser et al. 2001; Graves 1980; Greenwood and Hunt 1989; Partridge 2010; Steinnes 1977, 1982; White 1999; and Vermeulen and Van Ommeren 2009 for an analysis for the Netherlands).

The economic magnet function of cities makes them a pulling force for international immigrants. We are witnessing a rapid growth of thriving and increasingly culturally diverse cities in the entire world. The spatial clustering patterns of immigrants are so strong that one can reliably predict where new immigrants will locate on the basis of the location of the stock of immigrants, with the same ethnicity, in the destination country (Bartel 1989; Card 2009). One of the reasons for this is the presence of immigrant social networks that can be used to find a job or a house. The presence of other immigrants in an area also creates a network that supports the demand and supply of their home-country products.

On the one hand, these effects are hypothesized to be positive for economic development. Diversity of labour increases productivity because different people tend to 'learn from each other' (Jacobs externality). Immigrants increase the supply of a specific kind of amenity in cities – we will refer to these amenities as immigrant induced (cultural) amenities – that influence the product variety for all people living in that area.

On the other hand, the diversity of people in cities does not come about without some frictions and costs. Too much social variety may lead to interaction problems by increasing the costs of communicating, thus leading to the fractionalization of society (the "Babylon effect": see Florax et al. 2005). Heterogeneity can also generate costs if it results in racism and prejudice (Abadie and Gardeazabal 2003), and if it leads to conflicts of preferences over public spending (Alesina et al. 1999, 2004). Social cohesion in a diverse society might also decline, because the development of trust is more difficult to foster (Alesina and La Ferrara 2002). Putnam (2007) argues that diversity might decrease both bonding and bridging social capital. This means that people living in a diverse society turn neither to their fellow natives nor to other ethnicities, but just refrain from participating in society altogether.

In this chapter we will examine the contribution of immigrant heterogeneity to the attractiveness of cities (for a general immigrant impact analysis see Bakens and Nijkamp 2010). To this end, we will discuss various aspects of the literature on city (consumer) amenities and productivity amenities and the role of immigrants in these processes. In the remainder of this chapter, we will first discuss the impact of immigrant diversity on labour productivity and the various mechanisms through which this relation is established. Then we will discuss amenity-led location behaviour, and the impact a diverse city population might have on the variety of products offered in a city. Finally, we will outline a research agenda, and describe the data needed for further research.

2 Connecting People: Cultural Diversity and Productivity

The general claim that diversity has a positive impact on productivity is based on the work of Jane Jacobs (1961, 1969). She states that the clustering of heterogeneous industries leads to higher economic growth in cities, because ideas or insights from one industry are exchanged and applied in another industry, leading to new ideas and innovations. Glaeser et al. (1992) used these insights to test for agglomeration externalities in cities and found affirmative evidence for Jacobs externalities. De Groot et al. (2009) find in their meta-analytic study on agglomeration externalities that industry diversity (Jacobs externalities) tends to have a positive effect on economic growth in cities.

In research about the impact of immigrants on productivity, the above mentioned Jacobs externalities are adapted to the diversity of workers on the city or firm level. As with the diversity of industries or firms, a labour force[1] consisting of diverse workers will have more diverse ideas and insights which will likely lead to for example more innovations, and better solutions to existing problems (see Ottaviano and Peri 2006). If Jacobs externalities are considered to have an impact on labour productivity as a result of the exchange of ideas, social structures should be the main focus of research. The ability of people to interact in social networks is considered to be their social capital. Following the definition of Akçomak (2009), social capital is created by the relations between individuals. Thus, it is not the human capital individuals possess personally like skills and knowledge that is the focal point of research, but the capital they can acquire in interacting with other people. The social contacts that workers have and the social groups with which they interact are a crucial part of the human capital accumulation and human capital spillovers across people. To our knowledge, not much research within the field of urban economics has been conducted on the social networks between immigrants and native workers in urban areas in order to analyse Jacobs externalities.

These ways of looking at the impact of immigrants on labour productivity is complementary to the neoclassical approach of (im)migration. From the neoclassical point of view, migration is seen as the result of the mobility of production factors and different factor prices between distinct regions. The impact of immigrants on the host country's labour market can then be described quite elegantly: labour markets respond to a positive labour supply shock through the adjustment of wages and employment (i.e. a crowding-out effect if immigrants are substitutes for the native workers).

In a series of meta-analytical studies,[2] Longhi et al. (2005a, b, 2008, 2010) have estimated the average impact of immigration on the labour market by looking at the

[1] In most of the research concerning productivity impacts, the degree of diversity in the labour market is expressed by a fractionalization index (see, for example: Alesina and La Ferrara 2005; Bellini et al. 2012; Ottaviano and Peri 2005, 2006; Ozgen et al. 2010, 2011; Suedekum et al. 2009). A culturally diverse region is then a region in which the probability is relatively high that two randomly selected individuals from the region's population are from a different nationality or cultural background. Although the fractionalization index is most often used in this type of research, one can think of other measures that differentiate among groups and/or regions to indicate concentration, specialisation or inequality (like Herfindahl index or GINI-coefficient, to name a few) that could be applied in this field of research. More on this topic can be read in, for example, Maignan et al. (2003) who compare bio-ecological measures of diversity with economic measures of diversity. In Sect. 4.2 of this chapter we discuss issues concerning the measurement of culture itself.

[2] Meta-analysis requires the acquisition of a cluster of applied modelling studies concerned with the same research question and the use of a common econometric specification, in order to draw a general quantitative synthesis conclusion.

adjustment in wages and employment. The studies show that the impact is, if significant at all, only very small. If the ratio of immigrants to native workers increases by 1 percentage point, the wages of native workers decrease by 0.119 % (Longhi et al. 2005a), and a 1 % increase in the ratio of immigrants to native workers decreases the employment of native workers by 0.024 % (Longhi et al. 2005b). So, in theory, wages and employment of the native workers would, ceteris paribus, decline due to a immigrant influx, but in practice, this does not seem to be the case due to heterogeneity and complementarity among production factors, and due to complex wage adjustments if there is a tight labour market.

In the above mentioned research it becomes clear that the impact of immigrants on the labour market crucially depends on the geographical coverage and scope of the labour market (see also Longhi et al. 2010). Statistically significant negative impacts of immigration occur relatively more frequently in those studies that focus on large geographical labour markets (e.g. a country versus the different regions within that country). When focusing research on a small labour market, the adjustment effects in the economy such as native regional out-migration, changes in sectoral and trade composition, an increase in consumption expenditures, but also impacts on productivity, innovation and social structures, might be better identified with the data (see Longhi et al. 2008, 2009).

In accordance with neoclassical theory, the impact of immigrants on the host country's national labour market thus seems rather trivial, but recently the focus of research on the impact of immigrants has shifted towards a less aggregate level: the differential impact of immigrants on local or regional scales. On a more disaggregate level, migration – or in a more general sense, increasing cultural diversity – may have an impact on socio-economic dynamics in the local or regional economy beyond the neoclassical framework.

Some research has been conducted that focuses on Jacobs-like externalities of immigrants. Based on their research on the US labour market, Ottaviano and Peri (2006) conclude that productivity and economic growth rises when the diversity of workers increases in US cities. When focussing on the relationship between the diversity of nationalities in a population and innovation, Ozgen et al. (2011) find a positive relation between a diverse population and the number of patent applicants beyond a critical treshold. The results indicate that the composition of the immigrant population, i.e. the nationalities of the immigrants, influences the innovative capacity of European regions. Their research shows that a unit increase in the diversity index (see footnote 1), increases patent applications per million inhabitants by 0.2 %.

Suedekum et al. (2009) analysed the impact of immigrant heterogeneity on German NUTS-3 level labour markets. They find that cultural diversity has a positive impact on the German labour market, and particularly highly skilled immigrants increase local labour productivity. Low-skilled immigrants have a negative effect on local wage and employment. An interesting finding in this research is that the largest negative effects of immigrants on German labour markets are to be found in regions with a large and culturally homogeneous group of unskilled immigrants. In the US, Borjas (2005), CCSCE (2005) and

Orrenius and Nicholson (2009) find that, on average (in the long run), the impact of immigrants on the receiving region's employment and wage rates is positive, while only very small substitution-effects between natives and immigrants are observed. The largest negative effect of immigrants is on the wages of earlier immigrant cohorts. In New Zealand (Strutt et al. 2008), the influx of immigrants has caused wages and employment to decrease for some native-born workers and earlier immigrant cohorts for whom new immigrants are close substitutes in the labour market. Research by Econtech (2006) in Australia shows that an increase in the influx of high-skilled immigrants has led to negative transition effects in the short run, but to positive labour market effects in the long run; labour force participation, employment and skill level have increased.

The influence of cultural diversity on labour productivity is thus determined by the possibility for people from different cultures to meet and interact with each other in their workplace, neighbourhood, or school. Some of the abovementioned literature measures the presence of immigrants in areas and concludes that the presence of immigrants increases productivity. But, the presence of immigrants alone does not lead to productivity externalities of the kind described in this section. Interaction among people from diverse cultures or between the firms in which these people work is a requisite for these kind of productivity externalities. This assumes that on an disaggregate micro-level – the neighbourhood and workplace – research should focus on the probability of people from a different culture to meet and interact to determine the influence of cultural diversity on labour productivity.

3 Quality of Life: Cultural Diversity and Consumption

As described above, much research has focused on the impact of cultural diversity on labour productivity. Externalities of cultural diversity, especially in urban regions, can also be directed towards consumption. Although earlier work on agglomeration externalities has (also) focussed on the amenities of product diversity in cities (e.g. Jacobs 1961, 1969; Roback 1982), the interest in consumption amenities in cities has revived in the last decade. This is mainly because some cities in the US and Europe saw an unprecedented rise in prices that cannot be fully explained by an increase in productivity or income (Glaeser et al. 2001; Glaeser 2010). This means that besides the productivity amenities that cities can provide, people are willing to pay high rents for other externalities in cities, in particular quality of life.

In general, in areas with immigrant communities, a supply chain of goods from the home country of these migrants arises. So, the presence of immigrants who demand and/or produce their home-country products increases the product variety in that location. Examples are numerous, but one can especially think of non-tradable private goods, such as Italian or Japanese restaurants, Chinese supermarkets, or ethnic clothing stores. Consequently, the diversity of the goods and services offered in cities with immigrants might increase the utility of living in those cities for all workers.

If the utility of living in heterogeneous cities is higher, then more people may want to live and work there, thus boosting the economic development of the city.

In their research, Glaeser et al. (2001) show that the valuation of amenities in cities has increased over the years. In London and Paris, the amenity premium has risen due to a faster increase in rents than in wages. If wages had risen faster, this would have suggested that the willingness to live in cities is caused by the gain in productivity. Partridge (2010) has indicated that over the past 50 years, migration within the US can indeed best be explained by natural amenities-led growth. The Sunbelt and parts of the Rocky Mountains appear to have attracted relatively many more workers than the largest US metropolitan areas, whereas the former standard core-periphery patterns explained according to the NEG (New Economic Geography)-models have not been confirmed in this research.

The attractiveness of product heterogeneity is also the concept behind the 'love of variety' in Fujita et al. (1999). Other models that look at the willingness-to-pay for amenities or the utility derived from living in a certain area are spatial equilibrium models based on the work of Roback (1982). Spatial equilibrium models inspired by Roback (1982) have been used in different modifications to identify whether the productivity or amenity externality prevails in urban areas, and to calculate the valuation of a specific amenity. The general insight of the model is that the value people assign to city amenities is a premium over the rent minus the wage (see also Glaeser et al. 2001).

The basic Roback model (1982) is a model with many cities that vary according to the quantity of an endowment amenity. Both the indirect utility and the production are a function of wages, rents, and amenities. Wage and rent differences can thus be characterized as functions of the amenity in the equilibrium situation (with a given distribution of firms and workers across cities). The model can be extended by including a non-tradable goods sector. In the initial model, workers are identical in their tastes and skills, but Roback (1988) allows for heterogeneity in individual preferences. Dalmazzo and Blasio (2010) use this model to analyse the valuation of urban amenities among different skill-groups in Italy. They find a substantial urban rent premium, but no urban wage premium, indicating that urban amenities are important drivers for urban agglomeration. This seems especially the case for high-skilled people, who apparently attribute a higher utility to shopping possibilities and cultural consumer goods (theatres, etc.). This is in line with the general view that consumer amenities only become important for location choices after a certain income threshold is reached.

Ottaviano and Peri (2006) took the Roback model as a starting point to model the influence of cultural diversity on production and consumption. Estimating the wage and rent specifications, and controlling for city fixed effects results in positive and significant estimates for cultural diversity (measured by country of birth) in the US. For the simplest model, an increase in the diversity index of 0.1 is associated with an increase in average real wages of US natives of 13 %, and an increase in real rents of 19 %. The same model specification is used in Bellini et al. (2012), but applied to European data where the authors find roughly the same results.

Quigley (1998) lists some more amenities specifically linked to urban diversity. Because of the economies of scale in a city, the provision of public goods, such as parks and sports stadiums, are more easily provided. Shared inputs produce more differentiated goods in theatres, restaurants, fashion, and the like. Search costs for consumers are lower in cities that allow for the agglomeration of products in, for example, shopping districts. And cities are better able to provide substitute goods in cases of economic fluctuations.

In this section, we have shown that theoretical frameworks have been developed that can test whether workers value certain amenities, and whether the urban amenities or productivity amenities prevail in the location choice of workers/consumers. The role of immigrants in providing product heterogeneity has received some attention within the field of economic research, but the majority of the research on migration has focussed on productivity effects. To our knowledge the research that has been conducted does not take into account data of product variation in cities. Thus the implicit assumption in this research is that the presence of immigrants increases product heterogeneity and people value this. Hence, an area that needs further research is the impact of immigrant diversity and immigrant-induced product heterogeneity in cities on the utility of living in that city by using data on product variation.

4 Conceptual Framework and Methodology

4.1 General Framework

In the previous two sections, we have explained that cultural diversity in a city can be seen as an amenity for its inhabitants from a consumer point of view, and that cultural diversity influences productivity by means of human capital accumulation and the exchange of ideas. Figure 1 gives a summary of the causalities and relations described in the previous sections.

Arrow 1 indicates the effect of cultural diversity in a city on product heterogeneity due to the supply of foreign products initially primarily demanded by immigrants. If shops and restaurants arise to sell these products, the product variety increases for all people living in that area. According to the 'love of variety' (Fujita et al. 1999), and the work of, among others, Glaeser (2010), product variety can be seen as a positive feature of a location, and can thus be labelled a urban amenity, as is indicated by arrow 2. Arrow 3 shows this causality and indicates that urban amenities increase the attractiveness of a location.

The productivity side of the model shows that cultural diversity in a city influences the networks and the social capital of people either in the workplace or in the neighbourhood (arrow 4). Arrow 5 shows that this can lead to urban productivity gains if Jacobs externalities arise (Jacobs 1961, 1969; Ozgen et al. 2010, 2011; Ottaviano and Peri 2005, 2006). A more productive city is an attractive area for other people to locate and find a job, as is indicated by arrow 6. Arrow 7

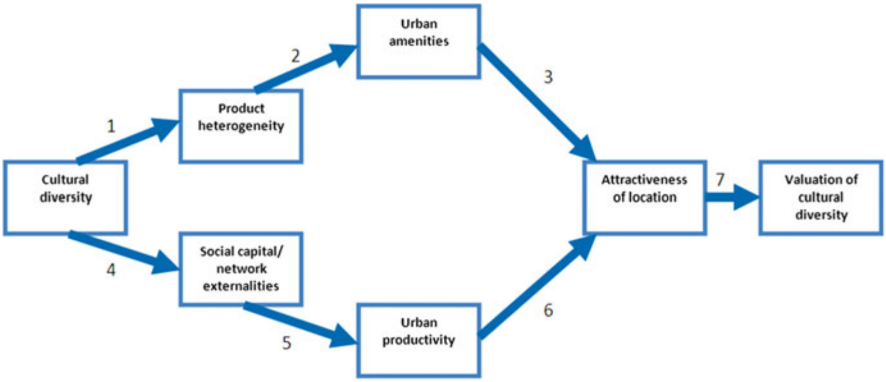


Fig. 1 Relation between cultural diversity and attractiveness of locations

shows that, within this figure, the attractiveness of a culturally diverse location ultimately indicates people's value of cultural diversity.

The different hypotheses stated in the literature and implicitly stated in Fig. 1 are summarized in Table 1. The two main hypotheses that indicate positive effects deduced from the literature are:

1. Interaction of people from different cultural groups in neighbourhoods and/or at work increases productivity.
2. People attribute a higher utility to living in areas with immigrant-induced culturally heterogeneous goods.

In the wage equation, cultural diversity is based on the probability of interacting with people from a different cultural group, and thus on the number of immigrants in the neighbourhood or city and at the firm level. In the rent equation, first the relationship between cultural diversity in an urban area and the number of culturally-induced different non-tradable (private) goods available in the city needs to proved. Next, the valuation of workers of these culturally-induced amenities needs to be calculated in accordance with the valuation of other amenities.

