Lecture Notes in Economics and Mathematical Systems 680

Anna-Lena Sachs

Retail Analytics

Integrated Forecasting and Inventory Management for Perishable Products in Retailing

Springer

Lecture Notes in Economics and Mathematical Systems 680

More information about this series at
http://www.springer.com/series/300

Anna-Lena Sachs

Retail Analytics

Integrated Forecasting and Inventory Management for Perishable Products in Retailing

Springer

Anna-Lena Sachs
Department of Supply Chain Management & Management Science
University of Cologne
Cologne
Germany

Dissertation at Technische Universität München, TUM School of Management, submitted on February, 5th 2014 and accepted on April 15th, 2014

ISSN 0075-8442 ISSN 2196-9957 (electronic)
ISBN 978-3-319-13304-1 ISBN 978-3-319-13305-8 (eBook)
DOI 10.1007/978-3-319-13305-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014958011

Printed on acid-free paper

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

Abstract

Based on real data from a large European retail chain, we analyze newsvendor decisions for perishable products. We suggest a data-driven approach that integrates forecasting and inventory optimization. The approach is distribution-free and takes external factors such as price and weather into account. Using Linear Programming, the model fits a linear inventory function to historical demand observations. It accounts for unobservable lost sales and stockout-based substitution by considering the timing of sales occurrences. We show that the model captures real-world retail characteristics well, in particular, if the assumptions of existing standard models are violated or if demand depends on external factors and is highly censored. Additionally, we determine the optimal policy for an aggregated service level target in a multi-product context. We investigate whether order quantities determined by a real decision maker consider the elements of the optimal policy. We analyze behavioral causes to explain suboptimal decision-making and find that behavioral biases observed in laboratory experiments are also present in the real world.

Acknowledgements

First and foremost, I am very grateful to my supervisor Prof. Dr. Stefan Minner for his invaluable support, advice, and encouragement throughout the course of my Ph.D. studies—both in Vienna and in Munich. Thanks to him, I had the chance to pursue research projects in different fields and I enjoyed working on these topics very much. I am also very thankful to Prof. Dr. Martin Grunow for giving me useful feedback and many insightful comments on my research as well as for being part of the examination committee. I would like to express my gratitude to Prof. Dr. Rainer Kolisch who took over the position as chairman of the examination committee.

I want to thank all current and former members of the chair of Logistics and Supply Chain Management at TU München and University of Vienna: Christian Bohner, Maximilian Budde, Alexandra Ederer, Pirmin Fontaine, Evelyn Gemkow, Miray Közen, Dr. Nils Löhndorf, Katharina Mariel, Dr. Arkadi Seidscher, Dr. Lena Silbermayr, Dr. Martin Stößlein, Dariush Tavaghof Gigloo, and Michael Weingärtner. Thank you all for the many fruitful discussions.

I would also like to thank Prof. Dr. Mirko Kremer, Dr. Christian Lang, and Prof. Dr. Sandra Transchel for their guidance throughout the course of my Ph.D. studies.

Furthermore, I am very grateful for the fruitful cooperations during the past years. I would like to thank Prof. Ulrich Thonemann, Ph.D., and Dr. Michael Becker-Peth for our cooperation in several interesting projects in the field of behavioral operations management. I would like to express my deepest gratitude to Prof. Philip Kaminsky, Ph.D., for inviting me to work on our joint project on data-driven inventory management at UC Berkeley. I thank Dr. Birgit Löhndorf and Prof. Dr. Rudolf Vetschera for our joint work on probability effects in utility elicitation. I gratefully acknowledge the cooperation and most valuable input from Bernd Binder, Jens Daniel, and Madeleine Hirsch. Furthermore, I appreciate the great support by Achim Frauenstein and Leschek Swiniarski without whom the data collection would not have been possible. I also thank Walter Gossmann from the Bakery Guild for the insights I gained on bakery business in Germany. Thanks to Dr. Stefan Appel for providing technical support on the number crunching.

Finally, I would like to thank my parents for their critical comments and proofreading, my brother Johannes for sharing his expertise in Econometrics, my husband Kai for his patience and encouragement, as well as my daughter Elisa for bringing so much joy to my life.

Contents

List of Tables

List of Figures

Acronyms

AIC	Akaike information criterion
ARIMA	Autoregressive integrated moving average
BIC	Schwarz's Bayesian information criterion
BLUE	Best linear unbiased estimator
CR	Critical ratio
cv	Coefficient of variation
EM	Expectation-maximization
FIFO	First in first out
Known	Known parameters
LIFO	Last in first out
LP	Linear Program
LR	Likelihood ratio
MCS	Model confidence set
MM	Method of moments
MNL	Multinomial logit
MSE	Mean squared error
Non	Non-parametric
OLS	Ordinary least squares
OOS	Out-of-stock
Par	Parametric
POS	Point-of-sale
RMSE	Root mean square error
SI	Safety inventory
SKU	Stock-keeping unit
SL	Service level

Chapter 1
Introduction

1.1 Motivation

Recent studies on food losses have caught attention to the large amounts of food discarded in industrialized countries due to consumer behavior and lack of supply chain coordination which leads to expired excess inventories (Gustavsson et al. 2011). A study by the United States Department of Agriculture found that 11.4 % fresh fruit, 9.7 % fresh vegetables, and 4.5 % fresh meat and seafood are wasted annually (Buzby et al. 2009). In Germany, 11 million tons of food are wasted per year (Kranert et al. 2012). The large amount of discarded food not only represents a waste of natural resources but causes also a monetary loss to consumers and retailers.

Despite the large amounts of excess inventories, retailers experience frequent out-of-stock (*OOS*) situations. According to Corsten and Gruen (2003) *OOS* rates amount to 8.3 % worldwide considering non-perishable product categories. *OOS* situations are estimated to occur even more frequently for perishable products due to their short shelf-lives that make full product availability less desirable (ECR 2003).

Food waste and *OOS* situations are caused by mismatches between supply and demand. Supply chain management is especially challenging for food products due to demand uncertainty, short shelf lives, highly demanding customers, and low profit margins (Akkerman et al. 2010; Hübner et al. 2013). In out-of-stock situations, Corsten and Gruen (2003) estimate that about 50 % of the customers leave the store without buying the product, while the other half opts for substitution. Substitution leads to an increase in sales of other products. Empirical studies of customer reactions to stockouts indicate that customers finding poor availability in a store on a regular basis will not only incur short-term lost sales but will also decide not to return to this store in the long-run (Anderson et al. 2006). If stockout situations are ignored and ordering decisions are only made based on past sales data, a store will most likely also be understocked in the future if there was poor availability in

A.-L. Sachs, *Retail Analytics*, Lecture Notes in Economics and Mathematical Systems 680, DOI 10.1007/978-3-319-13305-8_1

the past. Furthermore, sales of the substitute products will be overestimated thus incorporating demand switching behavior even though it may be more beneficial to stock more of the first choice product.

Another reason for suboptimal decision making eventually causing waste and stockouts are the decision makers themselves. Human decision makers are subject to behavioral biases. By anchoring on reference points such as mean demand or chasing demand, the decision makers systematically deviate from the optimal order quantities (Schweitzer and Cachon 2000; Bendoly 2006; Bolton et al. 2012).

Retail companies have access to large amounts of data collected by point-of-sale (POS) scanner systems, but lack the tools to process the data and improve their current practices (Fisher and Raman 2010). Retail analytics aim to achieve a better alignment of supply and demand by leveraging this data.

This work is motivated by challenges faced when solving an inventory optimization problem for perishable products at a large European retail chain. We have collected POS data for fruit, vegetables, and bakery products of more than 60 stores at the retail chain starting 2008. Based on the real data, we analyze behavioral causes of suboptimal decision-making in practice and develop novel models to account for the prevalent retail characteristics.

1.2 Problem Statement

While there is a large body of literature on the inventory optimization and assortment planning problem, many theories lack applicability to real-life situations in retailing (Fisher 2009). Other than the classical text book approach with given demand distributions, the practical application poses several further challenges:

- Demand is non-stationary and depends on external factors such as price and weather.
- The parameters and the type of the demand distribution are unknown.
- Unsatisfied demand may result in lost sales that are unobservable.
- Substitution may take place if a product is out-of-stock.
- Decision makers do not always choose profit-maximizing order quantities and are prone to behavioral biases.

Existing approaches often ignore causal relationships between external factors and demand that could be leveraged for improved forecasting and rather rely on simplifying distributional assumptions that may not fit real data. Some of them rely on extensive data requirements that are not available to store managers, but disregard important information in the data contained such as time of the last sale that can be used to obtain estimates for unobservable lost sales or demand substitution. Stockout situations are often ignored and decisions are only made based on past sales data. Consequently, future order decisions may be biased due to unobservable lost sales and substitution behavior.

Furthermore, decision makers do not always choose the profit-maximizing solution eventually causing more waste or stockouts. Human decision makers are subject to behavioral biases (e.g., Schweitzer and Cachon 2000; Bolton and Katok 2008; Becker-Peth et al. 2013; and many more). Several laboratory experiments have shown that decision makers systematically deviate from the optimal order quantities, but the question arises whether these biases can also be observed in real-world data.

Ignoring the requirements posed by real-world environments either prohibits their practical application or results in inaccurate inventory control and consequently, loss of profit. Consequently, the following research questions will be addressed in this work:

1. How can external factors that influence demand be incorporated in inventory optimization to better align supply and demand?
2. Can forecasting and inventory optimization be integrated instead of performing two sequential tasks?
3. How can the time of the last sale be used to estimate unobservable lost sales and substitution behavior?
4. What are the optimal order quantities in a multi-product setting with an aggregated service level target and does the empirical decision maker behave profit-maximizing?
5. Are decision makers in real-world environments subject to the same behavioral biases as in laboratory experiments?

1.3 Outline

This research aims to use the data available to retailers in order to better balance excess inventory and out-of-stock situations for perishable products. Using real data containing daily and hourly sales from 2008 on, the inventory replenishment decisions for fruit, vegetables, and bakery products of more than 60 stores were analyzed.

Its contribution is to overcome limitations of existing research such as assumptions on the theoretical demand distribution not fitting the demand observations, to improve order decisions by taking additional information into account and to analyze the ordering behavior of a real decision maker from a behavioral perspective.

We first review related work on unobservable lost sales estimation, assortment planning and behavioral operations management in Chap. 2.

We then develop a data-driven model for single-period problems that integrates demand forecasting and inventory optimization in Chap. 3. The model is distribution-free and takes external variables such as price and weather into account that influence demand. A linear inventory function of the external variables is fitted to historical demand observations to determine its coefficients using Linear

Programming (*LP*). We compare the model to a time-series forecast and regression analysis in a numerical study and analyze the results based on real data. This chapter is based on Beutel and Minner (2012).

In Chap. 4 we extend the model from Chap. 3 by assuming that the retail manager has no full demand information and only observes sales. If a stockout occurs, the lost sales have to be estimated. Otherwise, the retail manager would underestimate true demand in future periods and order too few units. We establish sales patterns from days with full availability based on Lau and Lau (1996) to estimate the demand and compare the data-driven model to other parametric and non-parametric approaches in a numerical study. This chapter is based on Sachs and Minner (2014).