Though these hypotheses cannot considered to be a very novel way of looking at the effects of immigrants on local economies, there are different, empirically novel ways of testing them that will certainly give better and new insights into the described mechanisms. One of these ways is to test the hypotheses simultaneously in a Roback-like model in order to look at the impact of cultural diversity on production and consumption. Another way is to use data on product variety to research the impact of immigrants on amenities in cities, and determine whether these culturally-induced amenities increase the utility of living in these cities. Additionally, the use of micro-data allows for the analysis on different spatial levels to see on which levels (cities, firms, neighbourhoods) these externalities occur. With micro-data one can actually measure whether people from different cultures interact in the workplace or neighbourhood, and whether city inhabitants attribute a high utility to immigrant-induced product heterogeneity.

Table 1 Effects of cultural diversity on production and consumption

	Positive effect	Negative effect
Urban productivity	Jacobs externalities	Babylon effect
	Bridging or bonding of social capital	No bridging and no bonding of social capital
Urban amenities (consumption)	'Love of Variety'	Suboptimal decisions on public good provisions

Based on the effects mentioned in this chapter and on Baycan-Levent (2010)

Using micro-data marks the emergence of a new field within economic research. Detailed 'micro'-administrative databases, sometimes longitudinal, are becoming more and more available and allow research on a far more detailed level. These databases allow us to better describe the adjustment processes at work in, for example, local labour markets, but also other dynamics associated with migration.

Additionally, using micro-data allows for the observation of spatial sorting. When focussing on the location behaviour of individual workers on a neighbourhood level, all kinds of sorting effects can be controlled for. Not only can the amenities be considered, but the negative effects of cultural diversity in neighbourhoods can be addressed, and differences between groups of people in their valuation of amenities can be accounted for, and one can also control for the productivity premiums in different regions.

Below we elaborate on two methodological aspects of this kind of research: the measurement of culture and endogeneity problems, because in the research described in this chapter, in general, these two aspects form a weaker spot or at least allow for intense debate on what is exactly measured. The measurement of culture and endogeneity problems each need more attention in future research from a theoretical, empirical, and methodological point of view.

4.2 Measuring Culture

Other disciplines like sociology and anthropology have a tradition of research on culture and cultural differences on different levels of aggregation.[3] In the field of economics, the impact of 'culture' on economic processes has led to a focus on concepts like institutions (for example, North 1990), trust (Alesina and La Ferrara 2002), or social capital (Putnam 2000). In the research described in this chapter very simple, easy observable, and objective measures of culture are used. One of the reasons is that the analytical tools of economic analysis do not easily allow for endogeneity problems that stem from measures of culture that are less easy to disentangle from other individual characteristics or environmental influences.

[3] It is beyond the scope of this chapter to discuss extensively different definitions and views on culture. We only focus on measuring culture relevant for the impact of immigrant diversity on economic outcomes.

So the question is: what is a meaningful and possible way of measuring culture given how it is hypothesized to affect economic processes? Culture is often referred to as being very persistent over time. Therefore, to distinguish between cultures in research on the effects of cultural diversity on economic outcomes, concepts of culture that are consistent over time like nationality, country of birth, language, race, or religion are often used. These concepts are not used to estimate the impact of different countries of birth on productivity per se, but country of birth is used as a proxy for differences in norms, beliefs and all kinds of unobservable variables that constitute a culture and might lead to economic externalities.

Identifying cultural groups like this has some theoretical and methodological drawbacks. Nationality and country of birth are rather objective and easily obtainable measures. However, nationality is subject to intergenerational erosion making it likely to underestimate diversity in a society.[4] Nationality and country of birth are predominantly self-claimed identities and can change over time. In countries with large groups of immigrants from, for example, former colonies, measuring cultural diversity by nationality might underestimate the effect of immigrants since generally immigrants from former colonies have obtained (or obtain) the host country's nationality quite easily. Whereas language spoken at home (Ottaviano and Peri 2005) is a good indicator for one's identification with a subculture, this might not be the case in all countries, especially with languages like Spanish and English. Falck et al. (2010) show that on the basis of data from Germany, dialects, i.e. language variation within a country, can be a very powerful tool for explaining cultural differences within a country.

This research shows that even a 'simple' variable such as language has different dimensions. Religion and race roughly have the same advantages and disadvantages as language, meaning that, in general, they can be measured reasonably simply, but this can imply, on the one hand a very broad indication (religions and races spread across continents), and, on the other, a very narrow indication (religions have numerous bifurcations and mixtures). The use of social value surveys can however help to identify cultures based on less general measures, but these measures are less objective and quantifiable leading to possible endogeneity bias in estimating the effect of cultural diversity on economic processes. Future research that will further conceptualize the concept of culture from an economically useful point of view would be very welcome.

4.3 The Endogeneity Bias and Unobserved Heterogeneity

In most of the research done in this field, the main source of bias in identifying the effect of cultural diversity on productivity and consumer amenities is endogeneity of independent variables. Whether productive people locate in urban areas or urban areas make people more productive, and whether firms follow workers or workers

[4] Sometimes, country of birth of the parents is used to overcome underestimation of diversity.

follow firms, are all difficult causalities to investigate.[5] When the causality between an independent and dependent variable in a model can also be reversed, for example the correlation between the immigrant heterogeneity in a firm and the productivity of that firm, the independent variable is correlated with the error term. In regression models, this makes the Ordinary Least Squares (OLS) estimator inconsistent and biased. The econometrical solution for the correlation between the independent variable and the error term is to use a randomized controlled experiment (likely not feasible in this setting because of ethical objections), quasi-natural experiment, or an instrumental variable in a two stage-least-square (2SLS) estimation.

The best way to deal with the endogeneity bias is to base research on effects of immigrant diversity on the receiving economy on a quasi-natural experiment. An experiment can be a policy shift, natural disasters or any exogenous shock to the system. A much cited quasi-natural experiment is the Mariel boatlift (Card 1990), in which the impact on the Miami labour market of a sudden increase of the labour force by 7 % by the arrival of about 125,000 Cubans in 1980 is analysed. In chapter 9 of this book "Explicitly Implicit: How Institutional Differences Influence Entrepreneurship", Bauernschuster et al. (2012) use the reunification of Berlin as a quasi-natural experiment to analyse the impact of the difference in implicit institutions between east and west Germany on individual decisions to become an entrepreneur.

Since quasi-natural experiments are often difficult to encounter, many researchers produce interesting instrumental variables to deal with the endogeneity bias. A good instrument is a variable that is uncorrelated with the disturbance term, so not correlated with the dependent variable, but is correlated with the endogenous independent variable. In a 2SLS estimation, each endogenous regressor needs at least one instrument. Time-lagged variables can serve as instruments although one can argue whether these instruments are exogenous. Card (2001) introduced a shift-share methodology for constructing instruments that build on, for example, past shares of immigrants from a certain background in a city to predict the future inflow of these immigrants in the city (as used in Bellini et al. 2012; Ottaviano and Peri 2005, 2006) Ottaviano and Peri (2005, 2006) also use the distance of a city to the main gateways into the US as instruments for the diversity of that city. Ozgen et al. (2011) use the presence of McDonalds restaurants to instrument the openness and international connectedness of regions to deal with the endogeneity bias of immigration. One of the instruments Suedekum et al. (2009) use, is the regional vote shares for the Green party in national elections to proxy for the tolerance level of natives towards foreigners in a region.

The access to micro-data also provides opportunities for research set-ups that can (partly) deal with the endogeneity bias. With longitudinal micro-data, the location decisions of workers can be analysed over time. Secondly, modelling wage and rent equations simultaneously provides some insight into which effect

[5] See for example Combes et al. (2011) on these issues concerning the identification of agglomeration effects.

is stronger: productivity or amenities. Thirdly, focussing on reverse commuting and different aggregate spatial levels might be a good identification strategy for regressing the impact of cultural diversity on productivity or consumer amenities.

It is also important to note that, in research on productivity or urban amenities, there are many worker and city specific characteristics that we do not observe (unobserved heterogeneity) but might explain part of the correlation between cultural diversity on the one hand and urban productivity and amenities on the other hand. It is obvious that controlling for education, income, skill level or any form of human capital of workers when estimating the impact of cultural diversity on productivity, is crucial for the results (see, for example, Alesina and La Ferrara 2005; Bellini et al. 2012; Münz et al. 2006; Ottaviano and Peri 2005, 2006; Ozgen et al. 2010; Pekkala Kerr and Kerr 2008; Saxenian 2007; Strutt et al. 2008; Suedekum et al. 2009). Likewise, cities have certain characteristics besides cultural diversity that make them attractive locations for workers or consumers. When data on these characteristics is unknown (or unobservable), but these characteristics vary across cities and/or workers and are invariant over time, the unobserved heterogeneity can be controlled for by using fixed effects in panel data.

5 Conclusions

The appeal of cities as magnets for productive firms and workers, and as places with a broad product variety for consumers, has encouraged many researchers to examine the processes at work that make cities such attractive locations to settle. International migrants have been shown to have a preference for settling in urban regions, and, in this chapter, we have shed light on the effects of the presence of immigrants on the economic development of urban areas. On the basis of an extensive literature review, we have hypothesized that the interaction of people from different cultural groups at the neighbourhood or firm level in cities will increase labour productivity in line with the concepts of Jacobs externalities and bridging social capital. For the consumption side of the model – a far less researched issue – we have hypothesized that urban cultural diversity increases the heterogeneity in the private goods provided, which will increase the utility of living in that area. This is based on the concept of ‘love of variety’.

By simultaneously measuring the immigrant-induced urban amenity premium and the immigrant-induced productivity premium, the dominant effect can be identified. This is an interesting approach which needs further research, and is a relatively new approach within the economic field of immigrant studies. Future research should also focus more on the amenity effect of immigrants in cities. Research on immigrant entrepreneurship and clustering shows that immigrants tend to supply regions with home-country goods, thus increasing the product variety. The utility that natives assign to these products is an interesting research area. The research agenda laid out in this chapter stresses that, with the use of micro-data, the disaggregate level at which the effects can be analysed is a promising development.

With the use of micro-data, we can explicitly measure consumption amenities by focussing on product varieties, especially immigrant-induced product variety, on a neighbourhood or city level. With micro-data, the interaction between cultural groups in the neighbourhood or firm level can be measured, and thus the likelihood of Jacobs externalities can be better estimated. In this way we are better able to measure the influence of cultural diversity on location behaviour and the valuation of urban cultural diversity.

References

Abadie A, Gardeazabal J (2003) The economic costs of conflict: a case study for the Basque country. Am Econ Rev 93:113–132

Akçomak IS (2009) Bridges in social capital: a review of the definitions and the social capital of social capital researchers, UNU-MERIT working paper, 2009–2002

Alesina A, La Ferrara E (2002) Who trusts others? J Public Econ 85:207–234

Alesina A, La Ferrara E (2005) Ethnic diversity and economic performance. J Econ Lit 43:762–800

Alesina A, Baqir R, Easterly W (1999) Public goods and ethnic division. Quart J Econ 111:1243–1284

Alesina A, Baqir R, Hoxby C (2004) Political jurisdictions in heterogeneous communities. J Polit Econ 112:384–396

Bakens J, Nijkamp P (2010) Lessons from migration impact analysis. Rev Portuguesa de Estudos Regionais 24:5–16

Bartel AP (1989) Where do the new U.S. immigrants live? J Lab Econ 7:371–391

Bauernschuster S, Falck O, Gold R, Heblich S (2012) Explicitly implicit: how institutional differences influence entrepreneurship. In: Crescenzi R, Percoco M (eds) Geography, institutions and regional economic performance. Springer

Baycan-Levent T (2010) Diversity and creativity as seedbeds for urban and regional dynamics. Eur Plan Stud 18(4):565–594

Bellini E, Ottaviano GIP, Pinelli D, Prarolo G (2012) Cultural diversity and economic performance: evidence from European regions. In: Crescenzi R and Percoco M (eds) Geography, institutions and regional economic performance. Springer

Boarnet MG (1994) The monocentric model and employment location. J Urban Econ 36:79–97

Boarnet MG, Chalermpong S, Geho E (2005) Specification issues in models of population and employment growth. Pap Reg Sci 84(1):21–46

Borjas GJ (2005) Foreign-born domestic supply of science and engineering workforce. Am Econ Rev 95:56–60

Card D (1990) The impact of the Mariel boatlift on the Miami labor market. Ind Labor Relat Rev 43(2):245–257

Card D (2001) Immigrant inflows, native outflows and the local market impacts of higher immigration. J Labor Econ 19(1):22–64

Card D (2009) Immigration and inequality. Am Econ Rev 992:1–21

CCSCE (2005) The impact of immigration on the California economy. Centre for Continuing Study of the California Economy, Palo Alto

Combes P, Duranton G, Gobillon L (2011) The identification of agglomeration economies. J Econ Geogr 11:253–266

Dalmazzo A, de Blasio G (2010) Amenities and skill-biased agglomeration effects: some results on Italian cities. Pap Reg Sci. doi:10.1111/j.1435-5957.2010.00327.x

de Groot HLF, Poot J, Smit M (2009) Agglomeration, innovation and regional development: theoretical perspectives and meta-analysis. In: Capello R, Nijkamp P (eds) Handbook of regional growth and development theories. Edward Elgar, Cheltenham, pp 256–281
Deitz R (1998) A joint model of residential and employment location in urban areas. J Urban Econ 44(2):197–215
Econtech (2006) The economic impacts of migration: a comparison of two approaches. Department of Immigration and Multicultural Affairs, Kingston/Sydney
Falck O, Heblich S, Lameli A, Suedekum J (2010) Dialects, cultural identity, and economic exchange. IZA Discussion paper series, Bonn, no 4743
Florax RJGM, de Graaff T, Waldorf BS (2005) A spatial economic perspective on language acquisition: segregation, networking, and assimilation of immigrants. Environ Plan A 37:1877–1897
Fujita M, Krugman P, Venables AJ (1999) The spatial economy: cities, regions, and international trade. The MIT Press, Cambridge
Glaeser EL (2010) The paradox of urban triumph. CPB Lecture, June 2010
Glaeser EL, Kallal HD, Scheinkman JA, Shleifer A (1992) Growth in cities. J Polit Econ 100:1126–1152
Glaeser EL, Kolko J, Saiz A (2001) Consumer city. J Econ Geogr 1:27–50
Graves PE (1980) Migration and climate. J Reg Sci 20:227–237
Greenwood MJ, Hunt GL (1989) Jobs versus amenities in the analysis of metropolitan migration. J Urban Econ 25:1–16
Jacobs J (1961) The death and life of great American cities. Vintage Books, New York
Jacobs J (1969) The economy of cities. Vintage Books, New York
Longhi S, Nijkamp P, Poot J (2005a) A meta-analytic assessment of the effect of immigration on wages. J Econ Surv 19:451–477
Longhi S, Nijkamp P, Poot J (2005b) The fallacy of "job robbing": a meta-analysis of estimates of the effect of immigration on employment. J Migr Refugee Issues 1:131–152
Longhi S, Nijkamp P, Poot J (2008) Meta-analysis of empirical evidence on the labour market impact of immigration. Rég et Dév 27:161–191
Longhi S, Nijkamp P, Poot J (2009) Joint impacts of immigration on wages and employment: review and meta-analysis. J Geogr Syst 12(4):355–387
Longhi S, Nijkamp P, Poot J (2010) Meta-analyses of labour-market impacts of immigration: key conclusions and policy implications. Environ Plan C Gov Pol 28:819–833
Maignan C, Ottaviano G, Pinelli D, Rullani F (2003) Bio-ecological diversity vs. socio-economic diversity: a comparison of existing measures. Fondazione Eni Enrico Mattei, working paper 13.2003
Münz R, Straubhaar Th, Vadean F, Vadean N (2006) Costs and benefits of European immigration. HWWI Policy Report, No. 3, http://hdl.handle.net/10419/46220
North DC (1990) Institutions, institutional change and economic performance. Cambridge University Press, Cambridge
Orrenius PM, Nicholson M (2009) Immigrants in the U.S. Economy: a host country perspective. J Bus Strategies 26:35–53
Ottaviano GIP, Peri G (2005) Cities and cultures. J Urban Econ 58:304–337
Ottaviano GIP, Peri G (2006) The economic value of cultural diversity: evidence from U.S. cities. J Econ Geogr 6:9–44
Ozgen C, Nijkamp P, Poot J (2011) Immigration and innovation in European regions. In: Nijkamp P, Poot J (eds). Migration impact assessment. Edward Elgar (forthcoming)
Partridge M (2010) The duelling models: NEG vs amenity migration in explaining US engines of growth. Pap Reg Sci 89(3):513–537
Pekkala Kerr S, Kerr WR (2008) Economic impacts of immigration: a survey. HBS working paper 09–013
Putnam RD (2000) Bowling alone: the collapse and revival of American community. Simon & Schuster, New York

Putnam RD (2007) E Pluribus Unum: diversity and community in the twenty-first century. The 2006 Johan Skytte Prize lecture. Scand Polit Stud 30:137–174

Quigley JM (1998) Urban diversity and economic growth. J Econ Perspect 12(2):127–138

Roback J (1982) Wages, rents, and the quality of life. J Polit Econ 90(6):1257–1278

Roback J (1988) Wages, rents, and amenities: differences among workers and regions. Econ Inq I 26(1):23–41

Saxenian A (2007) Brain circulation and regional innovation: the silicon valley-Hsinchu-Shanghai triangle. In: Polenske K (ed) The economic geography of innovation. Cambridge University Press, Cambridge

Steinnes DN (1977) Causality and intra-urban location. J Urban Econ 4(1):69–79

Steinnes DN (1982) Do ‘people follow jobs’ or do ‘jobs follow people’? A causality issue in urban economics. Urban Stud 19:187–192

Strutt A, Poot J, Dubbeldam J (2008) International trade negotiations and the trans-border movement of people: a review of the literature. Population studies centre discussion paper no.68, University of Waikato, Hamilton

Suedekum J, Wolf K, Blien U (2009) Cultural diversity and local labour markets. Regional Studies (forthcoming)

Vermeulen W, van Ommeren J (2009) Does land use planning shape regional economies? A simultaneous analysis of housing supply, internal migration and local employment growth in the Netherlands. J Hous Econ 18:294–310

White MJ (1999) Urban areas with decentralized employment: theory and empirical work. In: Cheshire PC, Mills ES (eds) Handbook of regional and urban economics, vol 3. North-Holland, Amsterdam

What Drives Chinese Multinationals to Italy?

Carlo Pietrobelli, Roberta Rabellotti, and Marco Sanfilippo

Abstract This study investigates the motivations driving Chinese outward direct investment to Italy. The analysis is based on secondary sources and in-depth interviews with key informants and senior managers of Chinese affiliates in Italy. The evolution of the Chinese pattern of entry in Italy confirms the pattern followed by Chinese firms in other European countries, adding some additional interesting results. Starting from small-scale operations in trade-related activities, Chinese FDI have evolved towards the acquisition of tangible and intangible resources that are

JEL Classification: F21, P45

This work has greatly benefited from comments from Lin Liu, Martin Chrisney, and participants at the conference on 'Emerging Multinationals' at the Copenhagen Business School, 2008, the 2009 Asialics conference at the Hong Kong Science and Technology University, as well as Chatham House and CASCC research seminars in London, September 2008 and January 2009, Harvard University in October 2010, and in Brussels, June 2009. The authors gratefully acknowledge the collaboration of Thomas Rosenthal of Fondazione Italia Cina who provided useful contacts with Chinese enterprises in Italy, the assistance of Marta Farina with the Chinese-language translations, as well as the support of the MOFCOM branch in Milan, the Bank of China and the Association of Chinese Firms in Italy during the fieldwork. The authors are particularly grateful to the people interviewed for their time and knowledge. Financial support from Compagnia di San Paolo, CASCC (Centro Alti Studi sulla Cina) and PRIN 2007 is gratefully acknowledged.

C. Pietrobelli (✉)
Inter-American Development Bank, Washington, US and CREI, University of Roma Tre, Rome, Italy
e-mail: carlop@iadb.org

R. Rabellotti
Department of Political and Social Sciences, Universitty of Pavia, Pavia, Italy
e-mail: roberta.rabellotti@unipv.it

M. Sanfilippo
Robert Schuman Centre for Advanced Studies, European University Institute, Firenze, Italy
e-mail: marco.sanfilippo@eui.eu

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_19,

deemed necessary to improve China's presence in international markets and to upgrade their technological and production capacities.

Chinese investments in Italy are increasingly targeting the acquisition of technological capabilities and of design skills and brands to tap local competences available in specialized manufacturing clusters in sectors such as automotives and home appliances. The main industries of specialization of Chinese OFDI in Italy reflect this approach and appear to be related to China's strategy to increase the sophistication of its exports and to move away from standardized commodities and intermediate manufactures and components.

Keywords Chinese FDI • Chinese investments in Italy • FDI motivations

1 Introduction

The expansion of Chinese enterprises abroad is occurring at an ever-increasing pace and though this has generated worldwide considerable interest, concern and controversy, too little is still known about why these companies are going abroad. In particular from Europe, Chinese investments are often seen with a mix of fears and hopes. There are reservations because the Chinese State is often behind FDI and because Chinese companies are often backed by political and financial support; moreover, when takeovers are involved, there is a lot of uncertainty about the future survival of companies acquired by Chinese investors, their impact on employment, and the risk of loss of key technological capabilities. On the positive side, the input of fresh capital is very attractive, especially in these times of low growth, and, moreover in the case of acquisitions, often enterprises facing financial difficulties are rescued. Nevertheless, the key point is that these mixed sentiments are mainly based on scanty information, personal interpretations, at the best on case studies of a few well-known companies (e.g. Haier, Huawey) or acquisitions (e.g. IBM by Lenovo and more recently, Volvo by Geely).

In the literature, it has been stressed that the internationalization activity of firms from countries such as China in the developed world reflects attempts to acquire or augment strategic assets, such as new technologies and brands and to secure access to distribution networks, rather than exploiting existing assets, as envisaged by the so-called Ownership-Localization-Internalisation (OLI) model (Dunning 1981). This view has inspired an alternative to the OLI model, the Linkage-Leverage-Learning (LLL) framework proposed by Mathews (2002) to capture the idea that latecomer firms use foreign direct investment and global linkages to leverage their existing cost advantage and learn about new sources of competitive advantages.

Focusing on Europe, the expansion of Chinese investors is taking place in countries with economic systems as diverse as France, Germany, Italy, the Netherlands and the UK. The diversity in terms of sectoral specialization of these economic systems calls for empirical research to investigate the motivations of Chinese investments.

In this chapter, we study what motivates Chinese companies to invest in Italy, being the Italian context of particular interest because of the peculiarity of its economic system, which share with the Chinese economy features such as a strong presence of small and medium enterprises (SMEs) and a specialization in "traditional" industries agglomerated in industrial districts (Amighini and Chiarlone 2005). In terms of stock, in 2009 Italy is ranked sixth among the countries of Europe (excluding Russia and the Luxembourg) with more than 190 million US$, corresponding to a share of 2.2%. Considering flows, these have consistently increased from 2003 to 2007, with just a small decline in 2008 and a considerable recovery in 2009.[1] This trend is also confirmed by a very recent survey undertaken by the China Council for the Promotion of International Trade (2010) in cooperation with the European Commission and UNCTAD on a sample of more 1,300 Chinese companies interviewed with a questionnaire in nearly 30 Chinese provinces. According to this survey, when Chinese companies invest in the EU, they mainly locate in Germany, the United Kingdom, France, and Italy, and they consider the same countries for future investment plans.

For the empirical analysis, we have compiled an original proprietary database from multiple sources, including the whole population of Chinese companies. Moreover, we have undertaken in-depth interviews with key informants and senior managers of some of these Chinese affiliates. This allows us to draw a detailed map of Chinese investments in Italy, exploring their characteristics in terms of size, choice of location, sector of specialization and activities undertaken. With regard to their motivations, we discover that what is happening today mirrors, in the opposite direction, what happened centuries ago when Marco Polo visited China in the XIII Century and lived there for over a decade. The famous Venetian traveler was astonished by the level of civilization achieved by the Asian country and brought back important scientific and technological discoveries, like the use of compass, money and coal. In these days, Chinese enterprises appear to be increasingly interested in acquiring and learning knowledge and technology being developed by companies in advanced countries through a rapid increase of their foreign investments. In Italy, Chinese companies are attracted by the high skills in design-intensive sectors, by the high sophistication of customers, whom global rising stars like Haier would like to learn how to satisfy, by well known brands and by the small size of the enterprises, which makes acquisitions a relatively easy target for the new wave of Chinese companies in their international strategy.

[1] Data about stocks and flows are provided by the Chinese Ministry of Commerce (MOFCOM). These statistics underestimate the real value of investments because they did not include the financial sector up to 2006 and are based on the value arising from approval procedures rather than the effective value of bids (thus excluding non-approved investments and private transactions not formally recorded). In addition, these data do not take into account the investments channeled via offshore (such as Cayman and Virgin Islands) or financial centers (Hong Kong) and thus not officially recorded in Chinese balance-of-payments information. Notwithstanding these limitations, MOFCOM data represent today the most up-to-date source of information on Chinese FDI, disaggregated by destination country and sector.

Moreover, Italian industrial districts with their agglomeration economies also represent an element of attraction, although there are doubts that Chinese companies are really able to become locally embedded and take advantage of the collective efficiency. [2]

The chapter is organized as follows. Section 2 provides a review of the literature on the motivations behind Chinese FDI. Section 3 gives some background information on Italy as a location for FDI. Sections 4 and 5 focus on Chinese FDI in Italy, presenting the findings of our empirical analysis and exploring the motivations of Chinese companies for investing in Italy. Section 6 concludes.

2 The Literature on the Motivations for Chinese FDI

Much of the work that investigates the motivations for FDI refers to the four categories identified by Dunning (1993): resource seeking; efficiency seeking; market seeking and strategic asset seeking. Resource seeking FDI are mainly directed to resource-rich countries, especially in Africa and Latin America. Chinese enterprises have so far had little incentive to seek cheap input factors abroad, and particularly in Europe, given the large domestic supply of low-cost labor, land and capital (Buckley et al. 2008), and their motivation to seek efficiency by exploiting economies of scale and scope and/or securing access to cheaper input factors has been therefore rather weak. Thus, these two motivations are not very relevant in the case of Chinese FDI going to European countries and our focus is on the remaining two driving forces as the main attractors to Europe.

Starting with market seeking investments, the host market size appears as one of the significant determinants of Chinese FDI in an econometric exercise undertaken by Buckley et al. (2007). Therefore, we can expect this motivation to play a positive and significant role also for investments going to the European countries. The literature is also stressing that in the early 1990s most of Chinese FDI were mainly defensive (i.e. FDI following trade) as firms set up foreign affiliates in order to serve their customers better and to increase customer loyalty (Buckley et al. 2008). In the case of FDI towards developed countries, investments are used as a springboard to bypass trade barriers and may be motivated by the attempt to avoid quota restrictions and accusations of dumping products (Luo and Tung 2007). These are, for example, the cases of Haier, which built a manufacturing plant in the US to avoid quota restrictions, and TCL, which acquired an insolvent German television maker, Schneider Electronics, to elude dumping accusations in the EU market. A further reason for defensive market seeking investments is the attempt to escape from the excessive competition at home, given the large number of foreign MNEs in China and the obligation to open the Chinese domestic market under the WTO accession

[2] Collective efficiency is defined as the competitive advantage derived from local external economies and consciously pursued joint action (Schmitz 1995).

terms. This has caused profit margins to fall and resulted in overcapacity in some mature industries, such as textiles and clothing, pushing Chinese firms to find new markets overseas by establishing local sales and distribution centres, but also overseas production bases (OECD 2008). More recently, market seeking investments have also increasingly become offensive (i.e. trade following FDI), aimed at developing new markets, improving brand recognition, adapting products to market requirements and at raising company profiles in markets with growth potential (Buckley et al. 2008).

The other main attractor of Chinese firms to developed countries is access to strategic assets such as technology, know-how, managerial and marketing skills, recognized brand names, distribution networks and reputation. Chinese companies use these investments to rapidly overcome their disadvantages in terms of technology, knowledge and skills (Hong and Sun 2006; Luo et al. 2010). This is also an expressed goal of state-directed Chinese FDI (Deng 2009). Evidence from the UK confirms that the need to acquire new and advanced management skills and to tap into pools of local knowledge is a key reason for Chinese internationalization (Cross and Voss 2008). Further empirical evidence on these motivations is provided by case studies on well-known Chinese MNEs such as Haier, Lenovo, BOE and TCL (Li 2007; Liu and Buck 2009). In a study on the white goods industry, Bonaglia et al. (2007) stress that the success of leading companies from emerging countries such as China and Turkey depends as much on firms' internal resources as it does on the agglomeration advantages of the clusters in which they operate and are embedded. In fact, the choice of the offshore locations is driven both by demand and cost considerations, and by the presence of suppliers of specialized components. Therefore, among the strategic assets Chinese companies are searching abroad we can also include the exploitation of collective efficiency in clusters (Schmitz 1995; Giuliani et al. 2005; Rodriguez-Pose and Comptour, chapter "Evaluating the Role of Clusters for Innovation and Growth in Europe" in this book).

The intensification of cross-border merger and acquisition (M&A) activities by Chinese companies is a confirmation of the importance of the strategic asset seeking motivation (Cui and Jiang 2009). As new international players, Chinese firms generally carry out cross-border M&A primarily to speed up acquisition and control of strategic assets. Based on case studies on companies such as Lenovo, Huawei, Haier and TCL, Deng (2009) and Rui and Yip (2008) analyse the rationale for foreign acquisition activity, emphasizing that it provides a tool to compensate for competitive disadvantage and is a low cost way of leveraging advantages in production capabilities (e.g. the case of Lenovo). In other cases, the firms acquired are loss-making businesses, which are purchased to use their brand names. This, together with the little prior international experience of many Chinese firms and with their cultural distance from western companies, may raise some doubts about their ability to successfully manage the acquired enterprises (Buckley et al. 2008).

There are a number of studies focusing on how cultural factors could significantly affect the location choice and the performance of Chinese companies abroad. Some empirical analyses such as Buckley et al. (2007), Kand and Jiang (2010) and Quer et al. (2011) have introduced various proxies for measuring *cultural proximity* or, rather, *cultural distance* between China and the host countries. Their findings highlight that cultural similarities, such as the presence of a strong Chinese community or the existence of a similar institutional framework, positively and significantly influence the location choice of Chinese MNEs. In particular in their study on Chinese firms in the UK, Cross and Voss (2008) find that that *cultural and language proximity* ranks third among the investment motivation, just behind the market seeking motive and above the asset seeking one. In a survey study investigating the entry mode choice of Chinese MNEs, Cui and Jiang (2009) conclude that Chinese investors perceive cultural differences as barriers to overcome by entering foreign markets through JVs rather than by other entry modes. The cultural distance may result in communication problems affecting human resources management and the ability to understand local markets and to deal with the local business and institutional environment (Boston Consulting Group 2006; Schüller and Turner 2005). Further evidence is provided by company case studies. Thus, for instance, the joint venture between TCL and the French company Alcatel did not prove successful and finally broke up also because of the cultural differences between the two companies were huge (Wu 2005). Similarly, the joint venture between CITIC and Fletcher in New Zealand failed for a range of motivations including inadequateness of communication, lack of mutual trust and scarce cooperation between the two parts due to cultural differences (Ding et al. 2009). In the literature there are also examples of good practices such the well-known cases of Lenovo and Huawei. Rui and Yip (2008) emphasize that the capacity to integrate and combine Chinese culture with world-class Western management systems is the key to the success of these acquisitions. In the case of Lenovo, for instance, the acquisition of IBM PC division was followed by the decision to keep the business in the US and to let it be run by an international management team with at the top position a Chinese American pair of managers (Goldstein 2007).

In the rest of this chapter, the motivations of Chinese companies to invest in Italy are researched in detail. On the basis of the existing findings discussed above, we can expect Chinese companies be attracted by Italy for a mix of market seeking and strategic asset seeking motivations:

- With regard to market seeking: the size of the domestic market, being part of the EU market, the opportunity to learn about market requirements and to improve brand reputation are expected factors of attraction;
- With regard to strategic asset seeking: brand names, design skills, technologies in mature industries and agglomeration economies in clusters are expected to be driving attractors.

3 FDI in Italy

Traditionally, Italy has not been very attractive as a destination for foreign investment. The reasons for this poor performance are many: structural factors such as the fragmentation of the private sector dominated by small and medium sized enterprises, specialization in traditional sectors characterized by low R&D expenditure, and the large size of the public sector have been suggested as discouraging foreign investors (Bronzini 2007; Committeri 2004; Mariotti and Mutinelli 2009; OECD 1994). Other disadvantages include poor infrastructures, high levels of criminal activity in some areas of the country, high levels of bureaucracy and rigid labour market regulation (Daniele and Marani 2008).