In Chap. 5 we consider a two-product model where customers either choose not to purchase a product if it is out-of-stock or purchase a substitute if their first-choice is sold out. We determine the additional demand due to substitution from the sales patterns that we used in Chap. 4 for the unobservable lost sales estimates. We compare the performance of the model to a parametric approach based on Poisson-distributed and real data.

We analyze empirical decisions in Chap. 6. We first develop a normative model to determine optimal order quantities for an aggregated service level target over several products and identify elements of the optimal policy. We then compare the performance of the model to the decisions of a single manufacturer who supplies retail stores with bakery products. Finally, we investigate whether the decision maker is subject to behavioral biases that were found in laboratory experiments. This is joint work with Michael Becker-Peth (University of Cologne), Stefan Minner (Technische Universität München) and Ulrich W. Thonemann (University of Cologne).

Chapter 7 concludes this thesis by summarizing the main findings. We herein state the limitations and areas for future research.

Chapter 2
Literature Review

In the following we will review literature on unobservable lost sales estimation, assortment planning and behavioral operations management.

With our focus on retail, we exclude literature on assortment planning in production from our analysis. In contrast to retail where substitution is consumer-driven, demand substitution in production systems is usually controlled by the supplier side. If demand cannot be filled, downward substitution takes place and excess demand is satisfied from a superior product (see for example Hsu and Bassok 1999). In the assortment planning literature, three main types of substitution can be distinguished based on their causes. If substitution occurs as a response to a temporary out-of-stock situation, it is called stockout-based substitution. If customers substitute one product for another to realize savings from price differences, it is referred to as price-based substitution. Assortment-based substitution takes place when a product is not carried by the store at all and thus, a customer chooses another variant from the available assortment set instead of her favorite product. For a comprehensive review on the assortment planning problem and related aspects, the interested reader is referred to Kök et al. (2008) and Pentico (2008).

2.1 Unobservable Lost Sales

The existing literature distinguishes between parametric and non-parametric approaches. The former is based on theoretical demand distributions such as the normal in Nahmias (1994), who suggests to approximate mean and standard deviation from sales data given an order-up-to level inventory policy. The censored part of the right tail of the normal distribution is calculated by taking into account that the distribution function is symmetric about its mean. A major advantage of the normal distribution function is its wide applicability to large datasets due to the

A.-L. Sachs, *Retail Analytics*, Lecture Notes in Economics and Mathematical Systems 680, DOI 10.1007/978-3-319-13305-8_2

Law of Large Numbers. In contrast, a Poisson process is generally more suitable for discrete and small datasets as in Conrad (1976) or the compound Poisson in Springael and van Nieuwenhuyse (2005).

But, as Agrawal and Smith (1996) emphasize, another important requirement is that the distribution chosen is also capable of capturing the effects of demand variation as present in retailing. They derive the parameters of the negative binomial distribution by matching the sample mean and the frequency of observing zero demand to the observed frequency of demand.

Lau and Lau (1996) propose a nonparametric model that does not require any prior distributional assumptions based on the product limit method (Kaplan and Meier 1958) and daily sales patterns obtained from previous observations.

Berk et al. (2007) use Bayesian updates for obtaining the parameter values of the negative binomial, Gamma, Poisson and normal distribution for the censored newsvendor problem. They rely on an approximation of the posterior distribution by matching the first two moments given that one parameter is known (e.g. mean or variance). Lu et al. (2006) consider Bayesian updates in the context of durable goods for a general distribution function and apply their findings to the normal distribution with known variance. Lu et al. (2008) analytically investigate the benefits from overstocking to learn about the true demand in a Bayesian setting. Tan and Karabati (2004) suggest an updating mechanism to achieve a desired service level by iteratively adjusting the inventory level. Jain et al. (2013) also use Bayesian updates, but additionally take the timing of sales transactions before a stockout occurs into account.

2.2 Assortment Planning

Pentico (1974) is among the first to address the assortment planning problem with stochastic demand. A single-period newsvendor solution is obtained under the assumption that customer arrivals occur before any demand is filled. A "no-crossover" assumption prohibits stockout-based substitution.

van Ryzin and Mahajan (1999) gain theoretical insights on the trade-off between inventory costs and product variety benefits. There is only assortment-based but no stockout-based substitution. Their analysis is restricted to cases where all variants offered have the same retail price-cost ratio. By adding variety to an assortment, total demand increases but comes at the expense of potential cannibalization effects if demand of the other items in the assortment decreases. Fast and costly replenishment strategies are desirable for fashion items for which demand is characterized by purchase behavior that depends on previous demand and therefore allows forecasting future sales once the season has started. In contrast, demand is assumed to be independent for casual items and scale economies should be leveraged by using large-scale store formats.

Another extension of van Ryzin and Mahajan (1999) is proposed by Maddah and Bish (2007) who additionally take pricing into account in a newsvendor setting with assortment-based substitution. Demand is represented by a multinomial logit (*MNL*) model with a mixed multiplicative and additive form. For a high customer arrival rate or large mean demand, the optimal prices in an assortment have equal profit margins.

Topaloglu (2013) also builds on the model of van Ryzin and Mahajan (1999). He extends the model by varying the assortments offered over a selling period. Topaloglu (2013) sets up a nonlinear model to determine which products to offer in an assortment and for how long each product should be offered. He uses the *MNL* model in order to solve the nonlinear model.

Miller et al. (2010) address the assortment planning problem for infrequently purchased goods. They assess the robustness of the optimal assortment for changing customer preferences. Therefore, they develop a *MNL* model with heterogeneous utilities as well as two other choice models. In addition to the retailer's objective function with choice probabilities, they establish upper and lower bounds on the expected profit by assuming that the customer purchases the most respectively least profitable product. They apply adaptive conjoint analysis to online purchase data in order to obtain individual product utilities and form consideration sets for each customer. By comparing results on customer preferences to a retailer's market share, they show that their estimates are reasonable. They find that increased heterogeneity leads to higher profit uncertainty but at the same time allows the shift of customers to more profitable products, thus resulting in higher expected profits.

Sauré and Zeevi (2013) study a dynamic assortment planning problem where the retailer learns about customer preferences by varying the set of products offered. There is only limited shelf space and the retailer must select which products to offer. They address the trade-off between learning out about customer preferences versus offering the best set of products (when no learning takes place anymore). They suggest policies to quickly find the best set of products and how to identify products as suboptimal that should not be carried.

2.3 Assortment Planning with Stockout-Based Substitution

Smith and Agrawal (2000) develop a base-stock inventory model with stockout-based substitution that determines the optimal assortment to be carried as well as inventory levels subject to a service level constraint. They show how further constraints such as shelf space can be incorporated into their approach. A logit choice model is used to determine substitution probabilities. Demand is also dependent on the inventory policy. In a comparison of single and multiple substitution attempts, they find that the more items are stocked the smaller is the effect of allowing for more substitution attempts since the probability of finding a suitable item approaches one.

Mahajan and van Ryzin (2001b) extend the model of Smith and Agrawal (2000) by introducing dynamic consumer substitution where the number of substitution attempts is not restricted and substitution rates depend on the availability of substitutes in a given assortment. They model demand and substitution as a general choice process. The profit-maximization problem is solved with a stochastic sample path gradient algorithm which is compared to heuristic policies. The setting with dynamic substitution is compared to static substitution where demand is independent of the current on-hand inventory levels. An important finding of their analysis is that the profit function is not quasi-concave in inventory levels. Furthermore, larger amounts of popular items and fewer amounts of unpopular items should be stocked in an inventory system with substitution compared to a traditional newsvendor.

Kök and Fisher (2007) develop a practice-motivated approach to determine the optimal assortment from sales data. Given their focus on products with long shelf life and high service level, the demand function is obtained from log-linear regression, ignoring unobservable lost sales. Parameters for assortment-based substitution are estimated from stores with varying assortments calibrated on full-assortment stores. The approach is then extended to possible out-of-stock situations. Stockout-based substitution rates are derived from individual store sales data using the expectation-maximization (*EM*) algorithm. Input data required include time of purchase, customer arrivals at different levels and number of product units sold. Other factors influencing purchase behavior such as price, weather and promotional activities are also incorporated. Finally, an iterative heuristic combined with a local search algorithm is applied to solve the assortment optimization problem. They find that stores should aim at higher inventory levels of goods with high demand variance thus hedging against potential lost sales. The amount of inventory to be carried of products with large case sizes depends on the available shelf space.

Hopp and Xu (2008) formulate an attraction model with a factor for each product that depends on quality and price. Multiple substitution attempts are modeled by a static approximation as a simplification of the dynamic substitution approach of Mahajan and van Ryzin (2001b). Different settings of price, service and assortment competition are studied. In a duopoly with price, service and assortment competition, product variety diminishes compared to a monopoly in order to avoid price competition whereas the total number of products and thus inventory level increases.

Honhon et al. (2010) consider the assortment planning problem with stockout-based substitution. Demand is classified into different customer types whereas each type has a certain ranking of purchase preferences. Prices remain fixed in this model. The optimal assortment is determined for a fixed proportion of each customer type and a heuristic is provided for the more general case with random proportions. They find that the optimal set of assortment possesses a certain structure in terms of newsvendor fractiles and underage cost.

Yücel et al. (2009) combine assortment planning with the supplier selection problem in the presence of quality issues and dynamic substitution behavior. For each of these aspects, a cost function is included in the overall objective function that is furthermore subject to shelf space constraints and constraints on the quantity of each product that can be supplied. In their analysis, they show that ignoring one of the factors substitution, supplier selection or shelf space limitations results in a significant loss of profit.

2.4 Stockout-Based Substitution in a Fixed Assortment

Parlar and Goyal (1984) address the newsvendor problem with two products and stockout-based substitution with exogenous rates. The assortment is assumed to be fixed and demand is represented by some type of density function. They show that the expected profit as a function of order quantities of the two products is strictly concave. The concavity can only be shown if the retail price of one product lies within a certain range of the substitute's price. Upper and lower bounds of this range can be calculated from the substitution probabilities of the two products.

Mahajan and van Ryzin (2001a) consider inventory competition in an oligopoly. Each firm offers one product and customer demand depends on the availability of the substitute product, i.e. stockout-based substitution takes place. Demand is modeled as a utility-maximizing choice process from which a *MNL* model is chosen for the numerical example. The authors prove the existence and uniqueness of a Nash equilibrium. They provide support for a competitive overstocking effect: In a symmetric setting where all firms face the same costs, prices and choice probability, the firms would be better off if they all decreased their inventory levels in the Nash equilibrium. However, none of the firms would decide to do so because each of them would have an incentive to choose a higher inventory level than its competitors (prisoners' dilemma). Consequently, they all end up with higher inventory levels and lower profits. The profits of each firm decrease with an increasing number of firms in the market.

Netessine and Rudi (2003) compare two settings where stockout-based substitution takes place. In the centralized setting, one decision maker determines the order quantities of all products that may be mutual substitutes. In the decentralized setting, a decision maker only determines the order quantities of one product. Order quantities of potential substitutes are determined by other decision makers. Given deterministic substitution rates, excess demand is reallocated to another product in a given assortment. In the decentralized setting, expected profits are optimized for each product separately. Compared to the standard newsvendor solution, the optimal order quantity is increased by the substituted demand from other products. Therefore, the decentral decision makers have to take the anticipated inventory levels of the other products into account which results in a Nash equilibrium. In the centralized setting, inventory levels are chosen to maximize the expected profit for all products. Demand switching into both directions—to and from a certain product—results in adjusting the newsvendor inventory up- and downward. While inventory levels in the decentralized setting are usually higher than the centralized solution, the authors show also a counterexample where decentralized inventory levels may also be lower for some products.

2.5 Joint Pricing and Inventory Planning with Substitution

Aydin and Porteus (2008) consider the joint inventory and pricing problem for multiple products based on Petruzzi and Dada (1999). Investigating the effect of prices in a given assortment, they allow for price-based substitution. Given several assumptions on the relationship between product price and attractiveness, they establish different demand models, e.g., comprising of the logit and multiplicative competitive interaction functions. The profit function is not jointly quasi-concave in prices and inventory levels, but based on the first-order conditions, they find a price-inventory level vector that maximizes their profit function. They find that a unit cost increase of one product results in lower optimal prices for the other products. Additionally, if a product's attractiveness is quality-dependent, higher quality of one product leads to higher optimal prices for all products.

The problem considered by Zhao and Atkins (2008) is an extension of Petruzzi and Dada (1999) and closely linked to Hopp and Xu (2008) but evaluating simultaneous price and inventory competition. Products are sold by competing vendors, and customers are allowed one substitution attempt to the product offered by another seller in case that their favorite variant is out-of-stock. This type of substitution behavior is represented by price-independent spill rates. Demand is a general function of price competition with a stochastic component. They demonstrate that the profit function is jointly quasiconcave in inventory and price and establish a Nash equilibrium under certain assumptions. For a linear demand function, they show that increasing spill rate raises prices, inventory and safety stock and results in a positive effect on profits. This positive effect can be enhanced by price competition if the spill rate is high. The direction of the effect of increased price competition on prices and safety stocks depends on the level of the spill rate.

Karakul and Chan (2008) address the joint procurement and pricing problem with substitution. Given the fixed price of an existing product, the (higher) price for a newly-introduced product is determined taking into consideration potential cannibalization effects from the existing product. They show that considering substitutability results in higher prices and safety stock of the new product but lower safety stock of the existing one and overall increased profitability.

2.6 Behavioral Operations Management

Human decision makers often deviate from the normative solution. In the past decade, several studies analyzed this behavior and potential causes in the newsvendor setting. Starting with Schweitzer and Cachon (2000), newsvendor decisions were analyzed in laboratory experiments. Schweitzer and Cachon (2000) provide evidence that human decision makers consistently order too few units in high profit settings and too many in low profit settings. This bias towards mean demand was named the pull-to-center bias. Bolton and Katok (2008) show that the pull-to-center bias also persists after many rounds and (almost) no learning takes place (Bolton and Katok 2008; Ho et al. 2010). The pull-to-center bias is no matter of experience since it is not only observed in experiments with students but also with managers who already have experience with similar ordering tasks (Bolton et al. 2012).

Several potential causes for this pull-to-center bias have been identified: One potential explanation by Schweitzer and Cachon (2000) is demand chasing where the decision maker anchors on his order quantity and adjusts towards observed demand.

A second potential explanation is ex-post inventory error minimization which describes the minimization of anticipated regret from not matching demand (Schweitzer and Cachon 2000). Kremer et al. (2014) find additional evidence for decision maker's preference to avoid ex-post inventory errors. If participants in an experiment are offered additional information on demand which would help to reduce the mismatch between demand and supply, i.e. the ex-post inventory error, they are willing to pay a price that exceeds the benefit from eliminating risk based on several risk utility functions.

A third potential explanation is anchoring on mean demand and insufficient adjustment (Schweitzer and Cachon 2000; Schiffels et al. 2014). It describes the tendency of human decision makers to facilitate the decision making process by choosing mean demand as reference point that is then adjusted.

Another potential explanation was developed by Su (2008). Su (2008) assumes that the newsvendor decision maker chooses the order quantity from a set of alternative solutions where more attractive alternatives (i.e. order quantities yielding higher profits) are chosen with higher probability. This decision process is modeled as logit choice model where the decision makers make random errors.

Kremer et al. (2010) show that ex-post inventory minimization, anchoring and adjustment, and demand chasing are valid in the presence of framing. They present a newsvendor situation in two different ways: one group of participants is aware of facing a standard newsvendor problem whereas the other group plays a lottery experiment. The first group shows a significantly stronger tendency towards mean anchoring and demand chasing than the second group. This result contrasts Su (2008)'s findings since the explanation of random errors should hold for both cases and not depend on the frame.

Building on the ex-post inventory error minimization framework and prospect theory (Kahneman and Tversky 1979), Ho et al. (2010) include reference-dependent preferences in a multilocation newsvendor model. They model the reference-dependent preferences as psychological costs which the decision makers associate with leftovers and stockouts. Decision makers associate higher psychological costs with leftover inventory than with stockouts.

Ren and Croson (2013) identify overconfidence as an explanation for the pull-to-center bias. If the newsvendor is too confident in his estimates, he underestimates the variance which results in an order quantity closer to mean demand than optimal. The authors suggest corrective measures based on coordinating contracts.

Becker-Peth et al. (2013) show that knowledge about behavioral biases helps to improve contract design. They find that order quantities under a buyback contract do not always lie between mean demand and the optimal order quantity. Further, they study how decisions under a buyback contract are affected by the contract parameters finding that the chosen order quantities not only depend on the critical ratio. The decision makers value revenues from sales differently than from returning items to the supplier. Becker-Peth et al. (2013) use these findings and the data collected in their experiments to fit parameters of adjusted buyback contracts. They test these adjusted buyback contracts in another set of experiments and show that they capture the decision making process better than without behavioral parameters.

Other studies investigate, for example, the effect of censoring (Feiler et al. 2013; Rudi and Drake 2014) or forecasting (Kremer et al. 2011). For more general reviews on behavioral operations not only in the newsvendor context refer to Katok (2011) and Bendoly et al. (2010).

Chapter 3
Safety Stock Planning Under Causal Demand Forecasting

Mainstream inventory management approaches typically assume a given theoretical demand distribution and estimate the required parameters from historical data. A time series based framework uses a forecast (and a measure of forecast error) to parameterize the demand model. However, demand might depend on many other factors rather than just time and demand history. Inspired by a retail inventory management application where customer demand, among other factors, highly depends on sales prices, price changes, weather conditions, this chapter presents two data-driven frameworks to set safety stock levels when demand depends on several exogenous variables. The first approach uses regression models to forecast demand and illustrates how estimation errors in this framework can be utilized to set required safety stocks. The second approach uses (Mixed-Integer) Linear Programming under different objectives and service level constraints to optimize a (linear) target inventory function of the exogenous variables. We illustrate the approaches using a case example and compare the two methods with respect to their ability to achieve target service levels and the impact on inventory levels in a numerical study. We show that considerable improvements of the overly simplifying method of moments are possible and that the ordinary least squares approach yields better performance than the LP-method, especially when the data sample for estimation is small and the objective is to satisfy a non-stockout probability constraint. However, if some of the standard assumptions of ordinary least squares regression are violated, the LP approach provides more robust inventory levels.

3.1 Introduction

Forecasting demand is undoubtedly one of the main challenges in supply chain management. Inaccuracy of forecasts leads to overstocks and respective markdowns or shortages and unsatisfied customers. To secure supply chain performance against

A.-L. Sachs, *Retail Analytics*, Lecture Notes in Economics and Mathematical Systems 680, DOI 10.1007/978-3-319-13305-8_3

forecast inaccuracy, an important countermeasure is safety stocks. The size of safety stocks required to obtain a certain customer service level depends on the degree of demand uncertainty and the corresponding forecast errors. Demand misspecification significantly affects the whole supply chain but its costs can be reduced by minimising forecast errors (see Hosoda and Disney 2009). An improvement of forecasts therefore directly results in inventory savings and service level improvements.

Mainstream inventory management models require specification of a demand distribution (e.g., normal, gamma; for an overview, see Silver et al. 1998) and are solved using stochastic calculus to derive required (safety) inventory levels. Alternatively, time series forecasting techniques are used and different measures for the resulting forecast error are then utilized to set safety stocks (see, e.g., Zinn and Marmorstein 1990; Krupp 1997). In practice, both the type of demand distribution and its parameters are unknown and need to be estimated. As a result, theory and practice often diverge since theory does not sufficiently address the needs of practice and practice uses overly simplistic approaches to overcome the problem (Lee and Billington 1992; Wagner 2002; Tiwari and Gavirneni 2007).

Accounting for estimation errors from having only a limited sample of historical demand observations is addressed by correcting the required estimation of the true demand standard deviation in Ritchken and Sankar (1984). The idea of estimating inventory control parameters treating historical demand as deterministic values and then optimizing performance in hindsight was proposed for the safety stock planning problem by Spicher (1975), further evaluated in Kässmann et al. (1986) and developed for an (s, S) inventory system in Iyer and Schrage (1992). Even though most forecasting and inventory models assume that demand is sampled from one distribution, it is more realistic that the kind of distribution may vary (Scarf 1958). In order to circumvent the requirement to assume (and test) a theoretical demand distribution, a different approach is followed by robust, distribution free inventory models where the inventory is set to maximize worst case performance (for a review see Gallego and Moon 1993). By giving less weight to outlier observations (e.g., using trimming), higher profits are sacrificed for obtaining a solution with lower variability. Lower variability thus reduces the risk associated with an inventory decision. More recently, Bertsimas and Thiele (2005, 2006) introduced a robust, data-driven approach that uses Linear Programming to obtain the required inventory level decisions based on historical data without making distributional assumptions. However, their approach works with a sample of observed demands only and does not take other factors such as price into account that might explain demand variations (Fildes et al. 2008).

Our research is motivated by an inventory planning problem in the retail sector where demand forecasting just based on historical demand did not provide satisfactory results due to strong demand dependency on other factors like prices, price changes, weather and others. For example, apart from price effects, retailers observe a strong increase in demand for lettuce on warm summer days in the barbecue season. Since time series models are unable to capture these effects, other forecasting methods were required. Econometrics provides a huge toolbox for estimation and statistical analysis, especially with respect to forecast errors. These methods allow decreasing safety stock by being able to explain a larger portion of the demand variability.