According to a survey conducted by the Bank of Italy, among the factors making Italy an attractive location to foreign investors, there are the size of the domestic market and the lower labour costs compared to other EU countries (Committeri 2004). Moreover, another factor driving foreign investors to Italy is the opportunity to exploit the economies of agglomeration in its famous industrial districts, characterised by strong sectoral specialization and by the existence of efficient and competitive networks of suppliers (Becattini 1990). According to Dunning and Lundan (2008), this results in a combination of endowment and agglomeration effects, which in turn lead to location advantages for foreign MNEs. Indeed, some recent empirical works using different sources of data have provided evidence that the strong specialization at the local level and the existence of agglomeration economies are key determinants in the case of inward FDI to Italy. In two different, but related, works Bronzini (2005 and 2007) finds that the sectoral specialization at the provincial level has a strong impact on the inward FDI in Italy. Nevertheless, he finds no evidence of an "industrial district effect" on the location choice of foreign MNEs, attributing this occurrence to the existence of strong linkages between the local productive structure and the social community, factor that could have little or no value to foreign investors (Bronzini 2005). Nonetheless, this last finding could be challenged if we look at the results of other studies. The work by Majocchi and Presutti (2009) shows that foreign firms investing in Italy are attracted in industrial districts, when other foreign enterprises have already invested there. This result is coherent with the more general finding by Mariotti et al. (2010), who use a more disaggregated territorial unit, the local labour systems,[3] finding that foreign MNEs investing in Italy tend to agglomerate with other MNEs to reduce the transaction costs of the investment. Moreover, they also find that foreign MNEs are generally reluctant to agglomerate with domestic firms, unless they hold significant comparative advantages at the sectoral level. Finally, a work by De Propris et al. (2005) shows that the existence of both local industrial systems (defined as local areas characterized by a cluster of highly specialized domestic firms), especially in high tech industries, and the traditional Marshallian industrial districts, especially in low tech industries, significantly affect the location choice of foreign MNEs to invest in Italy. According to these authors, in the former case this is due to the aim of foreign firms to acquire uncodified knowledge, while in the latter case the better division of labour and a stronger performance of industrial districts compared to the rest of the country is an attractive factor for foreign MNEs in lower technology sectors.

[3] The local labour system (LLS) is defined on the basis on information about home-to-work commuting flows (ISTAT 1997).

The Italian specialization in such sectors as automotive, textiles and clothing, machinery and home appliances may indeed be an attractor for multinationals from emerging countries, such as China, which are currently trying to upgrade their production and technological capabilities and build their own global champions in these industries. This is confirmed by the increasing presence in Italy of investors from emerging economies in these sectors as highlighted by Mariotti and Mutinelli (2008), based on firm level data from the ICE-Reprint database. In the next Sections, based on original empirical evidence, we explore the main internationalization characteristics and strategies of Chinese companies operating in Italy.

4 Chinese FDI in Italy

4.1 Methodology

No comprehensive database of Chinese companies operating in Italy is so far available. Therefore, our first task is to compile an original database based on multiple sources, to enable an in-depth study of Chinese OFDI in Italy. The data sources used are *FDIMarkets.com*, previously called Locomonitor, now produced by the Financial Times Group, which is the leading source of intelligence on FDI and provides UNCTAD and the World Bank with data. This database includes information on mode and year of investment, employment, sector, activity and turnover. We also use the *European Investment Monitor* (EIM), or Euromonitor, produced by Ernst & Young on project investments across Europe, and data on M&A collected in the *Zephir* database compiled by Bureau van Dijk. In addition, from 2007 to 2011 we continuously monitored the specialist business press (including *Il Sole 24 Ore*, the main Italian financial newspaper and the *Financial Times*) to check for information on new projects. The list of companies in our database was cross-checked with the assistance of MOFCOM representatives in Milan and with the President of the Association of Chinese Firms in Italy.

As a result, we have identified 78 Chinese investment projects in Italy, including five ceased investments and two cases of relocation. Our database includes information on company name and address, parent company name, sector of specialization, main activity undertaken in Italy, number of employees, total sales, year of the investment, entry and ownership modes. Whenever possible, this information is complemented by additional sources such as company documentation available on the Internet (including main company websites in Chinese), research papers and press articles.

After building our database, we have conducted face-to-face interviews with ten senior managers of Chinese affiliates in Italy, based on a set of open-ended questions focusing on background information about the company, its strategies of internationalization, motivations for investment in Italy and opinions on the Italian business environment. Besides, we have interviewed key informants such as

Table 1 Employees in Chinese companies in Italy

	1986–94	1995–99	2000–04	2005–10	n.a.	Total
<10	3	2	9	8	3	25
11–49	1	1	7	12	..	21
50–99	..	1	1	3	1	6
>100	..	1	2	4	1	8
Total	4	5	19	27	5	60

Source: Author's database

MOFCOM representatives in Italy, senior managers of the Bank of China, managers of the main Italian law firm dealing Chinese FDI.

4.2 *Characteristics of Chinese FDI in Italy*

The first Chinese "flagship" investment in Italy occurred in 1986 when Air China opened a commercial office in Rome (AT Kearney 2008). From the mid 1980s to the end of the 1990s investments were sporadic, and included an office in Turin of the Nanjing Motor Corporation, a commercial office of Cemate Machinery Technology and a branch of the Bank of China in Milan. The majority of Chinese FDI in Italy occurred after 2000, and reveals a recent but rapidly increasing interest. The available information shows that the majority of Chinese companies located in Italy also have investments in other European countries,[4] confirming that the decision to invest in Italy is usually part of a broader European strategy.

In terms of investment size, data on employment is available only for 60 of the 78 companies. Table 1 shows that most companies are small or very small, which is in line with the results of the survey by Cross and Voss (2008) in the UK that the majority of Chinese operations have less than 25 employees. Note that, with two exceptions, the few companies that employ more than 50 people were all established during the 2000s.[5]

With regard to sectoral specialization (Table 2), the main sectors are household appliances and automotives, both industries in which Italy traditionally has strong production capabilities and in which China is rapidly increasing its competitiveness (Amighini and Chiarlone 2005). A further important sector is Transport and Logistics, whose relevance is due to the geographical position of Italy as a hub for the Mediterranean. Italy also represents an important market for electronics and telecommunications, which are other attractive industries for Chinese investors.

Geographically, Chinese investments are strongly concentrated in the North of Italy (Table 2). The region of Lombardy hosts 36 investments, 24 of which in the

[4] We checked for the existence of subsidiaries in other countries in the databases and in the Chinese websites of the parent companies.

[5] For 22 companies, data on revenues and assets are available from the AIDA database and confirm the small size of Chinese companies in Italy.

Table 2 Sectoral and geographical distribution of Chinese FDI in Italy

	Lombardy	Piedmont	Veneto	Lazio	Emilia	Rest of Italy[a]	Total
White goods	5	1	5	..	..	2	13
Automotive	..	5	..	2	..	1	8
Transport and logistics	2	..	1	1	..	4	8
Trade services	3	1	..	..	1	..	5
Textiles	5	1	..	..	..	..	6
Electronics	5	..	..	1	..	1	7
Telecommunications	1	1	..	2	..	..	4
Metal products	4	1	..	..	..	1	6
Machinery	4	1	..	..	1	..	6
Chemical products	2	..	1	..	1	..	4
Financial services	3	..	..	..	..	..	3
Others[b]	2	..	2	1	2	1	8
Total	36	11	9	7	5	10	78

Source: Author's database

[a]Includes Campania, Liguria, Marche, Tuscany, Puglia, Basilicata and Calabria

[b]Include other manufacturing (non-metallic minerals, bicycles, jewellery, toys, pens); food and tobacco and a diversified group

metropolitan area of Milan, the favourite destination of Chinese firms, reflecting the general attractiveness of this region which hosts half of total FDI projects in Italy (Mariotti and Mutinelli 2009). Milan is particularly attractive to firms in the service sector. The only Bank of China branch in Italy, established in 1998, is in Milan, and a second branch is planned for the heart of the city's Chinatown. Also in January 2011 the Industrial and Commercial Bank of China (ICBC), the largest bank in the world in terms of stock market capitalization, has opened its first Italian branch in Milan. Consulting firms set up to assist Chinese companies wanting to invest in Italy are also present. Among them, since 2007, the China Milan Equity Exchange (CMEX) operates as sole partner of CBEX (China Beijing Equity Exchange) in Europe, providing comprehensive advice on legal, fiscal, financial and organizational issues. The second Italian region attracting Chinese FDI is Piedmont. Due its traditional manufacturing specialization in the automotive sector, most Chinese investments are in this industry. Investments in other regions occur in different sectors of specialization, namely white goods in Veneto, machinery in Emilia Romagna and logistics in Campania and Liguria.

The disaggregation of investments by main activity offers some interesting insights (Table 3). In line with what has happened in other European countries (Hay et al. 2009), while in the past the prevailing activity was establishment of sales and marketing offices, investments in higher value added activities have recently increased, especially manufacturing and, to a lesser extent, R&D. Furthermore, key informants suggest that traditional trade-related investments are evolving towards more sophisticated services, such as the search for new markets and the acquisition of new brands.

Table 3 Chinese OFDI in Italy per activity

	1986–94	1995–99	2000–04	2005–10	n.a.	Total
Sales and marketing	6	4	14	16	6	46
Manufacturing	..	..	7	12	2	21
Headquarters	..	..	1	2	..	3
R&D	..	..	1	3	..	4
Logistics and distribution	..	2	1	..	1	4
Total	6	6	24	33	9	78

Source: Authors' database

Table 4 Main M&A operations by Chinese firms in Italy

Year	Target	Acquirer	Sector	Size (employees)	Stake (%)
2001	Meneghetti	Haier	White goods	100	100
2004	Wilson	Wenzhou Hazan	Textiles	..	90
2005	Benelli	Quianjiang	Automotive	100	100
2006	Elios	Feidiao electrics	White goods	54	..
2007	Hpm Europe Spa	Hunan Sunward Intellingent Machinery	Machinery	6	51
2007	Omas Srl	Xinyu Hengdeli Holdingns	Luxory goods	48	90
2008	Cifa	Changsha Zoomlion	Machinery	70	60
2009	Elba	Haier	White goods	150	20

Source: Author's database

The entry mode of Chinese investments in Italy has evolved gradually. The first wave of investments in representative offices was characterized mainly by small-scale greenfield investments; this has developed to larger greenfield investments directed to activities such as R&D and marketing. Since 2000 there has been an increase in M&A including the acquisition of the motorcycle manufacturers Benelli by Quianjiang, of Meneghetti and Elba by Haier, and the takeover of Cifa, specialized in the production of machinery for the construction sector by Zoomlion, which so far is the largest acquisition in Italy and one of the largest in Europe (Table 4).

The evidence presented on Chinese FDI in Italy confirms some of the findings of the existing studies on the UK (Cross and Voss 2008; Liu and Tian 2008), France (Nicolas 2010) and Germany (Schüler-Zhou and Schüller 2009) in terms of investment size, activities undertaken and mode of entry. The next section, based on first hand information from the interviews, will put the Italian case in perspective, exploring the main motivations that push Chinese companies to invest in the country.

5 Why are Chinese Companies Investing in Italy?

Chinese investments in Europe reflect a sustained effort to enter competitive European markets and get access to superior technologies, know-how and competence (Filippov and Saebi 2008; Hay et al. 2009; Nicolas 2009). Our research on Italy confirms that the main motivations for Chinese investments in the country are the search for new markets and other trade-related activities (market-seeking), and the search for strategic assets. In the latter case, given the peculiarity of the Italian economic system, we find that Chinese FDI in Italy look for brands, access to advanced skills and technological capabilities in specialized clusters. Moreover, we also find that Chinese MNEs in Italy do search to minimize the cultural distance by searching for local advanced managerial capabilities but they also face some cultural problems when trying to take advantage of agglomeration economies in clusters.

5.1 Market-Seeking Investments

With regard to market-seeking investments, Italy represents an important market for foreign investors as it is the seventh largest economy in the world and it is part of the European Union. For instance, in a sector such as telecommunications, Italy boasts one of the highest rates of mobile penetration in Europe. Huawei has established three subsidiaries in Italy: in Rome, Milan and Turin. As confirmed by the managers interviewed, the size and potential of the market has been highlighted as the most important factor affecting Huawei's decision to invest in Italy. Having started out as a distributor for global MNEs in the Chinese market, Huawei's globalization strategy started from neighbouring countries before entry into Russia and Africa. Its extension to more sophisticated markets is designed to raise its international profile (Simmons 2008). Since 2000, the company has set up several high value added activities including R&D, training and design, in several European countries such as Sweden, the Netherlands, France and Germany, and has established its regional headquarters in the UK. In Italy, Huawei has invested to seek a large market and raise its profile and strengthen its brand, but also to conduct research and product development activities in its recently established research centre in Turin.

Overall, the prevalence of market-seeking investments is confirmed by Table 4: 46 projects are in market related activities, such as trade supporting services or marketing offices. Some of these investments have been established to better serve customers and to strengthen loyalty, in other words "following trade". This is the case for the state owned trading company, Temax, which opened an import–export office in Milan in 1991. Some important investments in the logistic sectors also aim at supporting trade. With the rise in Chinese exports, the main logistics companies have begun to invest in Europe initially through joint ventures and strategic alliances with local enterprises, while establishing representative offices through greenfield investments. Having acquired new capabilities and market power, Chinese companies are keen to strengthen their positions by acquiring European companies and investing in new infrastructure projects (Hay et al. 2009). This pattern applies to Chinese investments in the Italian logistics sector: COSCO (China Ocean Shipping Group) and China Shipping Company, both of which are in the top ten world shipping companies, both invested heavily in Italy.

The "trade following" category includes more recent investments such as the opening in 2008 of a sales office in Milan by Suntech Power Holdings, the world's largest photovoltaic module manufacturer. As confirmed by one of our key informers, Suntech has invested in Italy motivated by the large growth potential offered by the Italian market, sustained by the economic incentives provided by government. Similarly, Hisense, a large company producing home appliances, invested in Italy to strengthen the company's position in Europe, improve product image and promote its brand. Several key informers have underlined that Chinese companies consider Italian consumers to be highly demanding and particularly sophisticated. Therefore, in sectors such as home appliances the Italian market is regarded as a test market for products adapted to European tastes and it is considered strategic to obtain product feedbacks.

5.2 Strategic Asset Seeking Investments

According to our key informants, the direct contact with the market jointly with the ambition to acquire strategic assets in design, manufacturing and management are the main factors of attraction to Italy of Haier, the Chinese giant specialised in the white goods sector. Haier is the second world producer after Whirlpool, which first entered the Western market as an original equipment manufacturer (OEM) exporter. In 2000 Haier Europe was established in Varese to coordinate sales and marketing across 13 European countries (Duysters et al. 2009). In 2003, Haier made its first acquisition in Italy buying Meneghetti, a refrigerator producer and in 2009 it acquired another Italian company, Elba, which produces cooking appliances. These acquisitions were motivated on the one hand by the need to overcome EU tariff barriers and on the other by the objective to improve the capacity to design, develop and manufacture products suitable for the European market, and for the high end of the Chinese import market (Liu and Li 2002). Moreover, the intention to acquire knowledge and managerial capacity was behind the decision to locate the headquarters in Varese, given the area's strong tradition in white goods manufacture. In fact, Varese is well known for its white goods production and is home to important companies such as Philips and Whirlpool, and many other firms specialized in components and intermediate products. The agglomeration of many specialized firms generates positive externalities, arising from the presence of a pool of specialized workers and suppliers and by specialized knowledge on markets and technologies. These agglomeration advantages attracted Haier and influenced its decision to establish its European headquarters there, as confirmed by Bonaglia et al. (2007) in their study on the global white goods sector.

The existence of a specialised automotive cluster concentrating all different phases of the production process is also behind the decision of two Chinese automotive companies, Jac Anhui Janghuai and Changan, to invest in Turin. In 2004 and 2005, the two companies established R&D and design centres in Turin, where Chinese researchers are working together with their Italian counterparts in strict collaboration with other local specialized firms and research organizations. In both cases, the target is to improve technical know-how, with a particular emphasis on design skills. Both companies are 'newcomers' to the global automotive market and see investment in Italy as a rapid and efficient way to improve their capabilities in design and product development. Compared to other possible locations such as Germany or the UK, the Turin cluster offers the advantages of excellent design skills, availability of highly qualified and cheaper human resources and a pool of specialized suppliers for outsourcing a wide range of activities including engineering, modelling, prototyping and mathematical analysis and calculation.

Other relevant strategic asset-seeking investments are those aimed at the acquisition of well-known brands. This is a strategy followed by many emerging country MNEs, given the unfamiliarity of their home brands in foreign countries (Makino et al. 2002). Due to their specialization in the lowest value-added activities in global value chains: "becoming original design manufacturers (ODMs) and further progressing into original brand manufacturers (OBMs), either through the firm's own efforts or through brand acquisitions from incumbents, is hence the most difficult phase for any latecomer or newcomer MNE" (Bonaglia et al. 2007, p. 8).

In Italy, the acquisition of recognized brands is a common motivation for Chinese FDI. There are several cases such as the acquisition of Wilson by Wenzhou Hazan, one of the main footwear producers in China, which has maintained design and production in Italy to produce shoes to export to the Chinese market. Another example is the case of Elios, an Italian company producing electrical items such as lamp holders, which was acquired in 2006 by Feidiao and the acquisition in 2007 of Omas, a producer of luxury pens established in 1925 in Bologna, by the Xinyu Hengdeli Group, a trading company linked to LVMH selling luxury goods in the Asian market. An important acquisition was made in 2005 by the Quianjiang Group, China's largest scooter manufacturer, with Benelli, an established motorcycle producer that, at the time of the acquisition, was in serious financial troubles. In this case, besides the willingness to acquire a historic and world famous brand, the deal aimed at getting access to and leveraging from Benelli's manufacturing and R&D facilities, and made it Quianjiang's European R&D centre for the production of high-quality production. Notwithstanding the original aim, Mucelli and Spigarelli (2010) report of several problems following the acquisition of the motorcycle company Benelli by Quianjiang. These include the different perception of fiscal rules, language barriers and different working practices, which led to continuous misunderstanding between the Chinese and Italian employees, delaying the development of some new projects managed in between the Italian and the Chinese facilities.