The contribution of this work is to discuss and promote integrated causal demand forecasting and inventory management in addition to the mainstream time series based approach. Our inventory planning approach accounts for the causal relationship between demand and external factors in order to explain larger portions of demand variability. We extend the data-driven approach to causal forecasting and compare it to existing methods such as regression analysis and the method of moments (*MM*). The data-driven approach directly estimates optimal inventory and safety stocks from historical demands and external factors whereas regression analysis and the method of moments is divided into demand prediction and inventory level determination. Demand variation is thereby separated into explained variation (with no need for protection by safety stock) and remaining (unexplained) variation which requires safety inventory.

In the following, Sect. 3.2 reviews the required basics from safety stock planning and the ordinary least squares framework and Sect. 3.3 presents the data-driven Linear Programming approach for integrated demand estimation and safety inventory planning. An illustrative example and a numerical comparison study including real data show the advantages and disadvantages of the application of the two proposed approaches in Sect. 3.4.

3.2 Safety Stock Basics and Least Squares Estimation

We consider a perishable item, newsvendor model with zero lead time and periodic review. For demand estimation, $i = 1, 2, \ldots, n$ historical demand observations D_i are available. We assume that demand is fully observable. Each demand observation i can be partially explained by a set of m explanatory variables X_{ji}, $j = 1, \ldots, m$.

In a newsvendor context under profit maximization, cost minimization, or a service level constraint, the target inventory level (single period order quantity) B and the implied safety stock level SI are set as follows. Assume that period demand follows a (continuous) theoretical demand distribution with probability density function f and cumulative distribution function F. Let μ and σ denote the respective mean and standard deviation. In a cost-based framework, let v denote the unit penalty cost for not satisfying a demand (underage cost) and h be the unit inventory holding cost (overage cost). In a service level (*SL*) framework, let P_1 denote a required non-stockout probability (α-service level) and P_2 denote a required fill-rate (β-service level).

$$B = \mu + SI \tag{3.1}$$

$$F(B) = P_1 = \frac{v}{(v+h)} \tag{3.2}$$

$$P_2 = 1 - \frac{1}{\mu} \int_B^{\infty} (x - B) f(x) dx \tag{3.3}$$

The target inventory level B thus consists of two components: mean demand and safety inventory SI (3.1) and depends on the target P_1 or P_2 service level or, equivalently, on penalty and holding costs. Equations (3.2) and (3.3) represent the standard formulas to set target inventory levels under service level constraints (see, e.g., Silver et al. 1998). For the widely used case of normally distributed demands (forecast errors), letting $f_{0,1}$ and $F_{0,1}$ denote the respective standardized normal density and distribution, the required safety factors under specified service level constraints become

$$B = \mu + k\sigma \tag{3.4a}$$

$$k_1 = F_{0,1}^{-1}(P_1) = F_{0,1}^{-1}\left(\frac{v}{v+h}\right) \tag{3.4b}$$

$$k_2 = G^{-1}\left(\frac{(1-P_2)\mu}{\sigma}\right) \tag{3.4c}$$

$$G(k_2) = f_{0,1}(k_2) - k_2(1 - F_{0,1}(k_2)) \tag{3.4d}$$

If only a sample of historical demand observations is available, besides making distributional assumptions about f, we need estimators for the parameters μ and σ in order to evaluate (3.1) and (3.2). These estimators can either directly be obtained from calculating mean and standard deviation of the sample as in the method of moments or as a function of other variables.

The classic approach in Econometrics to estimate the functional relationship between demand and any hypothesized factors is ordinary least squares (*OLS*) regression. Demand D_i is expressed as a (linear) function of explaining variables X_{ji} with the objective to minimize the sum of the squared errors. In the following the main results required for the purpose of safety stock setting are illustrated for a single explanatory variable first and then summarized for the multi-variable case using matrix notation.

3.2.1 The Single-Variable Case

With $m = 1$, demand can be expressed as the sum of an absolute term, a dependent term, and an error term u_i.

$$D_i = \beta_0 + \beta_1 X_i + u_i \tag{3.5}$$

Given the observations (X_i, D_i), the *OLS* estimators for β_0 and β_1 are (see, e.g., Gujarati and Porter 2009):

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(D_i - \overline{D})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \tag{3.6}$$

$$\hat{\beta}_0 = \overline{D} - \hat{\beta}_1 \overline{X} \tag{3.7}$$

where $\overline{D}$ denotes the average observed demand and $\overline{X}$ the average value observed for the explanatory variable.

For the purpose of inventory management and safety stock planning, the inventory for a given situation with known value of the explanatory variable X_0 (e.g., the next day's sales price) is required. For this purpose, the estimated standard deviation of $D(X_0)$ is of most interest.

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2}{n-2}} \tag{3.8}$$

$$\widehat{Var}(D(X_0)) = \hat{\sigma}^2 \left(1 + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2}\right) \tag{3.9}$$

Expression (3.9) consists of two parts, an estimator for the standard deviation of the error term u, as commonly used in inventory models with theoretical (normal) demand distribution, and the standard deviation of the estimator which accounts for the additional risk of the sample. Obviously, for $n \to \infty$ the latter term converges to zero. Compared to using the method of moments for estimating the standard deviation, the percentage reduction in safety stock by using regression can be stated by R^2 as the fraction of explained variation of total variation in observations. In an idealized setting where the regression function perfectly explains demand fluctuations, R^2 would be equal to 1 and no safety stock would be required. Any demand variation could be explained by changes in the exogenous variables that are known in advance. In the opposite case where the portion of explained variation approaches 0, results obtained by regression analysis are identical to those generated by the method of moments.

According to the classical *OLS* assumptions, we assume that the error term u_i is normally distributed with mean zero and constant variance σ^2 (homoscedasticity assumption). Furthermore, error terms of different observations are independent. Given that these assumptions hold, the least squares estimator $\hat{\beta}$ is the best linear unbiased estimator (*BLUE*) and linear in D_i. Its expected value equals the true value of β, i.e. it is unbiased. Furthermore, the estimator for the coefficient β has the smallest variance compared to other linear unbiased estimators, i.e. it is efficient or "best". For a detailed overview on assumptions and methods, see Gujarati and Porter (2009).

3.2.2 The Multi-Variable Case

In the general (linear) case with multiple regression variables, we present the results illustrated in the previous section in matrix notation. The demand model becomes

$$D_i = \beta_0 + \sum_{j=1}^{m} \beta_j X_{ji} + u_i \tag{3.10}$$

$$D = X\beta + u \tag{3.11}$$

where D is the vector of demands, β is the vector of m coefficients, X is the $m \times n$ matrix of observations of explanatory variables and u is the vector of errors. The estimator for the coefficients β_j is

$$\hat{\beta} = (X'X)^{-1}X'D \tag{3.12}$$

with X' denoting the transpose of X.

The required variance estimators become (see, e.g., Gujarati and Porter 2009).

$$E(u'u) = \sigma^2(X'X)^{-1} \tag{3.13}$$

$$\hat{\sigma}^2 = \frac{E(u'u)}{n-m} \tag{3.14}$$

$$\widehat{Var}(D(X_0)) = \hat{\sigma}^2(1 + X_0'(X'X)^{-1}X_0) \tag{3.15}$$

3.2.3 Violations of Ordinary Least Squares Assumptions

In real-world applications, several of the above stated *OLS* assumptions might be violated. In the following, we discuss heteroscedasticity, non-normal residuals, and errors with non-zero mean. Violations can be corrected either by adjusting the estimation procedure itself or by altering inventory planning.

If the zero mean assumption for the error terms does not hold, there are two possibilities. Either all error terms have a common mean other than zero. This simply adds to the constant, resulting in a biased estimate for the intercept $\hat{\beta}_0$ but with no effect on the estimation procedure. On the other hand, if the mean varies with each observation, there are more parameters than observations and the regression coefficients are biased which results in holding higher safety stocks to account for the larger portion of unexplained variability.

It is a common phenomenon in retail that one observes larger variations in demand when prices are especially low under promotions or extreme weather conditions or vice versa. As a consequence, the error term is heteroscedastic and the least squares estimator is no longer *BLUE*. It is still unbiased and linear, but not efficient. This means that standard *OLS* does not yield the minimum variance estimator (Gujarati and Porter 2009). A large variance then results in building up additional safety stock. To compensate this effect, we can adjust the estimation procedure according to Park (1966). This method assumes the following functional form for the error term:

$$\sigma_i^2 = \sigma^2 X_i^{\Theta} e^{r_i} \tag{3.16}$$

with r_i as the error in estimating the variance of the original error term u. Then, using $\hat{u}_i^2$ as an approximation for the unknown σ_i^2,

$$\ln \hat{u}_i^2 = \ln \sigma^2 + \Theta \ln X_i + r_i \tag{3.17}$$

serves to estimate parameter Θ with *OLS*-regression. The demand function is subsequently divided by $X_i^{\Theta/2}$ to establish homoscedasticity and regression analysis is performed on the transformed equation.

Finally, we assess the impact of non-normal demand on our model estimation approaches. Apart from normal demand, another widely used distribution in inventory theory is the gamma distribution in which case the *OLS* estimators are still *BLUE*. We can thus adjust the inventory planning approach for non-normal residuals in order to better achieve the required service level.

3.3 Data-Driven Linear Programming

Rather than following the indirect (parametric) approach to first estimate a demand model and then determine the inventory level B, we propose an integrated approach and assume that the required inventory level B is a linear function of the explanatory variables $X_0 = (X_{j0})_{j=1,\ldots,m}$ given by

$$B = \beta_0 + \sum_{j=1}^{m} \beta_j X_{j0}. \tag{3.18}$$

The decision variables are the parameters β_j and indirectly the resulting inventory levels y_i and satisfied demands s_i for each demand observation i. The problem can be formulated as a newsvendor model with a cost minimization objective or with a service level constraint.

3.3.1 The Cost Model

Assume a single period newsvendor framework with unit holding cost h and unit penalty cost v. The goal is to find the target inventory function parameters such that total costs comprising of inventory holding cost and shortage penalties are minimized. Other objectives like profit maximization can be incorporated in a straightforward manner. Further, logical constraints relate inventories, satisfied demands, and demand observations.

$$\min C = \sum_{i=1}^{n} (hy_i + v(D_i - s_i)) \tag{3.19}$$

s.t.

$$y_i \geq \sum_{j=0}^{m} \beta_j X_{ij} - D_i \qquad i = 1, \ldots, n \tag{3.20}$$

$$s_i \leq D_i \qquad i = 1, \ldots, n \tag{3.21}$$

$$s_i \leq \sum_{j=0}^{m} \beta_j X_{ij} \qquad i = 1, \ldots, n \tag{3.22}$$

$$s_i, y_i \geq 0 \qquad i = 1, \ldots, n \tag{3.23}$$

$$\beta_j \in \Re \qquad j = 1, \ldots, m \tag{3.24}$$

The objective function (3.19) is the sum of holding costs and penalty costs over all individual demand observation and their positioning above (inventory y_i) or below (shortage $D_i - s_i$) the linear target inventory level function. Equation (3.20) together with the minimization objective determines excess inventory y_i corresponding to each demand D_i. Equations (3.21) and (3.22) enforce that the sales quantity s_i under a given demand is equal to the minimum of demand D_i and supply.

3.3.2 The Service Level Model

In addition to the variables defined in Sect. 3.3.1, let γ_i define a binary demand satisfaction indicator for observation i. In the following we assume either a required non-stockout probability P_1, or a fill rate P_2. The goal is to find the target inventory function parameters such that inventory holding costs are minimized subject to a service level constraint.

$$\min C = \sum_{i=1}^{n} hy_i \tag{3.25}$$

s.t.

$$y_i \geq \sum_{j=0}^{m} \beta_j X_{ij} - D_i \qquad i = 1, \ldots, n \tag{3.26}$$

$$s_i \leq D_i \qquad i = 1, \ldots, n \tag{3.27}$$

$$s_i \leq \sum_{j=0}^{m} \beta_j X_{ij} \qquad i = 1, \ldots, n \tag{3.28}$$

$$D_i - \gamma_i M \leq \sum_{j=0}^{m} \beta_j X_{ji} \qquad i = 1, \ldots, n \tag{3.29}$$

$$\sum_{i=1}^{n} \gamma_i \leq n(1 - P_1) \tag{3.30}$$

$$\sum_{i=1}^{n} s_i \geq P_2 \sum_{i=1}^{n} D_i \tag{3.31}$$

$$s_i, y_i \geq 0 \qquad i = 1, \ldots, n \tag{3.32}$$

$$\gamma_i \in \{0, 1\} \qquad i = 1, \ldots, n \tag{3.33}$$

$$\beta_j \in \Re \qquad j = 1, \ldots, m \tag{3.34}$$

In (3.29) M is a large number (e.g., the maximum demand observation). If a non-stockout probability is required, then (3.29) ensures that the satisfaction indicator γ_i for demand D_i becomes one if demand exceeds supply. Summing over all observations n, (3.30) states that a maximum of $(1 - P_1)n$ observations are allowed to result in a stockout. Note that because of integer n, certain values of the required service level P_1 will result in an overachievement of service. Under a fill-rate service constraint expressed in (3.31), total sales have to exceed the fraction P_2 of total demand.

3.4 Numerical Examples

In a controlled simulation experiment, we compare the proposed methods. First, we generate data where all *OLS* assumptions are met, vary the sample size and then subsequently relax the *OLS* assumptions. Finally, we apply the proposed methods to real data.

In the following, we assume that true demand is a (linear) function of the sales price p plus some normally distributed error term u with zero mean and variance σ^2, $D = a - bp + u$. For a general review of price-dependent demand, see Petruzzi and Dada (1999). In practice, the parameters a, b, and σ^2 are unknown and need to be estimated. For each randomly generated problem instance, we normalize the price range to $p \in [0, 1]$ and draw the known parameters a, b and σ as follows:

- market size $a \sim U(1,000, 2,000)$,
- slope $b \sim U(500, 1,000)$,
- demand volatility is generated such that the coefficient of variation (cv) at mean price p equals 0.3 or 0.5, i.e. $cv = \sigma/\mu = 0.3$ or 0.5, respectively.

For a given instance, we sample $n \in \{50, 200\}$ normally distributed demand observations where the price is uniformly chosen from the interval [0, 1]. Figure 3.1 shows an illustrative example with 200 demand observations drawn for an instance with $a = 1471.8, b = 702.72$, and $\sigma = 360.7$.

From all n observations, the models are estimated and inventory levels are set such that a desired service level (either non-stockout probability or fill rate) of either 90 % or 95 % is met. For the cost model, we use a unit penalty cost of $v = 9$ instead of a non-stockout probability of 90 % and $v = 19$ for the 95 % case such that these two parameter settings yield the same critical fractile. Further, $h = 1$.

The safety stocks SI (except for the *LP* method) are set according to (3.1)–(3.4). For known parameters, given the price, mean demand is determined and the required safety stocks are obtained from (3.4). Ignoring demand dependency from prices and estimating demand parameters from the sample only (method of moments), μ and σ in (3.1)–(3.4) are replaced by the respective estimators $\hat{\mu}$ and $\hat{\sigma}$. Using the *OLS* method, $\hat{\mu} = \hat{\beta}_0 + \hat{\beta}_1 p$ and $\hat{\sigma}$ is estimated by (3.9). For the data set shown in Fig. 3.1 and $P_1 = 90\,\%$ this yields the following results:

i.) if all parameters are known, $Q = 1471.81 - 702.72p + 462.2$,
ii.) for the method of moments we estimate $\hat{\mu} = 1123.2$ and $\hat{\sigma} = 418.79$ and the inventory level resulting from (3.1)–(3.4) is $Q_{MM} = 1659.93$,

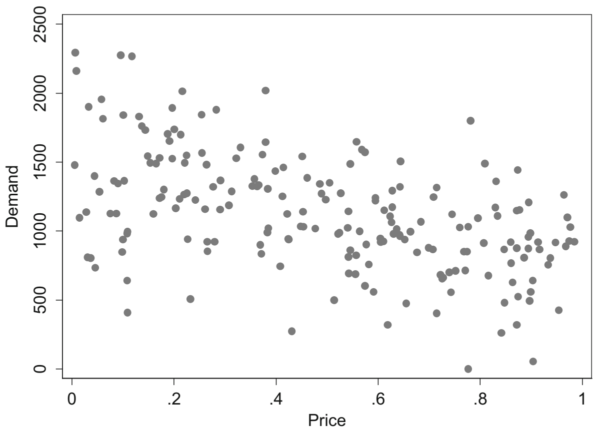


Fig. 3.1 Price-demand sampling example

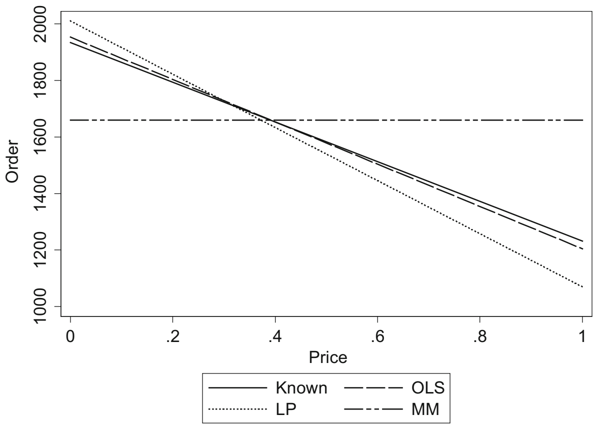


Fig. 3.2 Inventory level functions

iii.) *OLS* estimation yields $\hat{\beta}_0 = 1491.53$, $\hat{\beta}_1 = 748.46$, $\hat{\sigma} = 359.29$ with inventory level function

$$Q_{OLS} = 1491.53 - 748.46p + 459.89\sqrt{1 + \frac{1}{200} + \frac{(p-0.5)^2}{16.45}}$$
$$\approx 1491.53 - 748.46p + 461.61$$

iv.) the Linear Programming solution is $Q_{LP} = 2010.29 - 940.29p$. Figure 3.2 illustrates the resulting inventory levels as a function of price for the different methods of estimation.

After estimating the above inventory levels for n observations, further 100,000 demands are generated by first sampling a price uniformly from the interval [0, 1] and then sampling the demand under the true price-response function to test the estimates on out-of-sample observations.

The experiment was conducted for 500 randomly generated instances. Tables 3.1 and 3.2 summarize the average results with respect to achieved service level and average inventory level. Values given in Table 3.1 (3.2) are estimated from a sample of $n = 50$ (200). Note that the *LP Cost* approach was not applied to the cases with a fill rate constraint since the cost-service equivalence in this case depends on mean demand and standard deviation.

The method of moments not incorporating the price dependency performs worst with an average inventory increase of 20 % (6 %) for a required non-stockout probability and 33 % (9 %) for a fill-rate constraint if the coefficient of variation is 0.3 (0.5). The latter is caused by the fact that under a fill-rate constraint the safety factor in (3.4) depends on the demand forecast, too. Inventory level determination based on *OLS* slightly underachieves the service goals but reduces excess inventory. The *LP Service* level approach significantly underachieves the required service targets, however, this gap closes with a larger sample size and when a fill-rate rather

Table 3.1 Numerical results ($n = 50$)

$n = 50$		Average service levels (standard deviation)				
cv	SL (%)	Known	MM	OLS	LP Service	LP Cost
0.3	$P_1 = 90$	0.9 (0.0009)	0.8922 (0.0357)	0.8904 (0.0353)	0.8441 (0.0481)	0.8817 (0.0458)
0.3	$P_1 = 95$	0.95 (0.0007)	0.945 (0.0247)	0.9419 (0.0252)	0.9129 (0.0375)	0.945 (0.0375)
0.5	$P_1 = 90$	0.9 (0.0009)	0.8892 (0.0355)	0.8868 (0.036)	0.8465 (0.0479)	0.8821 (0.0453)
0.5	$P_1 = 95$	0.95 (0.0007)	0.9409 (0.0251)	0.9386 (0.0261)	0.9155 (0.0375)	0.9302 (0.037)
0.3	$P_2 = 90$	0.9001 (0.0004)	0.897 (0.0208)	0.8969 (0.0188)	0.8932 (0.0204)	–
0.3	$P_2 = 95$	0.95 (0.0003)	0.9477 (0.0152)	0.9469 (0.0141)	0.9433 (0.0155)	–
0.5	$P_2 = 90$	0.9008 (0.0006)	0.8935 (0.0242)	0.893 (0.0238)	0.8895 (0.0259)	–
0.5	$P_2 = 95$	0.9504 (0.0004)	0.9443 (0.0176)	0.9435 (0.0176)	0.9397 (0.0204)	–
$n = 50$		Average inventory levels (standard deviation)				
cv	SL (%)	Known	MM	OLS	LP Service	LP Cost
0.3	$P_1 = 90$	449.2 (122.6)	536.6 (117.8)	444.7 (132.4)	392 (123.2)	443.1 (139)
0.3	$P_1 = 95$	563.2 (153.6)	672.1 (145.6)	557 (164.5)	507.8 (156.3)	548 (171.2)
0.5	$P_1 = 90$	741.7 (203.9)	783.2 (203.2)	724.3 (217.9)	649 (207.2)	729.8 (230.7)
0.5	$P_1 = 95$	931.7 (255.5)	979.6 (251.6)	905.7 (270.2)	844.1 (263.9)	905.3 (284.2)
0.3	$P_2 = 90$	163.2 (42.1)	225.8 (47.3)	163.1 (49.1)	157.3 (51.1)	–
0.3	$P_2 = 95$	265 (70)	348.6 (73.3)	263.7 (79.7)	255.3 (81.0)	–
0.5	$P_2 = 90$	389.1 (103.6)	424 (105.1)	379.6 (114.6)	373.2 (118.8)	–
0.5	$P_2 = 95$	563.1 (151.3)	604.8 (152.6)	546.5 (165.3)	537.4 (169.5)	–

Table 3.2 Numerical results ($n = 200$)