A further area of competence that Chinese companies, particularly medium-sized firms with little international know how, are seeking abroad is managerial experience. Chinese companies with no international experience often find it hard to deal with Western management models and therefore invest in the creation of close linkages with Western managers for reducing the cultural gap. In Italy, a case in point is the Hengdian Group (HG), a family-owned business established in 1975 in the Zhejiang region, which has opened its first European subsidiary in Milan. According to its managing director, although HG ranks third among private enterprises in China with a very diversified business in industries such as electronics, pharmaceutical, film and entertainment, it is still a very local firm with little experience of even the Chinese market outside its home region, and no international experience. The reason for opening a European branch is to gradually learn the marketing skills required for exporting, and to identify new potential areas for investment, particularly related to post sales assistance and customer care. Its management lacks international experience, and the Italian managing director, who has a personal, long term, relationship with the son of the founder of HG, is playing a key role in transferring Western management culture to the Chinese managers in the group. Several key informers have recognized that the case of HG is very promising for the capability of Italy to attract a new wave of medium-sized Chinese companies, latecomers in the international market. These companies may be particularly attracted by the small size of Italian companies, given that these acquisitions will also be facilitated in the future by the recent changes in Chinese regulation of FDI, aimed at extending the facilitation of the "Go Global" policy beyond large companies to include small and medium enterprises.

6 Conclusions

This chapter provides new empirical evidence based on in-depth interviews aimed at contributing to understand the motivations of China's presence in Europe. The evolution of the Chinese pattern of entry in Italy confirms the pattern followed by Chinese firms in other European countries, adding some additional interesting results. Starting from small-scale operations in trade-related activities, Chinese FDI have evolved towards the acquisition of tangible and intangible resources that are deemed necessary to improve China's presence in international markets and to upgrade their technological and production capacities.

Chinese investments in Italy appear to reflect a "Marco Polo" effect, but in the opposite direction: like the Venetian merchant discovered, learnt and returned with the scientific and technological discoveries of the XIII Century China, today Chinese companies are seeking the original skills and knowledge available in Italian companies and localities, especially in design-intensive, high-quality productions. Thus, they are increasingly targeting the acquisition of technological capabilities and of design skills and brands to tap local competences available in specialized manufacturing clusters in sectors such as automotives and home appliances. They link with Italian firms, learn from them and leverage upon their skills and capabilities. The main industries of specialization of Chinese OFDI in Italy reflect this approach and appear to be related to China's strategy to increase the sophistication of its exports and to move away from standardized commodities and intermediate manufactures and components. Italy is considered a key market for investment because of its size and, especially, sophisticated demand. Gaining knowledge about how to satisfy very demanding customers in terms of design, style, branding, marketing and post-sales assistance is what Chinese companies are keen to learn from their activities in Italy and, particularly, in Italian industrial districts known globally for their production and design excellence, for the density of agglomeration economies and the competitiveness of the supplier networks. These are for instance the very successful cases of the specialized automotive cluster in Turin and the home appliances district in Varese, where Chinese companies have established R&D and design centres and headquarters to absorb foreign technology and improve their technical know-how, especially in design skills, and to benefit from the pool of specialized suppliers to outsource a wide range of activities ranging from engineering to modelling and prototyping. The peculiarities of the Italian model of specialization, in terms of both sectors and size of firms, appear to be important attractors for Chinese investments, following a novel "Marco Polo" effect to acquire and leverage upon foreign knowledge and capabilities.

Nonetheless, there are also several signs showing that the acquisition of uncodified capabilities and the exploitation of local agglomeration economies are not at all automatically obtained with the location in industrial districts, where there are old established strong linkages between the local productive structure and the social community. How are Chinese MNEs trying to increase their degree of embeddedness in the local context where they operate, and in particular in local innovation networks is a key issue for future empirical research.

References

Amighini A, Chiarlone S (2005) Rischi dell'integrazione commerciale cinese per il modello di specializzazione internazionale dell'Italia. Riv di Polit Econ 95(7–8):63–86

AT Kearney (2008) Indovina chi viene a cena? Imprese Russe, Indiane e Cinesi in Italia. AT Kearney, Milan

Becattini G (1990) The Marshallian industrial district as a socio-economic notion. In: Pyke F, Becattini G, Sengenberger W (eds) Industrial districts and inter-firm cooperation in Italy. International Labour Office, Geneva, pp 37–51

Bonaglia F, Goldstein A, Mathews J (2007) Accelerated internationalization by emerging multinationals: the case of white goods sector, MPRA paper no. 1485, Munich

Boston Consulting Group (2006) The new global challengers- how 100 top companies from rapidly developing economies are changing the world, BCG report

Bronzini (2005) Industrial Districts, Agglomeration Economies and FDI in Italy, in Bank of Italy, Local Economies and Internationalization in Italy, Rome. Published in Italian. In: Signorini LM, Omiccioli EM (eds.) Economie locali e competizione globale, Bologna

Bronzini R (2007) FDI inflows, agglomeration and host country firms' size: evidence from Italy. Reg Stud 41(7):963–978

Buckley PJ, Clegg J, Cross AR, Liu X, Voss H, Zheng P (2007) The determinants of Chinese outward foreign direct investment. J Int Bus Stud 38:499–518

Buckley PJ, Cross AR, Tan H, Xin L, Voss H (2008) Historic and emergent trends in Chinese outward direct investment. Manage Int Rev 48(6):715–748
China Council for the Promotion of International Trade (2010) Survey on current conditions and intention of outbound investment by Chinese enterprises. China Council for the Promotion of International Trade, Brussels
Committeri M (2004) Investire in Italia? Risultati di una recente indagine empirica, Banca d'Italia, Temi di Discussione no 491, Rome, March
Cross AR, Voss H (2008) Chinese direct investment in the United Kingdom: an assessment of motivations and competitiveness. Paper presented at the conference corporate strategies in the New Asia. University of Bremen, Germany, 1–2 Feb 2008
Cui L, Jiang F (2009) FDI entry mode choice of Chinese firms: a strategic behaviour perspective. J World Bus 44(4):434–444
Daniele V, Marani U (2008) Criminalità e Investimenti Esteri: Un'analisi per le province italiane, MPRA working paper no 6632
De Propris L, Driffield N, Menghinello S (2005) Local industrial systems and the location of FDI in Italy. Int J Econ Bus 12(1):105–121
Deng P (2009) Why do Chinese firms tend to acquire strategic assets in international expansion? J World Bus 44(4):74–84
Ding Q, Akoorie MEM, Pavlovich K (2009) Going international: the experience of Chinese companies. Int Bus Res 2(2):148–152
Dunning JH (1981) International production and the multinational enterprise. Allen & Unwin, London
Dunning JH (1993) Multinational enterprises and the global economy. Addison-Wesley, Wokingham
Dunning JH, Lundan SM (2008) Multinational enterprises and the global economy, 2nd edn. Edward Elgar, Cheltenham
Duysters G, Jacob J, Lemmerns C, Jintian Y (2009) Internationalization and technological catching-up of emerging multinationals: a comparative case study of China's haier group. Ind Corp Ch 18(2):325–349
Filippov S, Saebi T (2008) Europeanisation strategy of Chinese companies: its perils and promises, University of Maastricht, UNU-Merit working paper series no. 2008–055
Giuliani E, Pietrobelli C, Rabellotti R (2005) Upgrading in global value chains: lessons from Latin America clusters. World Dev 33(4):549–573
Goldstein A (2007) Multinational companies from emerging economies. Composition, conceptualization and direction in the global economy. Palgrave Macmillan, London
Hay F, Milelli C, Shi Y (2009) Chinese firms enter Europe: some empirical evidence. In: Wu Z (ed) Financial sector reforms and the international integration of China. Routledge, London, pp 198–214
Hong E, Sun L (2006) Dynamics of internationalization and outward investment: Chinese corporations' strategies. China Quart 187:610–634
ISTAT (1997) I sistemi locali del lavoro 1991. Istituto nazionale di Statistica, Rome
Kand Y, Jiang F (2010) FDI location choice of Chinese multinationals in East and Southeast Asia: traditional economic factors and institutional perspective. J World Bus. doi:10.1016/j.jwb.2010.10.019 (in press)
Li PP (2007) Toward an integrated theory of multinational evolution: the evidence of Chinese multinational enterprises as latecomers. J Int Manage 13:296–318
Liu H, Li K (2002) Strategic implications of emerging Chinese multinationals: the haier case study. Eur Manage J 20(6):699–706
Liu L, Tian Y (2008) The internationalisation of Chinese enterprises: the analysis of the UK case. Int J Tec Globalisation 4(1):87–102
Liu X, Buck T (2009) The internationalization strategies of Chinese firms: Lenovo and BOE. J Chinese EconBus Stud 7(2):167–181
Luo Y, Tung RL (2007) International expansion of emerging market enterprises: a springboard perspective. J Int Bus Stud 38:481–498

Luo Y, Xue Q, Han B (2010) How emerging market governments promote outward FDI: experience from China. J World Bus 45(1):68–79

Majocchi A, Presutti M (2009) Industrial clusters, entrepreneurial culture and the social environment: the effects on FDI distribution. Int Bus Rev 18:76–88

Makino S, Lau CM, Yeh RS (2002) Asset – exploitation versus asset seeking: implication for location choice of foreign direct investment from newly industrialized economies. J Int Bus Stud 33(3):403–421

Mariotti S, Mutinelli M (2008) Le Multinazionali dei Paesi Emergenti in Italia. Econ Polit Ind 35(3):181–187

Mariotti S, Mutinelli M (2009) Italia multinazionale 2009. Rubettino, Rome

Mariotti S, Piscitello L, Elia S (2010) Spatial agglomeration of multinational enterprises: the role of information externalities and knowledge spillovers. J Econ Geog 10:519–538

Mathews JA (2002) Dragon multinationals – a new model for global growth. Oxford University Press, Oxford

Mucelli A, Spigarelli F (2010) Integration problems in Chinese overseas M&As: a case study approach, paper presented at the second 'Emerging multinationals' conference, Copenhagen, 25–26 Nov 2010

Nicolas F (2009) Chinese direct investment in Europe: facts and fallacies, Chatham House, International economics briefing paper, 1 Jan 2009

Nicolas F (2010) Chinese direct investment in France: no French exception, no Chinese challenge, Chatham house, International economics programme paper, 2 Feb 2010

OECD (1994) OECD reviews of foreign direct investments – Italy. OECD, Paris

OECD (2008) China: encouraging responsible business conduct. OECD, Paris

Quer D, Claver E, Rienda L (2011) Political risk, cultural distance, and outward foreign direct investment: empirical evidence from large Chinese firms. Asia Pacific J Manage (Published online 13 Jan 2011), pp 1–16

Rui H, Yip GS (2008) Foreign acquisitions by Chinese firms: a strategic intent perspective. J World Bus 43(2):213–226

Schmitz H (1995) Collective efficiency: growth path for small-scale industry. J Dev Stud 31(4):529–566

Schüler-Zhou Y, Schüller M (2009) The decision-making autonomy in Chinese overseas subsidiaries: The German case. GIGA Institute of Asia Studies, Mimeo/Hamburg

Schüller M, Turner A (2005) Global ambitions: Chinese companies spread their wings. China aktuell J Curr Chinese Aff 4:3–14

Simmons MS (2008) Huawei technologies: the internationalization of a Chinese company. In: Alon I, McIntyre JR (eds) Globalization of Chinese enterprises. Palgrave Macmillan, Houndmill, pp 194–207

Wu F (2005) The globalization of corporate China. NBR Anal 16(3):1–29

Regional Development and Government Income Transfer Programs: Combining Input–Output Systems and Geoprocessing as Tools for Planning in São Paulo State, Brazil

Silvio Massaru Ichihara, Joaquim José Martins Guilhoto, and Denise Imori

Abstract To measure the regional impacts and possible development and planning policies of the "Bolsa Familia" Program, an income transfer program founded by the Brazilian Federal Government, this study innovates by combining geoprocessing with input–output theory. The analytical potential is showed through the estimation and analysis of the São Paulo State Municipalities supply and demand relations. These relations were estimated in the process of the construction of the state interregional input–output system. By having in view that the creation of accurate development strategies depends on regional peculiarities, the results show that this program must be understood not only as a form of income transference, but also as a catalytic agent for decreasing the regional inequality inside the state.

Keywords Input–output • Geoprocessing • Regional development • Brazil • São Paulo state • Income transfer programs

---

JEL codes: C67, R15, R58

S.M. Ichihara (✉)
DERSA, Secretary of Transport – State of São Paulo, São Paulo, Brazil
e-mail: ichihara@dersa.sp.gov.br

J.J.M. Guilhoto
Department of Economics, FEA, University of São Paulo, São Paulo, Brazil

Regional Economics Applications Laboratory (REAL), University of Illinois, Urbana-Champaign, IL, USA
e-mail: guilhoto@usp.br

D. Imori
Department of Economics, FEA, University of São Paulo, São Paulo, Brazil
e-mail: denise.imori@usp.br

R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*, Advances in Spatial Science, DOI 10.1007/978-3-642-33395-8_20,

1 Introduction

As seen throughout the Book, spatial factors – largely overlooked by traditional growth theories, as indicated by the *Introduction* – cannot be ignored when one analyses the differential in economic performances among regions. Therefore, while chapter "The Long Run Interplay Between Trade Policy and the Location of Economic Activity in Brazil Revisited" approaches the debate over regional inequality in Brazil, emphasizing its possible interplay with trade policy, the present chapter analyses a more specific region of this country and the trade flows originated from there, proposing a methodological approach that emphasizes the spatial aspects of economic relations. More specifically, in order to analyze the Federal income transfer program "Bolsa Família" in one of the poorest regions of the São Paulo State, the Ribeira Valley, and to formulate strategies of development for this region, this study proposes the combination of an interregional input–output system with geoprocessing.

The input–output systems describe the flows of goods and services among the many sectors of an economy, considering the characteristics of one or several regions. According to Leontief (1965), its dimension can be large enough to represent the world economy or adequately small so as to describe a metropolitan region. When assuming a greater degree of geographical focalization, these models are able to incorporate the regional peculiarities, enabling the elaboration of specific analysis and of urban planning strategies (Isard 1998).

However, one of the greatest problems faced by these studies is the difficulty in obtaining the appropriate data. Surveys with an adequate statistical relevance are expensive, demanding significant time and dedication to be executed. Moreover, there is the need for secrecy, aggravated when one analyses business data. In this way, the advantages of the regional analysis collide with the insufficiency of information, so that the researcher is compelled to formulate alternative methods for the estimation of nonexistent or unavailable data. In this framework, the present work intends to formalize a procedure sequence that enables the estimation of an interregional input–output system, considering a higher degree of regional disaggregation.

Employing such methodology, the results from the empirical analysis of this chapter emphasize the role of trade-related spatial externalities among regions. In order to conjecture its possible dynamics, one should examine chapter "Assessing Regional Economic Performance: Regional Competition in Spain Under a Spatial Vector Autoregressive Approach", whose authors measure these spillovers in the context of a spatio-temporal analysis of regional growth, for Spain.

Another important point to be considered in the present chapter is that, as indicated in the *Introduction* of the Book, transport costs indeed seem to modify the balance between the agglomeration and dispersion forces acting on regional economic performance. It would be of great interest to contemplate our results having in sight those of chapter "Input–Output Linkages, Proximity to Final Demand and the Location of Manufacturing Industries", which suggests the presence of strong nonlinearity in the underlying spatial process of the location of US manufacturing plants.

The following section presents the proposed method and, subsequently, the evaluation of the results is done by means of the geographical representation of the main elements of the input–output table. In order to do this, one utilizes geoprocessing techniques, in accordance with the initial ideas exposed in Guilhoto et al. (2003).

2 Methodology

Initially, this section presents the structure of the interregional input–output table and the notation used to represent it. In the subsequent subsection, one briefly describes the main methods utilized in the estimation of economic flows among regions, focusing on the gravitational input–output approach, which contains elements of the information theory and spatial variables. At last, the variables needed for the implementation of the chosen method are calculated and the proper adaptations and considerations on the modeling are characterized.

2.1 The Interregional Input–Output Model

The input–output model is composed by matrices and vectors based in the actual relations of the economy, considering the logical and quantitative linkages among the productive sectors. Its structure has to be in accordance with the international standards defined by the United Nations – *System of National Accounts* – SNA (United Nations 1993).

The denomination "input–output table", utilized in this study, corresponds to the group composed by the use matrix (***n*** sectors x ***n*** sectors) and the vectors of final demand (investments, exports, inventory changes, households and government's consumption), imports, taxes, value added and total production. Figure 1 presents this structure, considering **r** regions.

Where:

$\underset{(nxn)}{\mathbf{Z}}^{rr}$: matrix of intersectoral and interregional intermediate consumption,

composed by elements:

$\mathbf{z}_{ij}^{st}$ – sales of sector i of region s to the sector j of region t, or consumption of sector j of region t from the sector i of region s; $(s, t \in \mathbf{r})$, $(i, j \in \mathbf{n})$.