$n = 200$		Average service levels (standard deviation)				
cv	SL (%)	Known	MM	OLS	LP Service	LP Cost
0.3	$P_1 = 90$	0.9 (0.0009)	0.8967 (0.0173)	0.8982 (0.0175)	0.8805 (0.0225)	0.896 (0.0217)
0.3	$P_1 = 95$	0.95 (0.0007)	0.949 (0.0113)	0.9483 (0.0118)	0.9333 (0.0177)	0.9455 (0.0164)
0.5	$P_1 = 90$	0.9 (0.0009)	0.8945 (0.0177)	0.895 (0.018)	0.881 (0.022)	0.8964 (0.0213)
0.5	$P_1 = 95$	0.95 (0.0007)	0.9455 (0.0121)	0.9454 (0.0125)	0.9339 (0.0174)	0.9456 (0.0163)
0.3	$P_2 = 90$	0.9001 (0.0004)	0.8986 (0.0102)	0.8996 (0.0097)	0.899 (0.0103)	–
0.3	$P_2 = 95$	0.95 (0.0003)	0.9492 (0.0073)	0.9494 (0.007)	0.9487 (0.0077)	–
0.5	$P_2 = 90$	0.9007 (0.0006)	0.8963 (0.0122)	0.8967 (0.0121)	0.8982 (0.0128)	–
0.5	$P_2 = 95$	0.9504 (0.0004)	0.9468 (0.0087)	0.9469 (0.0087)	0.9478 (0.0095)	–
$n = 200$		Average inventory levels (standard deviation)				
cv	SL (%)	Known	MM	OLS	LP Service	LP Cost
0.3	$P_1 = 90$	452.7 (119)	541.7 (104.9)	453.3 (124.5)	431.1 (119.2)	452.6 (125.6)
0.3	$P_1 = 95$	567.5 (149.1)	678.4 (130.5)	567.5 (155.2)	536 (149.0)	566.1 (158.8)
0.5	$P_1 = 90$	745.3 (198.1)	793.5 (188.9)	737.4 (205.8)	710.4 (198.2)	746.3 (209.2)
0.5	$P_1 = 95$	936.1 (248.3)	992.5 (235.2)	922.1 (256.3)	885.5 (249)	934.8 (264.8)
0.3	$P_2 = 90$	164.4 (40.9)	226.2 (36.4)	165.2 (44.2)	161.3 (45.8)	–
0.3	$P_2 = 95$	274 (87.7)	350 (58.7)	272.1 (82.1)	267.8 (83.6)	–
0.5	$P_2 = 90$	390.8 (100.7)	427.9 (93.8)	384.4 (106)	383.8 (110)	–
0.5	$P_2 = 95$	565.7 (147.1)	611.2 (137.5)	554.4 (153.6)	556.5 (158.5)	–

than a non-stockout constraint is used. In contrast, the *LP Cost* approach performs significantly better with average service levels slightly below target. Note that the underachievement of required service levels comes along with lower inventory levels.

3.4.1 Sample Size Effects

In the previous section, we have observed that the *LP Service* approach results in a P_1 level more than 5 % below target for small sample sizes. We estimate the parameters from samples of 20, 50, 100, 200, and 500 observations for a coefficient of variation of 0.5 and $P_1 = 95\,\%$. The results from testing the estimates on 100,000 out-of-sample observations are shown in Table 3.3.

Table 3.3 Sample size effects

$cv = 0.5$ $P_1 = 95\,\%$	Average service levels (standard deviation)				
n	Known	MM	OLS	LP Service	LP Cost
20	0.95 (0.0007)	0.9304 (0.0479)	0.9255 (0.0501)	0.8541 (0.0766)	0.9107 (0.0639)
50	0.95 (0.0007)	0.9409 (0.0251)	0.9386 (0.0261)	0.9155 (0.0375)	0.9302 (0.037)
100	0.95 (0.0007)	0.9437 (0.0186)	0.9437 (0.0185)	0.9231 (0.0265)	0.9421 (0.0240)
200	0.95 (0.0007)	0.9455 (0.0121)	0.9454 (0.0125)	0.9339 (0.0174)	0.9456 (0.0163)
500	0.95 (0.0007)	0.9461 (0.0075)	0.946 (0.0072)	0.9419 (0.0097)	0.9484 (0.0094)
$cv = 0.5$ $P_1 = 95\,\%$	Average inventory levels (standard deviation)				
n	Known	MM	OLS	LP Service	LP Cost
20	931.3 (242.9)	970.4 (300.9)	900.8 (311.5)	716 (282.6)	922.2 (387.3)
50	931.7 (255.5)	979.6 (251.6)	905.7 (270.2)	844.1 (263.9)	905.3 (284.2)
100	922.2 (248.4)	977.2 (238.2)	907.2 (256.9)	841.1 (236.9)	916.9 (263.5)
200	936.1 (248.3)	992.5 (235.2)	922.1 (256.3)	885.5 (249)	934.8 (264.8)
500	918.2 (245)	976.2 (227)	902.6 (248.2)	891.9 (245.8)	918.9 (253.3)

For a sample of 20 observations, the P_1 level achieved with the *LP Service* approach is almost 10 % below target, whereas for a sample size of 500 the service level nearly matches its target. In contrast, the *LP Cost* approach and *OLS* estimation attain higher service levels even with small sample sizes. The *MM* constantly achieves a service level above *OLS* and *LP*, but at the cost of a very high inventory level in all cases.

Kässmann et al. (1986) have come to similar results when comparing a service level to a cost approach. The *OLS*-based model follows a feedforward mechanism which determines optimal inventory levels based on sales forecasts. As opposed to this method the service level approach works as feedback model that corrects for positive and negative target deviations resulting from the parameter estimates. Similar to our results, the service-oriented feedback approach also achieves lower service levels than the feedforward approach. We can thus conclude that the sample size plays an important role concerning the accuracy of the estimates. Since the *MM* hedges against uncertainty with high inventory levels, its average service and inventory level are less affected.

3.4.2 Violations of OLS *Assumptions*

To assess the impact of violations of *OLS* assumptions, data is generated similarly to the above procedure with the following modifications. Tables 3.4, 3.5, and 3.6 contain a summary of the results for $n = 200$.

In order to test the effect of error terms with non-zero mean, we draw normally distributed disturbances with uniformly distributed mean e:

$u \sim N(e, \sigma^2)$ with $e \sim U(150, 300)$.

Note that the mean varies with each observation. The results in Table 3.4 show that both *OLS* and the *LP Cost* model cope well with the zero mean assumption violation. They achieve average service levels very close to the ones obtained by the known function. The service level model performs significantly better for the target fill rate than for in-stock probabilities, where it achieves only 98 % of the target P_1 service level.

Heteroscedasticity can be represented by standard deviations σ_i which depend on the level of the explanatory variable price. We assume the functional form (see Park 1966):

$$\sigma_i^2 = \sigma^2 X_i^{\Theta} e^{r_i} \tag{3.35}$$

with $\Theta = 3$ so that variations in demand increase with higher prices as displayed in Fig. 3.3.

The *LP Service* and the *LP Cost* model underachieve the P_1 service level but at lower inventory levels (see Table 3.5). In contrast, *OLS* overachieves a P_1 target of 90 %, but underachieves the 95 % target. In both cases, controlling for

Table 3.4 Violations of *OLS* assumptions ($n = 200$)—*non-zero mean*

$n = 200$	Average service levels (standard deviation)			
SL (%)	Known	OLS	LP Service	LP Cost
$P_1 = 90$	0.8977 (0.0178)	0.8971 (0.0179)	0.8801 (0.0229)	0.896 (0.0219)
$P_1 = 95$	0.9481 (0.0121)	0.9476 (0.0122)	0.9323 (0.0181)	0.945 (0.0166)
$P_2 = 90$	0.8993 (0.009)	0.8991 (0.009	0.8984 (0.0095)	–
$P_2 = 95$	0.9493 (0.0067)	0.9491 (0.0067)	0.9484 (0.0071)	–
$n = 200$	Average inventory levels (standard deviation)			
SL (%)	Known	OLS	LP Service	LP Cost
$P_1 = 90$	452.3 (123.8)	452.4 (123.8)	431.4 (120.2)	453.8 (126.7)
$P_1 = 95$	566.8 (154.6)	566.7 (154.5)	535 (147.4)	565.8 (156.4)
$P_2 = 90$	140.0 (42.9)	140.0 (43.0)	137.0 (43.4)	–
$P_2 = 95$	239.8 (71.1)	239.9 (71.1)	236.7 (71.5)	–

heteroscedasticity according to Park (1966) leads to average P_1 service levels closer to the target at lower inventory levels (see Table 3.5, marked with an asterisk).

For P_2 service level constraints, the *LP Service* model achieves a slightly lower service level, but at significantly less inventory. The *OLS* and *LP* model achieve higher fill rates with lower inventory than the known case. This at first sight counterintuitive result can be explained by the determination of service levels for different prices. While service level optimization with known parameters and the correction for heteroscedasticity lead to a good fit for all observations due to the price-dependent structure of the error term, *OLS* and *LP* achieve high service levels for low prices and thus small σ, but only small service levels for high prices which in the end averages to the required target fill rate.

Next, we replace our normally distributed observations by demand following the gamma distribution (see, e.g., Burgin 1975). Demand is distributed according to a $\Gamma(\lambda, \rho)$ distribution with $1/\lambda$ being the scale and ρ as shape parameter. Standard deviation and mean can be computed according to

$$\sigma = \frac{\sqrt{\rho}}{\lambda} = cv(a - 0.5b) \tag{3.36}$$

$$\mu = \frac{\rho}{\lambda} = a - bp \tag{3.37}$$

Table 3.5 Violations of *OLS* assumptions ($n = 200$)—*heteroscedasticity*

$n = 200$	Average service levels (standard deviation)			
SL (%)	Known	OLS	LP Service	LP Cost
$P_1 = 90$	0.9 (0.0009)	0.9165 (0.017) 0.906* (0.0242)*	0.8857 (0.0231)	0.8951 (0.0226)
$P_1 = 95$	0.95 (0.0006)	0.9449 (0.0132) 0.9524* (0.0169)*	0.9383 (0.018)	0.9458 (0.0169)
$P_2 = 90$	0.9 (0.0002)	0.9008 (0.0076) 0.901* (0.0071)*	0.8994 (0.0101)	–
$P_2 = 95$	0.95 (0.0002)	0.9603 (0.0061) 0.9509* (0.0061)*	0.9493 (0.0071)	–
$n = 200$	Average inventory levels (standard deviation)			
SL (%)	Known	OLS	LP Service	LP Cost
$P_1 = 90$	177.4 (48.6)	222.7 (66.1) 183.7* (57)*	150.4 (44.2)	176.6 (52.3)
$P_1 = 95$	222.5 (60.9)	278.6 (82.8) 230.7* (72.2)*	196.4 (58.8)	223.8 (65.5)*
$P_2 = 90$	47.2 (9.2)	35.4 (9.3) 49.1* (13.2)*	20 (6.8)	–
$P_2 = 95$	81.9 (18.3)	67.8 (20.6) 85.6* (25.3)*	51 (16.2)	–