$\underset{(lxn)}{\mathbf{I}}^{r}$: vector of the volume of imports of the several sectors of region **r**,

composed by elements:

$\mathbf{i}_j^t$ – volume imported by sector j of region t.

$$\begin{array}{ccccccccccc}
\underset{(nxn)}{\mathbf{Z}^{11}} & \cdots & \underset{(nxn)}{\mathbf{Z}^{1t}} & \cdots & \underset{(nxn)}{\mathbf{Z}^{1r}} & \underset{(nx5)}{\mathbf{Y}^{11}} & \cdots & \underset{(nx5)}{\mathbf{Y}^{1t}} & \cdots & \underset{(nx5)}{\mathbf{Y}^{1r}} & \underset{(nx1)}{\mathbf{X}^{1}} \\
\vdots & \ddots & & & \vdots & \vdots & \ddots & & & \vdots & \vdots \\
\underset{(nxn)}{\mathbf{Z}^{s1}} & & \underset{(nxn)}{\mathbf{Z}^{st}} & & \underset{(nxn)}{\mathbf{Z}^{sr}} & \underset{(nx5)}{\mathbf{Y}^{s1}} & & \underset{(nx5)}{\mathbf{Y}^{st}} & & \underset{(nx5)}{\mathbf{Y}^{sr}} & \underset{(nx1)}{\mathbf{X}^{s}} \\
\vdots & & & \ddots & \vdots & \vdots & & & \ddots & \vdots & \\
\underset{(nxn)}{\mathbf{Z}^{r1}} & \cdots & \underset{(nxn)}{\mathbf{Z}^{rt}} & \cdots & \underset{(nxn)}{\mathbf{Z}^{rr}} & \underset{(nx5)}{\mathbf{Y}^{r1}} & \cdots & \underset{(nx5)}{\mathbf{Y}^{rt}} & \cdots & \underset{(nx5)}{\mathbf{Y}^{rr}} & \underset{(nx1)}{\mathbf{X}^{r}} \\
\underset{(1xn)}{\mathbf{I}^{1}} & \cdots & \underset{(1xn)}{\mathbf{I}^{t}} & \cdots & \underset{(1xn)}{\mathbf{I}^{r}} & & & & & & \\
\underset{(1xn)}{\mathbf{T}^{1}} & \cdots & \underset{(1xn)}{\mathbf{T}^{t}} & \cdots & \underset{(1xn)}{\mathbf{T}^{r}} & & & & & & \\
\underset{(1xn)}{\mathbf{V}^{1}} & \cdots & \underset{(1xn)}{\mathbf{V}^{t}} & \cdots & \underset{(1xn)}{\mathbf{V}^{r}} & & & & & & \\
\underset{(1xn)}{\mathbf{X}^{1}} & \cdots & \underset{(1xn)}{\mathbf{X}^{t}} & \cdots & \underset{(1xn)}{\mathbf{X}^{r}} & & & & & &
\end{array}$$

Fig. 1 Matrix notation of the interregional input–output table

$\underset{(lxn)}{\mathbf{T}}{}^{r}$: vector of net indirect taxes of the several sectors of region **r**, composed by elements:

t_j^t – net indirect taxes collected of sector j of region t.

$\underset{(lxn)}{\mathbf{V}}{}^{r}$: vector of value added of the several sectors of region **r**, composed by elements:

v_j^t – valued added of sector j of region t.

$\underset{(nxf)}{\mathbf{Y}}{}^{rr}$: matrix of final demand of region **r** for the production of s, composed by five columns (**f** = 5): (a) households' consumption; (b) government's consumption; (c) gross fixed capital formation; (d) exports; and (e) inventory changes.

y_i^{st} – consumption of setor i of region s by the vectors of final demand of region t.

$\underset{(lxn)}{\mathbf{X}}{}^{r}$: vector of total production of the several sectors of region **r**, composed by elements:

x_j^t – total production of sector j of region t (total sum of the columns).

$\underset{(nxl)}{\mathbf{X}}^{r}$: vector of total production of the several sectors of region **r**, composed by elements:

x_i^s – total production of sector i of region s (total sum of the rows).

2.2 Estimation of the Interregional Flows: Application of the Gravity Input–Output Model

The Brazilian national input–output table used in this study was estimated by means of the system of national accounts, using the methodology described by Guilhoto and Sesso Filho (2005b). The inexistence of publications similar to the system of national accounts in the level of regional administrations makes impossible to utilize the same method to construct municipal tables. Hence, algebraic techniques have been elaborated or adapted from other fields, in the attempt to approximate the estimated interregional flows to the real ones.

Presently, as emphasized throughout this Book, setting aside the spatial question is not justifiable. Together with the evolution of the geoprocessing framework, the lines of thought such as that Alfred Weber[1] achieved an even greater importance by means of responding to questions relative to the optimum localization of the productive activities, respectively.

Since it considers spatial elements, of information and of input–output theories, the present work has elected the gravity[2] input–output model, with restrictions of supply and demand, as the theoretical basis of the calculations meant for estimating the interregional flows for the São Paulo State.

2.2.1 Regional Supply and Demand Data

By means of the São Paulo Economic Activity Survey (PAEP), carried out by the State Data Analysis System Foundation (SEADE 2002), it is possible to analyze the main segments that compose the economy of São Paulo State. The survey can be indicated as a powerful tool, able to characterize the economic activity in regional level.

[1] The Central Place Theory of Weber evidences the hierarchical relation among cities and is based in scale economies and in the optimization of transportation costs (Fujita et al. 2002). In this way, the optimal localization of production in relation to demand is a factor that justifies even more the utilization of models related to the transportation network in the input–output analysis.

[2] The gravity approach utilized in many fields, also incorporated to the modeling of the distribution of demand for transportation (model of four stages), admits characteristics that are very close to those required in the process of estimating the interregional flows. It is inspired in the Newtonians observations about gravity, assuming that the movements of goods depend on the levels of demand in the destination region and of supply in the origin, but are inhibited by the attrition of distance.

The PAEP used in this study corresponds to the year 2001–02, being composed by information obtained in questionnaires applied to the several economic sectors. It covers the trade, the general industry (extractive and transformation industry), the construction industry, the financial institutions and the services.

In the notation of Fig. 1 of the input–output table, the *GPV* corresponds to the vector $\boldsymbol{X}$ of total production and can also be associated with the regional sectoral supply.

$$GPV_i^s = x_i^s = o_i^s \, GPV_i^s x_i^s o_i^s \tag{1}$$

Where:

GPV_i^s- Gross Production Value of sector i of region s;

x_i^s- total production of sector i of region s;

o_i^s- total supply of i of region s;
$s, t \in \mathbf{r}$ regions; $i, j \in \mathbf{n}$ sectors.

Concerning the *IC*, its elements do not represent the sectoral demand, because they are associated only to the consumption of the productive activities and not to the total consumption, which has to also consider the Final Demand ($\boldsymbol{Y}$). The *IC* represents the sum of the rows of the matrix $\boldsymbol{Z}$ (intersectoral and interregional flows) and of the vectors $\boldsymbol{I}$ (imports) and $\boldsymbol{T}$ (taxes), in accordance with Eq. 2.

$$IC_i^t = \sum_{i=1}^{n} \sum_{s=1}^{r} z_{ij}^{st} + i_j^t + t_j^t \tag{2}$$

Where:

IC_i^t- Intermediate Consumption of sector i of region t;

z_{ij}^{st}- transactions of sector i of region s with sector j of region t;

i_j^t- imports of j of region t;

t_j^t- taxes collected of j of region t;

The *IC* evaluates the quantity each sector in each region consumes of inputs (aggregation of commodities).

2.2.2 Calculation of the Demand, Based on the Leontief-Strout Model

The Leontief-Strout model (1963) defines that the entire production of sector *i* destined to region *t* is equal to the consumption of *i* for the sectoral production added to the final demand related with *i*, in region *t*, in accordance with Eq. 3. The basic idea is to form a pool, both of the demand of a region *t* for *i*, and the share of the supply of *i* from the entire system that is destined to region *t*. In this way, the system facilitates the estimation of the flows from one region to another, by way of reaching the equilibrium.

$$\sum_{s} f_i^{st} = \sum_{j} a_{ij}^t . x_j^t + y_i^t \tag{3}$$

Where:

f_i^{st}- trade flow of the good or service *i* from region *s* to *t* ($\forall s, t$); ($\forall i$);

y_i^t- aggregation of final demand for *i* in region *t*, either considering or not considering the exports;

x_j^t- total production of *j* ($\forall j$) in region *t*;

a_{ij}^t- technical coefficients of the input–output tables of each region *t*.

The technical coefficients of the input–output table (a_{ij}) are considered here in order to facilitate the estimation of the consumption of a given commodity that is necessary for the production of the activities in each region. The coefficient is obtained from the ratio between the intersectoral flow and the total production of *j* of region *t*, in other words, it is the proportion of *i* used in the production of one unit of *j* in *t*: $a_{ij}^t = \sum_s z_{ij}^{st} \Big/ x_j^t$.

In practice terms, using the Leontief-Strout model means to calculate how much of commodity *i* is demanded and how much is supplied of the same *i* in each region, from the elements of the input–output table (technical coefficients of the use matrix, final demand and total production of each region).

Utilizing the notation of the interregional input–output table (Fig. 1) and the formulation of the Leontief-Strout model, the expression that determines the vector of demand can be represented by:

$$d_i^t = \sum_{j=1}^{\mathbf{n}} \sum_{s=1}^{\mathbf{r}} z_{ij}^{st} + \sum_{s=1}^{\mathbf{r}} y_i^{st} \tag{4}$$

Where:

$i, j \in \mathbf{n}$ sectors;
$s, t \in \mathbf{r}$ regions;

d_i^t- total demand for i in region t;

y_i- corresponds to the sum of the five vectors of final demand for i;

st in z^{st} and y^{st}- represents the destination of the production of region s, to the intermediate consumption and the final demand of region t, respectively.

In Eq. 2, if one separates taxes and imports, the value added of the intersectoral and interregional transactions can be established by the IC, obtaining a value that, although algebraically similar, is quite different from the element of the sum of z in Eq. 4, since in one of the cases there is row (i) sum and in the other, column (j) sum.

Although IC is not exactly what one is looking for, it can be used to estimate the intersectoral matrices of aggregated intermediate consumption in each region ($\sum_s \mathbf{Z}^{st}$, with the notation of Fig. 1). By means of these matrices, one can obtain the column (j) sum of z (Eq. 4).

In this point, the procedure requires an input–output table representing the totality of the interregional system to be obtained ($\sum_t \sum_s \mathbf{Z}^{st}$, following the notation of Fig. 1). In other words, if the objective is estimating the relations among the municipalities, one needs to utilize a State table; if the objective is obtaining an interstate system, one needs to already have a national table, and so on. In the present work, in order to estimate a interregional system considering the municipalities of São Paulo, it was necessary to utilize the data of the input–output table of this State.[3]

In order to convert the state data of sectoral demand into municipal data, one makes use of, essentially, the participation in IC of each sector in each municipality in relation to the state IC of the corresponding sector, proportionally distributing the technical coefficients associated with the matrix of intersectorial consumption of São Paulo. The following development demonstrates the obtainment of the aggregated use matrix for each municipality:

1. Initially, one obtains the vector $\underset{(\mathrm{lxn})}{\mathrm{IC}}$, composed by the elements that correspond to the intermediate consumption of each sector j of the state. Its expression is given by the column sum of the intersectorial consumption of São Paulo (considering Eq. 9 for a single region);
2. The obtainment of the vector of participation of the IC of each municipality **r** in the state total, settled as $\underset{(\mathrm{lxn})}{\mathrm{pIC}^t}$, is given by Eq. 5:

[3] This table corresponds to a system of 42 productive sectors obtained for the São Paulo State, using the method presented by Guilhoto and Sesso Filho (2005a). The technique of estimation is based in the disaggregation of the state data from the national Brazilian table, estimated in accordance with Guilhoto and Sesso Filho (2005a), and in the use of data on the interstate flows presented in Vasconcelos (2006).

$$\underset{(\mathrm{lxn})}{\mathrm{pIC}^t} = \underset{(\mathrm{lxn})}{\mathrm{IC}^t} \cdot \underset{(\mathrm{nxn})}{\left(\hat{\mathrm{IC}}\right)^{-1}} \tag{5}$$

Where:
$t \in \mathbf{r}$ regions – São Paulo municipalities ($\mathbf{r} = 645$);

IC^t - vector corresponding to the values of intermediate consumption of each sector in each one of the municipalities of São Paulo State, obtained in the PAEP.

3. At last, the aggregated use matrix of each municipality is given by Eq. 6.

$$\underset{(\mathrm{nxn})}{\sum_{s=1}^{r} Z^{st}} = \underset{(\mathrm{nxn})}{\left(\sum_{s=1}^{r}\sum_{t=1}^{r} Z^{st}\right)} \cdot \underset{(\mathrm{nxn})}{\left(\mathrm{p}\hat{\mathrm{CI}}^t\right)^{-1}} \sum_{s=1}^{r} \tag{6}$$

This calculation sequence allows one to estimate the sectoral demand related with the intermediate consumption of each municipality, but under a strong hypothesis that the technical coefficients of production of each region are proportional to the state mean. [4]

Using the results of Eq. 6, the first element in the right side of Eq. 4 can be obtained and the second element, relative to the final demand, is constituted by means of specifying the aggregated final demand (**Y**) data in the municipal level.

In order to do so, one directly utilizes data from other statistical basis, besides the PAEP. The values of the five vectors that compose the state final demand were proportionally distributed to the municipalities, using variables that correspond exactly to the group of demand or to a *proxy*, in accordance with Box 1.

At last, the values of sectoral supply (o_i^s) and demand (d_i^t) for each municipality can be obtained. Briefly, the value of o_i^s follows from information about the GPV of each productive sector in each municipality, in accordance with Eq. 1. Obtaining d_i^t is a more complex process, which requires considerations of the Leontief-Strout model, data on the aggregated municipal *CI*, information about the municipal final demand and a non-regionalized input–output table considering all the regions to be considered.

[4] This hypothesis means that the production technology of all the municipalities is similar to the state mean. In other words, a company of sector A, in a municipality, consumes inputs from sectors B, C and D in the same proportion as the state mean, considering all the companies that compose sector A, in the State. For this reason, the results of the present model must not be applied in analysis intending to evaluate the technological differences of the sectoral production in the municipalities.

Box 1 Variables Used for Specifying the Final Demand of São Paulo at the Municipal Level

Group of final demand	Variables that are available for each sector and municipality	Observations about the use of municipal data for specifying state data
Exports	Exports by municipality (Brasil 2007a)	Variable of direct correspondence with the exports
Government consumption	Public expenditure (SEADE 2007)	Variable of direct correspondence with the government consumption
Households consumption	Estimated municipal population in 2002 (IBGE 2007), together with the sectoral households consumption structure of the input–output table	Variable utilized as *proxy* of the importance of households consumption in each municipality, together with data from the table
Investments	Gross Operating Surplus (GOS) deduced from the Value Added, present in PAEP – SEADE (2002)	The magnitude of the sectoral GOS was used as *proxy*, since the sectors with the largest GOS have the greatest investment capacity
Inventory changes	Gross Production Value (GPV), calculated in PAEP – SEADE (2002)	The GPV was used as *proxy* of the inventory changes, since it measures the total quantity of resources that are involved in each sector

2.2.3 Definition of the Variable Used as Impedance Factor in the Gravity Approach

By means of the geographic information systems it is possible to estimate transportation costs or other variables that can act as attrition factors. However, the calculation of the transportation costs, in order to evaluate the intersectoral relations among the regions, is not trivial, since two essential aspects need to be considered: the first is relative to the heterogeneity and to the modes of transportation question; the second concerns the functions and variables that should be used in the determination of the costs.

The problem of the heterogeneity in the activities of the productive sectors emerges if one or more productive activities utilize different modes of transportation and are grouped in the same sector of the input–output table. Concerning the function and its variables, Isard (1998) mentions that the cost can be considered as distance, time or monetary value. In the distribution models that use the gravity approach, one intends to determine a generalized cost function, in which linear functions are recommended and that may incorporate a variable denominated modal penalty. This parameter would be a form of considering all the variables that cannot be easily dimensioned, and are relative to the modes of transportation that describe them.

In the present work, one has opted for considering only one transportation modality for the majority of the productive sector: the road transportation. The sectoral productions of extraction and refinery of oil and gas, and the iron and steel industries largely utilize other modes besides the road transportation. However, these productions are concentrated in few localities in São Paulo State, so that the relations of these sectors can be specifically treated.

In relation to the cost function, its measure is expressed in time, considering the road distance, average velocity, kind of pavement and situation of the road stretch, obtained by means of the software *TransCad® – Caliper* (Geographic Information System applied to transportation). The georeferenced database, composed from the national road network, corresponds to the one of the same company.