Table 3.6 Violations of *OLS* assumptions ($n = 200$)—*gamma distribution*

$n = 200$	Average service levels (standard deviation)			
SL (%)	Known	OLS	LP Service	LP Cost
$P_1 = 90$	0.8931 (0.0013) 0.9005* (0.0012)*	0.8905 (0.0174) 0.8979* (0.0168)*	0.8816 (0.0208)	0.8961 (0.0205)
$P_1 = 95$	0.9322 (0.0009) 0.9498* (0.0007)*	0.9299 (0.0136) 0.9478* (0.0113)*	0.9342 (0.0163)	0.9456 (0.0149)
$P_2 = 90$	0.8891 (0.001) 0.8999* (0.0007)*	0.8875 (0.0139) 0.8986* (0.0113)*	0.8965 (0.0147)	–
$P_2 = 95$	0.9334 (0.0008) 0.9495* (0.0007)*	0.9318 (0.0108) 0.9482* (0.0081)*	0.9464 (0.0114)	–
$n = 200$	Average inventory levels (standard deviation)			
SL (%)	Known	OLS	LP Service	LP Cost
$P_1 = 90$	740.2 (206.1) 768.7* (214.1)*	737.2 (216) 765.4* (224.2)*	719.9 (218.2)	767.3 (231.5)
$P_1 = 95$	920.5 (256.4) 1039.2* (289.5)*	916.1 (268) 1033.9* (302.2)*	965.6 (292.1)	1038.4 (316.1)
$P_2 = 90$	394.4 (105.1) 419.5* (117)*	393 (113.5) 418.9* (123.4)*	415.5 (127.3)	–
$P_2 = 95$	568.1 (154.1) 661.5* (184.2)*	565.6 (164.6) 659* (193.3)*	654.8 (198)	–

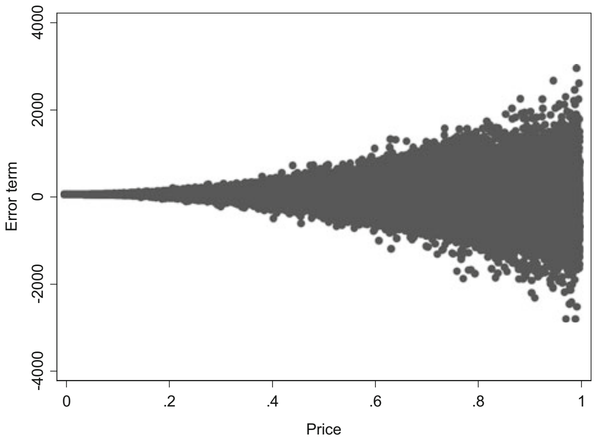


Fig. 3.3 Heteroscedasticity

and thus ρ and λ can be derived. For a coefficient of variation (cv) of 0.5, we compute required safety stock for normally and gamma distributed errors with safety factors based on Strijbosch and Moors (1999) (see Table 3.6). A comparison of the resulting inventory and service level shows that *OLS* also works well with gamma demand if the resulting inventory function is adjusted by a gamma safety factor (see Table 3.6, marked with an asterisk). Otherwise, if no attention is paid to the type of demand distribution, *OLS* underachieves the required service level. In contrast, the *LP* model does not require any prior knowledge on the type of demand distribution. Again, the *LP Cost* model proves to be superior to the service level model for P_1 concerning achievement of the target service level. For P_2, the *LP Service* approach performs well and produces results close to the target service level without accumulating large excess inventories.

In summary, the impact of violations of the standard *OLS* assumptions varies not only with the type of assumption, but also with the kind of service level used. While poor performance of *OLS* can be compensated by applying problem-specific remedies, the *LP* approach still outperforms these countermeasures in several cases in terms of inventory level.

3.4.3 Real Data

Since this research was originally inspired by an inventory management application at a large European retail chain, we assess the different models based on real data from 64 stores for a newsvendor-type product. The data contains daily sales for lettuce, customer demand, prices and weather information. Since true demand cannot be observed when demand exceeds supply, we exclude days when the product stocks out. After obtaining historical data for the first 50 days, the model parameters are specified in order to establish a function that is able to predict

Table 3.7 Example of the resulting order functions for one store

	Order function Q	P_1 service level	Inventory level
MM	84.01	0.87	34.42
OLS	$83.55 - 37.07p - 0.25w - 6.91d_1 + 13.98d_2 + 16.63$	0.89	28.75
LP Service	$82.61 - 25.57p - 0.15w - 10.31d_1 + 31.72d_2$	0.83	21.69
LP Cost	$95.03 - 22.5p - 19.25d_1 + 13.25d_2$	0.87	28.46

future demand. There are no structural breaks, seasonality or other factors during the time horizon considered. Therefore, it is reasonable to assume that customer demand is alike during the course of the experiment. Consequently, the functions established are then used to generate forecasts for the following days. For each one of these approximately 220 days (depending on product availability), predicted order quantities are compared to the observed demand realizations.

The explanatory variables are price, weather and weekdays. Prices recorded range from 0.29 € to 1.49 € and daily temperatures (w) lie between -6°C and 32°C. Furthermore, due to variations in demand during the week, weekdays are also included into the model. We segment weekdays into three categories: Tuesdays and Wednesdays generally exhibit the lowest demand levels (denoted by d_1), Fridays and Saturdays when maximum demand is observed (denoted by d_2) and Mondays and Thursdays. d_1 and d_2 are binary variables. Note that Mondays & Thursdays do not have an indicator variable to avoid collinearity.

Table 3.7 shows an example of a single store with the respective order functions for *MM*, *OLS*, *LP Service* and *LP Cost* and $P_1 = 90\,\%$. The last two columns contain the P_1 service level achieved and the resulting inventory levels.

Estimation of the multi-variable regression model yields goodness of fit measures R^2 between 0.25 and 0.74 with an average of 0.44. All estimated regression models are statistically significant. Results of our analysis are shown in Table 3.8. We observe that the method of moments best achieves the target P_1 service level of 90 %, but incurs inventory levels more than 30 % above those attained with the other approaches. The *LP Cost* approach exhibits the lowest variability compared to *OLS* and *LP Service*. Its constant service level comes at the cost of a slightly higher inventory level. In terms of underachievement, the *LP Service* model only achieves a P_1 10 % below target. This can be explained by sample size effects. Since the sample contains 50 observations, the *LP Service* model can be expected to perform worse than the other approaches due to the feedback mechanism. Using a larger sample size would solve this problem, but add new issues such as seasonality and trend.

The target fill rate is achieved equally well by all three approaches. The *MM* approaches hedge against uncertainty by accumulating larger inventory levels since it does not take any factors into account that allow to explain demand fluctuations.

Table 3.8 Results for real data

$n = 50$	Average service levels (standard deviation)			
SL (%)	MM	OLS	LP Service	LP Cost
$P_1 = 90$	0.8997 (0.0436)	0.8712 (0.0938)	0.8051 (0.0945)	0.8722 (0.064)
$P_2 = 90$	0.8653 (0.0430)	0.8579 (0.0602)	0.8561 (0.0709)	–
$n = 50$	Average inventory levels (standard deviation)			
SL (%)	MM	OLS	LP Service	LP Cost
$P_1 = 90$	50.2 (17.7)	36.5 (16.4)	28.5 (11.3)	38.5 (14.7)
$P_2 = 90$	26.6 (10.5)	17.6 (8.8)	17.2 (8.6)	–

For example, instead of stocking less on the low-demand weekdays and thereby reducing average inventory levels, the *MM* approach holds the same amount of inventory as on Fridays & Saturdays when demand is high. The *LP Service* and *OLS* approach take these effects into account which leads to overall lower inventory levels.

3.5 Conclusions

This work presents an integrated framework for demand estimation and safety stock planning in environments where demand depends on several external factors such as price and weather where a time series model would thus be inadequate. Using basic results from econometrics for causal demand forecasting and error estimation as well as a Linear Programming, data-driven method, target inventory levels to minimize cost or achieve desired service levels are calculated.

We illustrate and compare the proposed methods in a numerical experiment and a retail case application. As expected, the method of moments can achieve required service levels but at the cost of significantly higher inventory levels. The data-driven approach provides a robust method for inventory level determination. However, especially under non-stockout probability constraints, the data-driven approach exhibits an underachievement of required service levels if the size of the available data sample is too small. This problem can be avoided by applying the cost-service equivalence and using a cost minimization data-driven approach instead. The *OLS*-approach shows the best performance as long as the assumptions are valid. In case of misspecifications (heteroscedasticity, gamma distributed residuals), the robust approach dominates. Misspecification, however, can be overcome and appropriately addressed by adjusting the demand estimation or inventory level determination.

Chapter 4
The Data-Driven Newsvendor with Censored Demand Observations

Motivated by data from a large European retail chain, we tackle the newsvendor problem with censored demand observations by a distribution-free approach based on a data-driven approach. For this purpose, we extend the model introduced in Chap. 3. To improve the forecast accuracy, we simultaneously estimate unobservable lost sales, determine the coefficients of the exogenous variables which drive demand, and calculate the optimal order quantity. Since demand exceeding supply cannot be recorded, we use the timing of (hourly) sales occurrences to establish (daily) sales patterns. These sales patterns allow conclusions on the amount of unsatisfied demand and thus the true customer demand. To determine the coefficients of the inventory function, we formulate a Linear Programming model that balances inventory holding and penalty costs based on the censored demand observations. In a numerical study with data generated from the normal and the negative binomial distribution, we compare our model with other parametric and non-parametric estimation approaches. We evaluate the performance in terms of inventory and service level for (non-)price-dependent demands and different censoring levels. We find that the data-driven newsvendor model copes especially well with highly censored data and price-dependent demand. In most settings with price-dependent demand, it achieves similar or higher service levels by holding lower inventories than other benchmark approaches from the literature. Finally, we show that the non-parametric approaches are better than the parametric ones based on real data with several exogenous variables where the true demand distribution is unknown.

4.1 Introduction

In a lost sales inventory system, demand exceeding supply is usually not recorded. Studies have shown that *OOS* rates amount to 8.3 % of stock-keeping units (*SKUs*) per category worldwide considering non-perishable products

A.-L. Sachs, *Retail Analytics*, Lecture Notes in Economics and Mathematical Systems 680, DOI 10.1007/978-3-319-13305-8_4

(Corsten and Gruen 2003). *OOS* situations are estimated to occur even more frequently for perishable products, which is due to their short shelf-lives that make full product availability less desirable (ECR 2003). This leaves the store manager without sufficient knowledge on additional sales that could have been made had the inventory level been higher. Ignoring excess demand and sticking to the same order-up-to level results consecutively in demand misspecification and more lost sales (Nahmias 1994).Empirical studies of customer reactions to stockouts indicate that customers finding poor availability in a store on a regular basis will not only experience short-term lost sales but will decide not to return to this store in the long-run (Anderson et al. 2006). Retailers often apply rules of thumb to determine the optimal inventory level, underestimating forecast errors and levels which may affect the whole supply chain (Wagner 2002; Tiwari and Gavirneni 2007; Hosoda and Disney 2009).