The differences in altitude between origin and destination are also utilized. In order to calculate them, one uses interpolation techniques for treating the georeferenced altitude measures, available in the Integrated Cartographic Digital Base of Brazil to the Millionth Scale (IBGE 1997).

Using this group of information, the measures of the time consumed in covering the shortest route between a municipality (s) and the other (t) are obtained and attributed as value of the variable c_i^{st}.

2.3 Estimation of the Municipal Interregional and Intersectoral Input–Output Table

After obtaining o_i^s, d_i^t and c_i^{st}, it is then possible to estimate the trade flows among the regions (f_i^{st}) by means of Eq. 7.

$$f_i^{st} = a^s.b^t.o_i^s d_i^t.\exp\left(\gamma c_i^{st}\right)^{-1} \quad (7)$$

In this process, the balancing factors of the gravity model (a^s, b^t and γ) can be estimated through iteration and adjustment processes defined in Ortúzar (2004) and Nanne and Heydecker (1998). The procedures are based in iteration methods for searching the best values, and can be calibrated by auxiliary information about the distribution of the interregional flows. In this study, the algorithms were implemented and executed through the mathematical software *Matlab® – MathWorks*.

Having obtained the trade flows and the other data, it is possible to estimate the interregional input–output table, considering the attributions of a multi-regional model. The following development is based in the techniques presented in Miller and Blair (2009) for constructing the interregional table and the Leontief inverse matrix from the trade flows among the regions, the aggregated intermediate consumption and the aggregated final demand.

The solution of Eq. 7 produces a group of f_i^{st} values that can be represented by matrices of trade flows ($\underset{(\mathbf{rxr})}{\mathbf{F}_i}$):

$$\mathbf{F}_i = \begin{bmatrix} f_i^{11} & \cdots & f_i^{1t} \\ \vdots & \ddots & \vdots \\ f_i^{s1} & \cdots & f_i^{st} \end{bmatrix}$$

Where:

s, t $\in$ **r** regions (**r** = 645);
i $\in$ **n** sectors (**n** = 42)
f_i^{st}- trade flow of the commodity i from region s to t.

The Eq. 8 transforms the matrices $\mathbf{F}_i$ in matrices $\mathbf{P}_i$, of trade proportions, determining the participation of each commodity used in t proceeding from each region s, on the total utilized by t, including $t = s$.

$$\underset{(\mathbf{rxr})}{\mathbf{P}_i} = \underset{(\mathbf{rxr})}{\mathbf{F}_i} \cdot \underset{(\mathbf{rxr})}{\left(\hat{\mathbf{V} \cdot \mathbf{F}_i}\right)}^{-1} \tag{8}$$

Where:

s, t $\in$ **r** regions (**r** = 645);
i $\in$ **n** sectors (**n** = 42)

$$\underset{(\mathrm{lxr})}{\mathrm{V}} = \text{unity vector} = \underset{(\mathrm{lxr})}{[1 \ldots 1]}$$

Each $\boldsymbol{P}_i$ matrix is composed by elements p_i^{st}, that have to be organized in a single square trade matrix given by $\boldsymbol{G}$**,** of dimension $\boldsymbol{n.r}$.

$$\underset{(\mathbf{nr\ x\ nr})}{\mathbf{G}} = \begin{bmatrix} \mathbf{G}^{11} & \cdots & \mathbf{G}^{1t} & \cdots & \mathbf{G}^{1\mathbf{r}} \\ \vdots & \ddots & & & \vdots \\ \mathbf{G}^{s1} & & \mathbf{G}^{st} & & \mathbf{G}^{s\mathbf{r}} \\ \vdots & & & \ddots & \vdots \\ \mathbf{G}^{\mathbf{r}1} & \cdots & \mathbf{G}^{\mathbf{r}t} & \cdots & \mathbf{G}^{\mathbf{rr}} \end{bmatrix} \tag{9}$$

Where:

G^{st}- diagonal matrices composed by p_i^{st} (element of row s and column t of the $\mathbf{P}_i$ matrix).

$$\underset{(\mathbf{nxn})}{\mathbf{G}^{st}} = \begin{bmatrix} p_1^{st} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_n^{st} \end{bmatrix}$$

Another matrix (**SZ**) of same dimension has to be formed, using the matrices generated in Eq. 6 to compose the diagonal. Its other sub-matrices need to be null, as Eq. 10 illustrates.

$$\underset{(\mathbf{nr\ x\ nr})}{\mathbf{SZ}} = \begin{bmatrix} \sum_{s=1}^{\mathbf{r}} \mathbf{Z}^{s1} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & & & \vdots \\ \mathbf{0} & & \sum_{s=1}^{\mathbf{r}} \mathbf{Z}^{st} & & \mathbf{0} \\ \vdots & & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & \sum_{s=1}^{\mathbf{r}} \mathbf{Z}^{s\mathbf{r}} \end{bmatrix} \tag{10}$$

Where:

$\underset{(nxn)}{\sum_{s=1}^{r} Z^{st}}$ - matrix representation of the interregional relations, regionally aggregated, in the t regions.

In the same way, the matrix related with the final demand (**SY**) also has to be generated from the information of Box 1. In this case, the dimension of this matrix depends on the detail level required in the final demand; if the five vectors are separated, then the dimension of the matrix will be $\boldsymbol{n.r}\,x\,5.\boldsymbol{r}$, in accordance with Eq. 11.

$$\underset{(\mathbf{nr\ x\ 5r})}{\mathbf{SY}} = \begin{bmatrix} \sum_{s=1}^{\mathbf{r}} \mathbf{Y}^{s1} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & & & \vdots \\ \mathbf{0} & & \sum_{s=1}^{\mathbf{r}} \mathbf{Y}^{st} & & \mathbf{0} \\ \vdots & & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & \sum_{s=1}^{\mathbf{r}} \mathbf{Y}^{s\mathbf{r}} \end{bmatrix} \tag{11}$$

Where:

$\underset{(nx5)}{\sum_{s=1}^{r} \mathrm{Y}^{st}}$ - matrices of **n** rows and five columns, that represent how much the final demand of a region t consumes of the production of all other s regions, including $s = t$.

With the *G*, *SZ* and *SY* matrices, one can construct the interregional input–output table, basing it in the multiregional modeling, in accordance with Eqs. 12 and 13.

$$\underset{(\mathbf{nr\ x\ nr})}{\mathbf{ZR}} = \underset{(\mathbf{nr\ x\ nr})}{\mathbf{G}} \cdot \underset{(\mathbf{nr\ x\ nr})}{\mathbf{SZ}} \tag{12}$$

$$\underset{(\mathbf{nr\ x\ 5r})}{\mathbf{YR}} = \underset{(\mathbf{nr\ x\ nr})}{\mathbf{G}} \cdot \underset{(\mathbf{nr\ x\ 5r})}{\mathbf{SY}} \tag{13}$$

In these expressions, *ZR* is the Intermediate Consumption matrix, which describes both the interregional and the intersectoral relations, and *YR* is the Final Demand detailed for all the sectors in all regions.

$$\underset{(\mathbf{nr\ x\ nr})}{\mathbf{ZR}} = \begin{matrix} \underset{(\mathbf{nxn})}{\mathbf{Z}^{11}} & \cdots & \underset{(\mathbf{nxn})}{\mathbf{Z}^{1t}} & \cdots & \underset{(\mathbf{nxn})}{\mathbf{Z}^{1r}} \\ \vdots & \ddots & & & \vdots \\ \underset{(\mathbf{nxn})}{\mathbf{Z}^{s1}} & & \underset{(\mathbf{nxn})}{\mathbf{Z}^{st}} & & \underset{(\mathbf{nxn})}{\mathbf{Z}^{sr}} \\ \vdots & & & \ddots & \vdots \\ \underset{(\mathbf{nxn})}{\mathbf{Z}^{r1}} & \cdots & \underset{(\mathbf{nxn})}{\mathbf{Z}^{rt}} & \cdots & \underset{(\mathbf{nxn})}{\mathbf{Z}^{rr}} \end{matrix} \qquad \underset{(\mathbf{nr\ x\ 5r})}{\mathbf{YR}} = \begin{matrix} \underset{(\mathbf{nx5})}{\mathbf{Y}^{11}} & \cdots & \underset{(\mathbf{nx5})}{\mathbf{Y}^{1t}} & \cdots & \underset{(\mathbf{nx5})}{\mathbf{Y}^{1r}} \\ \vdots & \ddots & & & \vdots \\ \underset{(\mathbf{nx5})}{\mathbf{Y}^{s1}} & & \underset{(\mathbf{nx5})}{\mathbf{Y}^{st}} & & \underset{(\mathbf{nx5})}{\mathbf{Y}^{sr}} \\ \vdots & & & \ddots & \vdots \\ \underset{(\mathbf{nx5})}{\mathbf{Y}^{r1}} & \cdots & \underset{(\mathbf{nx5})}{\mathbf{Y}^{rt}} & \cdots & \underset{(\mathbf{nx5})}{\mathbf{Y}^{rr}} \end{matrix}$$

Following these stages, one then obtains the system initially presented, in Fig. 1. In the case of the intermunicipal table of São Paulo, the vectors *I* (imports), *T* (taxes), *V* (value added) and *X* were previously obtained by means of PAEP and other statistical basis. However, in other works the values of *T* could be associated with *Z*, or be obtained by the proportionality of the *IC*, while the vector *V* would be obtained by the difference of *IC* and the total production.

At last, the Leontief inverse matrix, which is basis of the impact analysis and other techniques of the input–output theories, can be obtained using the technical coefficients of the Intermediate Consumption matrix, as Eq. 14 indicates.

$$\underset{(\mathbf{nr\ x\ nr})}{\mathbf{IL}} = \left(\underset{(\mathbf{nr\ x\ nr})}{\mathbf{I}} - \underset{(\mathbf{nr\ x\ nr})}{\mathbf{AR}} \right)^{-1} \tag{14}$$

Where:

$$\underset{(\mathrm{nrxnr})}{\mathrm{AR}} = \underset{(\mathrm{nrxnr})}{\mathrm{ZR}} \cdot \left(\underset{(\mathrm{nrxnr})}{\hat{\mathrm{XR}}} \right)^{-1};$$

$\underset{(nrxnr)}{\mathrm{AR}}$ - matrix of interregional intersectoral technical coefficients;

$$\underset{(\mathrm{nrx1})}{\mathrm{XR}} = \left(\underset{(\mathrm{nrxnr})}{\mathrm{ZR}} \cdot \underset{(\mathrm{nrx1})}{\mathrm{V}'} \right) \cdot \left(\underset{(\mathrm{nrx5r})}{\mathrm{YR}} \cdot \underset{(\mathrm{nrx1})}{\mathrm{V}}' \right);$$

$\underset{(nrxnr)}{XR}$ - vector of total production – row sum of **ZR** and **YR**;

$\underset{(1xnr)}{V}$ - unity vector.

3 Regional Development and Income Transfer Programs

This work is based in the process of estimation of an inter-municipal input–output table, so that its results are characterized by the system formed by such information. However, the dimension of the obtained matrix is very large, being constituted by more than 43 million items,[5] a fact that raises difficulties to the visualization of the results by means of tables or graphics. However, while the vast number of regions makes it difficult to evaluate the results, it allows the utilization of tools that make use of the space, such as the geoprocessing techniques.

In the next subsection, the interregional relations are interpreted in relation to the households' consumption, enabling a hierarchical visualization of the municipalities and the identification of regional polarizing centres. Finally, the second subsection illustrates how these results can be used in the definition of strategic plans for reducing the regional inequality, having in mind the spatial distribution of the per capita amount of resources expended on the "Bolsa Família", a federal government transfer program, in the municipalities of São Paulo. The analysis highlights the role played by spatial spillover effects on economic activities.

Figures 2 and 3 show the location of the administrative regions (ARs) that aggregate the municipalities of São Paulo in 15 groups, also positioning the main access roads in the State. Considering the location of these roads is important in the present analysis since, as indicated previously, transport costs modify the balance between agglomeration and dispersion forces acting on regional economic performance and its dynamics.

3.1 Polarizing Centers for the Households Consumption

A possibility to understand the relationship among regions is to put the flows of goods and services such that only the most important ones, for each municipality, are represented. In this way, it is possible to represent at least one link for each city.

[5] The system was previously calculated considering the 42 productive sectors, but it was then aggregated into 8 groups, in order to facilitate the understanding of the data. Besides this, there are 5 groups of final demand for the 645 municipalities of São Paulo, determining a ***ZR*** square matrix with 5,160 rows and columns and a ***YR*** matrix with the same number of rows and 3,225 columns.

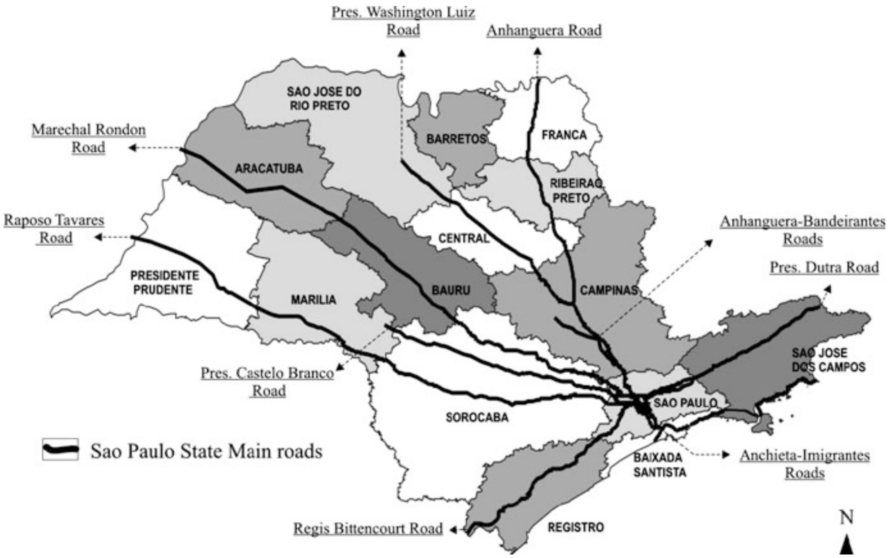


Fig. 2 Administrative Regions (ARs) and the main roads of São Paulo

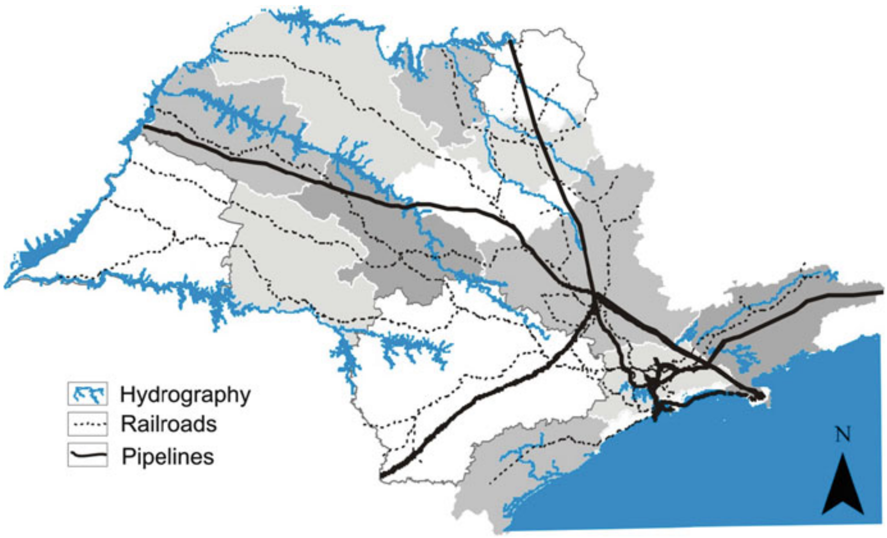


Fig. 3 Railroads, pipelines and hydrography of São Paulo

Figure 4 presents only this main link, considering the flows of the trade sector destined to the household consumption, among the municipalities with the largest supplies and all the others. This structure allows identifying the polarizing centers and its covered cities.

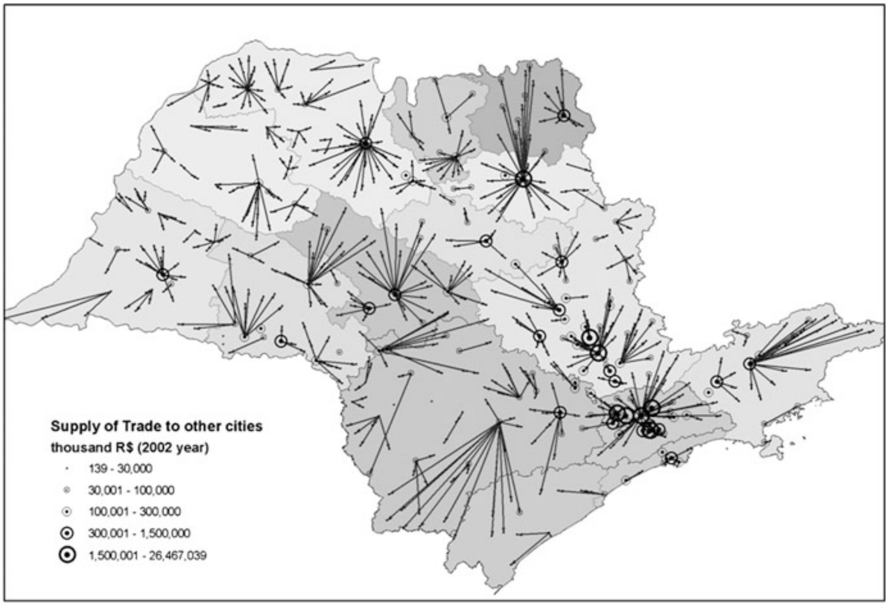


Fig. 4 Main trade supplying cities and covered municipalities

Works such as the Cities' Influence Region – REGIC (IPEA 2002) and the Proposal to a Regionalized Brazil (Brasil 2007c) have already determined hierarchies and poles for the whole Brazil, using very different methods that consider other variables beyond the economic data.

In this way, this study does not intend to compare the Fig. 4 with the results of the mentioned works, because this structure presents only an ordinal classification between the trade demand of households in the smaller cities and the production of the bigger cities.

However, is important to enhance that this type of information can be used in future studies that intend to include the economic intersectoral relations of the input–output table in the hierarchic definition among the cities, i.e., different presentations of polarizing cities can be defined considering each type of production and/or sectoral demand.

3.2 Considerations on the Reduction of the Regional Inequality

The objective of this section is to propose a localized action capable of reducing the regional inequality, exemplifying a possible use to an intermunicipal input–output table.

Using the data from the Ministry of Social Development (Brasil 2007b) about the municipal distribution of the amount spent by the "Bolsa Família" Program

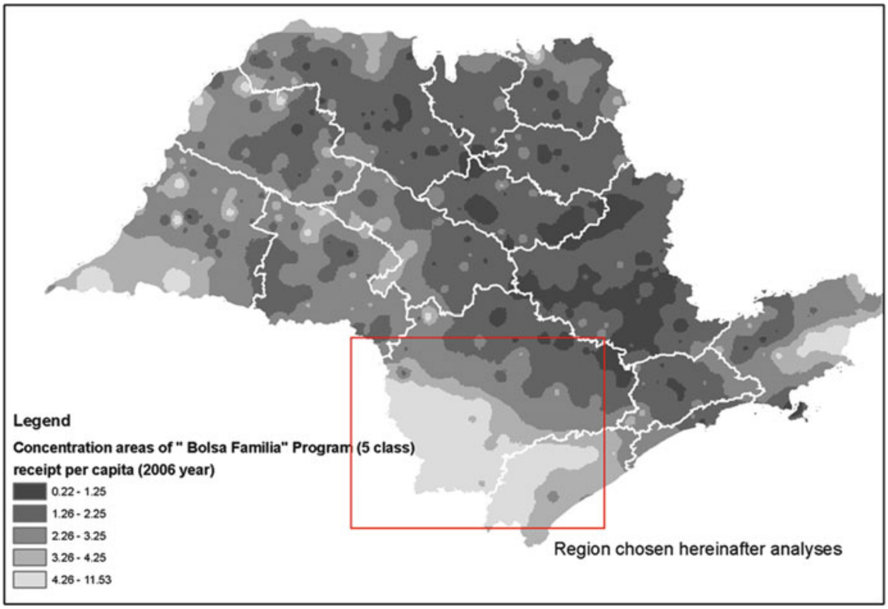


Fig. 5 Distribution of resources per inhabitant of the "Bolsa Família" Program in the São Paulo State

(BFP),[6] Fig. 5 illustrates the areas in relation to expenses per inhabitant. Once again, the areas that stand out are the cities of the Ribeira Valley and some cities of the Paraíba Valley.

Figure 5 does not intend to represent the areas where the transferences with the BFP are more expressive, but it does intend to represent the areas where the poor population is relatively higher. Considering only the three biggest cities of the State: São Paulo, Guarulhos and Campinas, the sum of transferences by the BFP corresponded to 25.7 % of the total destined to the State (about R$ 61.4 million in July of 2006). However, considering this volume divided by the total population, the per capita expense was about R$ 1.20 per person, according to the estimative of the population in 2006 (IBGE 2007). At the same time, the cities of the region of Ribeira Valley, previously pointed as the least developed in the State, have transfers rates 4–10 times superior per inhabitant, demonstrating that the population of this area is more dependent on the assistance program.

Among the goals of PBF, there is the promotion of an immediate relief for poverty (Brasil 2007b), what would characterize it as an emergency program, not a permanent one. However, if joint measures of social and economic development are

[6] The "Bolsa Família" Program (BFP) is a conditional program of direct income transfers of the Federal Government. The objective of this program is to guarantee that "poor" (monthly income per person from R$ 60 to R$120) and "extremely poor" (income per person lower than R$60) families receive a monthly benefit, according to Brasil (2007b).

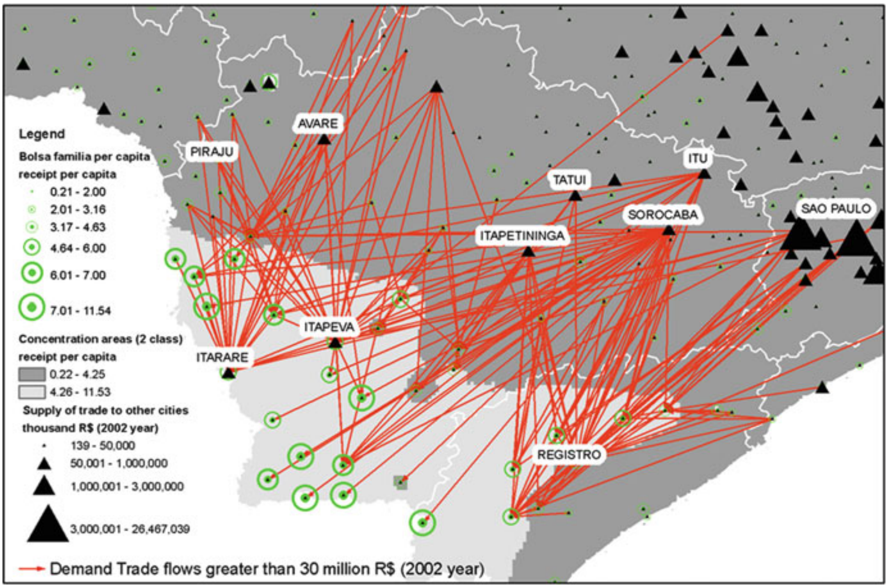


Fig. 6 Main demand flows of the municipalities with higher BFP per inhabitant rates

not taken, the PBF may not only become a permanent program but may also have to be expanded, determining a larger public expense and limiting the resources invested in infrastructure, what is essential for the economic growth.

It is expected that the solution is not to abruptly cut the funds for the program, but fortifying the development and improving the income distribution, so that the poor and extremely poor people no more receive the benefit because they no longer need it.

In this context, the concepts on the spillover of resources and the polarization centres can be used to search strategies destined to reduce regional inequality and to eventually fortify the cities of relatively more needy populations.

The Fig. 6 demonstrates only the main demand flows from the municipalities of the area that was highlighted in the Fig. 5, characterized by the biggest expense per inhabitant rates. One can notice that, basically, the municipalities of the highlighted area of the map present little interaction among them. Only Itapeva and Itararé are capable of supplying for the neighboring cities, all the other cities only demand from bigger cities as Itapetininga, Sorocaba, Tatuí, Registro and from the metropolitan region of São Paulo.

In this situation it is highly likely that the resources of the BFP destined to the families of the municipalities in the highlighted area of the map are generating income in the municipalities outside it.

The Federal Government destines R$ 2.4 million per month (value of July of 2006), or about 4 % of the state expense with the BFP, to approximately 1 % of the state population that inhabits the region of the Ribeira Valley. However,

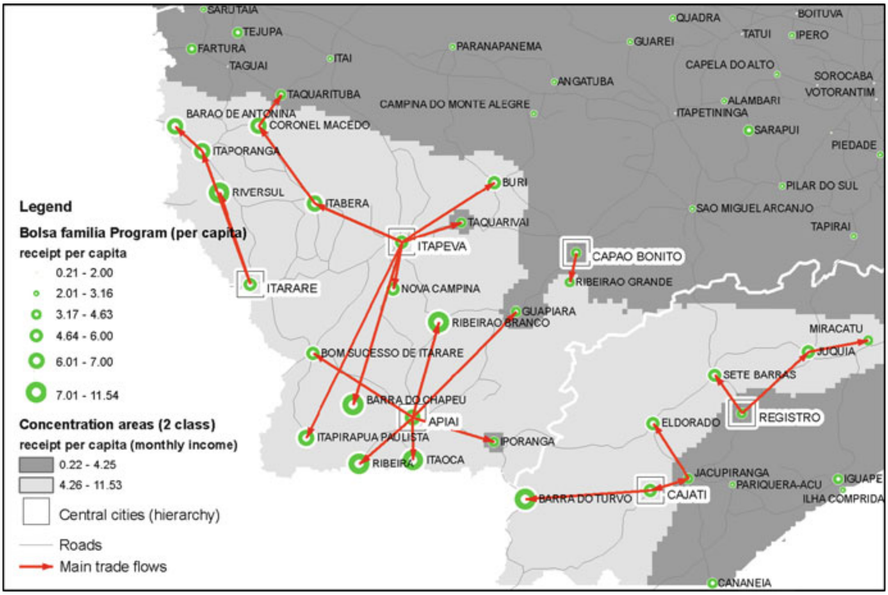


Fig. 7 Main demand flows of the municipalities with higher PBF per inhabitant rates

a great part of the multiplying effect generated by these resources is not kept in the region and these resources end up fortifying other more developed regions (north of the AR of Sorocada and AR of São Paulo) and even other States, such as Paraná.

Thus, the proposed question is not primordially related with the BFP, but with the possibility that the local income generation is being lost because of the spillover of these resources to other regions.

For this reason, it is possible to focus only in the relationship between the cities of the highlighted region and the bordering cities. The Fig. 7 represents the polarizing centres, considering the main flow of each city, as was done in Fig. 4.

Considering Fig. 7, six cities have been distinguished for offering its production to the other ones. Itapeva, Itararé and Registro are bigger cities, with more than 50,000 inhabitants, being more developed and capable of polarizing the regional trade. However, the city of Apiaí, located in the center of the highlighted region, has a population similar to many others (between 20,000 and 30,000 inhabitants) and appears to be relatively more important in the local trade.

Its position is interesting, because it is located in the intersection of the roads that link other cities that present a high PBF per inhabitant rate. According to this diagnosis, although superficial, the city of Apiaí could be elected as a strategic point for the local development.

Apiaí and the cities that integrate its range area (Ribeirão Branco, Bom Sucesso do Itararé, Ribeira, Itaoca, Iporanga and Guapiara) receive more than half million reais every month by means of the BFP, but they are not capable of making good use of the multiplying effects from these resources in their own economies.

In case that nothing is done, they can continue depending of the BFP and other forms of conditional direct income transfers for a long period of time. But if private and governmental incentives are destined to this region, maybe these cities will be able to centralize part of these resources, avoiding the spillover of production to places outside the region, in a way to vitalize and improve the economic relationship between the municipalities.

Since the industrial potential of the region is weak and, at the same time, the agricultural sector cannot expressively develop (this is an irregular terrain region, composed by areas of environmental protection), it is necessary to investigate other activities. The activities of the services sector can pointed for having, on average, a better capacity of fixing the multiplier effect.

The region can improve its economy through the exploration of the adventure tourism, as it has already been done in areas next to Apiaí, Iporanga and Eldorado, using the geological formation conserved by the Alto do Ribeira Tourist Park (PETAR). The government can, for example: introduce superior courses in tourism, geology and/or biology in the region and, at the same time, give incentives to the private initiative to extend the tourist infrastructure.

Measures as these, unfortunately, do not mean reducing the social inequalities, but they can reduce the regional inequalities.

This text does not have the intention of deeply exploring or offering a development plan for the region, but with this considerations one intends to exemplify how the data of a regionalized input–output table at the municipal level can be useful for understanding the economic relationships among the cities and how to use them to diagnose, direct or even recommend regional strategies with politics that already exists, as the BFP.

4 Conclusions

The estimation of input–output tables with greater regional specification can proportionate many benefits to regional studies. In this way, the present work proposes and evaluates a method in order to accomplish this task. Since the existing and available databases do not allow the construction of an intermunicipal input–output matrix based on census data, this work tries to approach the reality by using a large number of information, considering the production, the intermediate consumption and the final demand of the main activity groups in each municipality.

Accordingly with a point stressed by the present Book, that space and, more precisely, trade patterns deriving from its influence are key to economic performance and dynamics, this work presented an example of application of the input–output methodology, evaluating the spillover effects of resources from less developed areas that are included in the "Bolsa Família" Program, also demonstrating the possibility of identifying regional poles, to which one can design actions with the intention of vitalizing the local economies in a more precise way.

Anticipatively, it should be stressed that a similar study can be made for all of the Brazilian municipalities, what would make more sense, since the conditional direct income transfers is much more important in many other regions of the country, where the underdevelopment and the spillover of resources are even higher.

Finally, many other applications can use the data of an interregional input–output table, as well as this data can be linked to other methods, using spatial variables as the linking elements with other disciplines that use geographic information systems, like flow analysis, nets, geostatistics, geomarketing, among others, allowing new techniques to be created or adapted.

References

Brasil (2007a). Ministério do Desenvolvimento, Indústria e Comércio Exterior. Secretaria do Comércio Exterior. Disponível em: http://aliceweb.desenvolvimento.gov.br. Acesso em: 15 maio 2007

Brasil (2007b) Ministério do Desenvolvimento Social. Programa bolsa família. Disponível em: http://www.mds.gov.br/programas/transferencia-de-renda/programa-bolsa-familia. Acesso em: 27 ago. 2007

Brasil (2007c). Ministério Planejamento Orçamento e Gestão. Proposta de regionalização do Brasil - Ministério do Planejamento/Centro de Estudos de Gestão Estratégica - CCGE. Disponível em: http://www.mp.gov.br. Acesso em: 27 out. 2007

Fujita M, Krugman P, Venables AJ (2002) Economia espacial. Futura, São Paulo, p 391

Fundação Sistema Estadual De Análise De Dados – SEADE (2002) Pesquisa da Atividade Econômica Paulista – PAEP/Sala de acesso aos microdados

Fundação Sistema Estadual De Análise De Dados – SEADE (2007) Informações sobre os municípios paulistas. Disponível em: http://www.seade.gov.br. Acesso em: 12 Jan 2007

Guilhoto JJM, Sesso Filho UA (2005a) Estrutura produtiva da Amazônia: uma análise de insumo-produto. Belém: Banco da Amazônia S.A. (BASA), p 320

Guilhoto JJM, Sesso Filho UA (2005b) Estimação da Matriz Insumo-Produto a partir de Dados Preliminares das Contas Nacionais. Economia Aplicada**,** São Paulo 9(2):277–299, abr

Guilhoto JJM, Hewings GJD, Takashiba EH, Silva LMS (2003) Combining geographical information systems and input–output models: concept and initial ideas. In: Proceedings of the North American meeting of the regional science association international, 50, 2003, Philadelphia. Papers... : REAL, 2003. 1 CD-ROM, IL

Instituto Brasileiro De Geografia E Estatística – IBGE (1997) Base Cartográfica Integrada do Brasil ao Milionésimo Digital – 1997 – IBGE/DGC/CCAR. 1 CD-ROM

Instituto Brasileiro De Geografia E Estatística – IBGE (2007) Sistema IBGE de recuperação automática – SIDRA. Disponível em: http://www.sidra.ibge.gov.br. Acesso em: 25 maio 2007

Instituto De Pesquisas Econômicas – IPEA (2002) Caracterização e Tendências da Rede Urbana do Brasil, vol 2. IPEA, Brasília, p 390

Isard W et al (1998) Methods of inter-regional and regional analysis. Ashgate, Aldershot, p 490

Leontief W (1965) Input–output economics, 1st edn. Oxford University Press, New York, p 257

Leontief W, Strout A (1963) Multiregional input–output analysis. In: Barna T (ed) Structural interdependence and economic development. St. Martin's Press, New York, pp 119–150

Miller RE, Blair PD (2009) Input–output analysis: foundations and extensions, 2nd edn. Cambridge University Press, New York, p 750

Nanne JZ, Heydecker BG (1998) Improved algorithms for calibrating gravit models. In: Bell MGH (ed) Transportation networks: recent methodological advances. Pergamon, Amsterdam, p 319

Ortúzar JD, Willumsen LG (2004) Modelling transport, 3rd edn. Wiley, New York, p 499

Vasconcelos JR, Oliveira MA (2006) Análise da matriz por atividade econômica do comércio interestadual no Brasil – 1999. Texto para discussão no 1159. Instituto de Pesquisa Econômica Aplicada – IPEA, 2006. Disponível em: http://www.ipea.gov.br/pub

United Nations. System of National Accounts – SNA (1993) Disponível em: http://unstats.un.org/unsd/sna1993/toctop.asp