Consequently, unobservable lost sales estimation is a key factor in inventory planning when it comes to determining optimal order quantities. Existing approaches can be categorized into parametric and non-parametric approaches. Parametric approaches assume some kind of underlying demand distribution. It is often questionable, whether this demand distribution is appropriate in practice. It is thus useful to formulate a non-parametric estimation approach that relies on data readily available to retailers and takes external factors with a strong impact on demand into account. One such factor is price which is usually tracked and linked to the sales quantity. Even though the inventory planning literature has accounted for price-dependent demand in numerous settings (Petruzzi and Dada 1999; Khouja 2000), it has not yet been considered in lost sales estimation. If we assume that demand is some function of the sales price, we can incorporate additional information into our approach to better explain demand variations (Fildes et al. 2008).

We suggest a novel approach based on data-driven optimization which overcomes the limitations of existing parametric and non-parametric approaches. There are no prior distribution assumptions and it only requires *POS* scanner data as typically available in retail stores. Observations including explanatory variables obtained from store-level scanner data are directly incorporated in inventory optimization.

4.2 Related Work

In the following, we will focus on the distribution-free newsvendor model and unobservable lost sales estimation.

The distribution-free newsvendor model goes back to Scarf (1958). The reader is referred to Gallego and Moon (1993) for a review. Bertsimas and Thiele (2005, 2006) solve the inventory planning problem as a Linear Program (*LP*) that works directly with historical demand observations. This robust, data-driven approach does not require any distributional assumptions. In Chap. 3 we extend this approach to causal forecasting to explain demand variations and directly estimate the optimal inventory level from historical demand data with several explanatory variables.

Lau and Lau (1996) propose a non-parametric model that does not require any prior distributional assumptions based on the product limit method (Kaplan and Meier 1958). They suggest establishing sales patterns based on the complete demand observations. If a stockout occurs, demand is only partly observed. Based on the level of the demand observations of previous hours with full availability, unobservable lost sales are then estimated for the remainder of the day according to the sales patterns. A Tocher-curve is then fitted to the estimated fractiles using regression. Huh et al. (2011) also use the Kaplan-Meier estimator in a newsvendor-type data-driven inventory model to show that it converges to the optimal inventory policy.

Parametric approaches are based on theoretical demand distributions such as the normal in Nahmias (1994), who suggests to approximate demand mean and standard deviation from sales data given an order-up-to level inventory policy. A Poisson process might be more suitable for discrete and small demands as in Conrad (1976) or the compound Poisson in Springael and van Nieuwenhuyse (2005). Agrawal and Smith (1996) emphasize that another important requirement of the distribution is to be capable of capturing the effects of demand variation as present in retailing and they suggest the negative binomial distribution by matching the sample mean and the frequency of facing zero demand to the observed frequency of demand.

Berk et al. (2007) use Bayesian updates for obtaining the parameter values of the negative binomial, Gamma, Poisson and normal distribution for the censored newsvendor problem. They rely on an approximation of the posterior distribution by matching the first two moments, given that one parameter is known (e.g., variance). Lu et al. (2006) consider Bayesian updates in the context of durable goods for a general distribution function and apply their findings to the normal distribution with known variance. Lu et al. (2008) analytically investigate the benefits of overstocking to learn about the true demand in a Bayesian setting. Tan and Karabati (2004) suggest an updating mechanism to achieve a desired service level by iteratively adjusting the inventory level.

4.3 Data-Driven Model with Unobservable Lost Sales Estimation

A retailer observes historical sales for a newsvendor product. Demand variability can be partially explained by external variables. Therefore, the data-driven newsvendor model determines the optimal order quantity as a function of the external variables (see Chap. 3). As an extension of this model, the retailer now only observes censored demands and has to estimate unobservable lost sales before

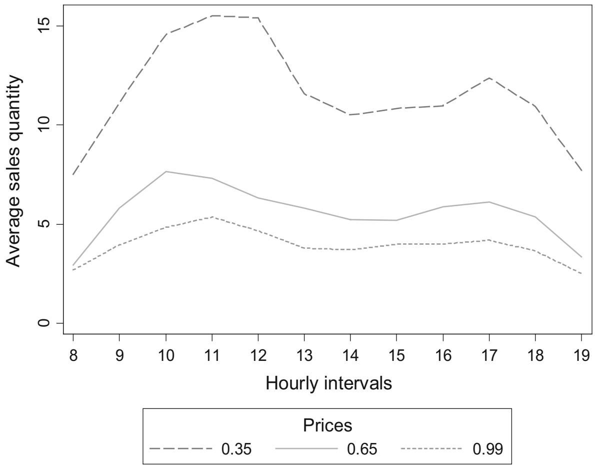


Fig. 4.1 Sales patterns for different prices

placing an order. Excess demand is lost so that sales data contain incomplete demand information. Any demand occurring after the product is out-of-stock remains unobserved.

Assuming that demand per day follows a similar pattern across all observations, we base our approach to estimate unobserved lost sales on Lau and Lau (1996). We assume that the level of the demand before the stockout occurs also reflects the influence of the external variables on demand. Figure 4.1 contains data on average hourly sales quantities of lettuce at one retail store as an example for a product with a short life-cycle, although this is not a newsvendor-product in the strong sense. The level of demand increases with lower prices, but the overall sales pattern is similar for low and high prices. Demand reaches its peak in the morning hours and around noon, then decreases, and in the late afternoon increases again.

4.3.1 Cost Model

The retailer has collected a set of $i = 1, \ldots, N$ sales observations and data on external variables X_{ji} per period i. The number of external variables to be considered is denoted by $j = 1, \ldots, m$.

If we assume that the product does not stockout, i.e. it is fully available in each period i, we determine the inventory level as a linear function of external factors influencing demand for a single-period newsvendor problem (see Chap. 3). The newsvendor's objective is to minimize total costs consisting of holding costs h for leftover inventory that is salvaged at the end of the day and penalty costs v per unit of unmet demand. The main idea is to fit a linear inventory function that determines the optimal order quantity depending on the values of the external variables X_{ji} to the sample data. The coefficients β_j correspond to these external variables including one coefficient β_0 for $X_{0i} = 1$ for the intercept. As a result, we obtain the target inventory level B_i as a product sum of the external factors with their respective coefficients

$$B_i = \sum_{j=0}^{m} \beta_j X_{ji}. \tag{4.1}$$

The coefficients of this linear inventory function are set such that a cost minimization objective or a target service level objective is achieved. The idea of fitting the coefficients to past demand observations for cost objectives or service level constraints (in-stock probability or fill-rate) is discussed in Chap. 3. The models therein assume that the retailer observes complete demand. This assumption may be violated if a retailer cannot track unobservable lost sales as is typically the case with perishable products such as fruits and vegetables.

In the following, we assume that stockouts are possible and lost sales are unobserved. The retailer then only observes sales, not demand. The decision-maker's objective is to minimize the sum of the penalty cost for unmet demand and inventory holding costs, but unmet demand is only known for days with complete demand observations. The sample of sales observations is grouped into full ($F = \{1, \ldots, c\}$) and censored ($C = \{c+1, \ldots, N\}$) demand observations. Each observation in $i = 1, \ldots, N$ is either assigned to F or C, depending on whether a stockout occurs on day i or not. If a product stocks out, then demand can only be observed until the time of the last sale which is usually recorded by point-of-sale scanner systems. If no stockout occurs, then demand can be observed for the entire day.

For all observations in set C (censored demand observations), the product is sold out before the store closes. Cumulative sales H_{ti} and sales h_{ti} at times t (usually per hour) are recorded by the *POS* scanner system until the product stocks out. The time point of the last sale, that is when the product stocks out, is recorded as $t = k_i$. This is an approximation because part of the demand in k_i is already lost.

Based on the complete demand observations in set F, sales patterns are defined according to the following procedure. Each day i is divided into $t = 1, \ldots, T$ discrete time intervals, e.g., hours. For each time interval, sales h_{ti} are recorded. Cumulative demand H_{ti} at the end of the day matches D_i when $t = T$, i.e., $H_{Ti} \equiv D_i$. Therefore, we know that the ratio of the mean of cumulative sales $\bar{H}_T$ to mean demand $\bar{D}$ of all demand observations is:

$$\frac{\bar{H}_T}{\bar{D}} = 1 \tag{4.2}$$

The ratio of cumulative demand H_t to total demand 1 h before closing ($t = T - 1$) can be calculated as the complete ratio less demand $h_T s$ that will occur in the meantime:

$$\frac{\bar{H}_{T-1}}{\bar{D}} = 1 - \frac{\bar{h}_T}{\bar{D}} \qquad \text{with } \bar{h}_t = \frac{\sum_{i=1}^{c} h_{ti}}{c} \tag{4.3}$$

The ratios of the preceding time intervals t can then be obtained recursively as

$$\frac{\bar{H}_t}{\bar{D}} = \frac{\bar{H}_{t+1}}{\bar{D}} - \frac{\bar{H}_{t+1}}{\bar{D}}\frac{\bar{h}_{t+1}}{\bar{H}_{t+1}} = \frac{\bar{H}_{t+1}}{\bar{D}}\left(1 - \frac{\bar{h}_{t+1}}{\bar{H}_{t+1}}\right) = \frac{1}{K_t} \tag{4.4}$$

It follows that

$$K_t = \frac{K_{t+1}}{\left(1 - \frac{\bar{h}_{t+1}}{\bar{H}_{t+1}}\right)} \qquad \text{with } K_T = 1. \tag{4.5}$$

The cumulative demand for all days with censored observations can then be estimated by interpolating the corresponding ratios before and after the stockout occurred in k_i. Since cumulative demand is multiplied with a term greater than one, it inflates the sales of the respective day.

$$D_i^k = H_{k_i i}\frac{(K_{k_i} + K_{k_i-1})}{2} \tag{4.6}$$

Based on these definitions, we can summarize the problem in a Linear Program. The decision variables are the coefficients of the inventory function, and sales s_i as well as leftover inventory y_i resulting from the chosen level of the inventory function.

$$\begin{aligned} \min C =& \sum_{i=1}^{N} h y_i + \sum_{i=1}^{c} v(D_i - s_i) \\ &+ \sum_{i=c+1}^{N} v\left(\frac{H_{k_i i}(K_{k_i-1} + K_{k_i})}{2} - s_i\right) \end{aligned} \tag{4.7}$$

s.t.

$$s_i \leq D_i, \forall i \in F \tag{4.8}$$

$$s_i \leq \frac{H_{k_i i}(K_{k_i} + K_{k_i-1})}{2} \qquad \forall i \in C \tag{4.9}$$

$$s_i \leq \sum_{j=0}^{m} \beta_j X_{ji} \qquad \forall i \in F \cup C \tag{4.10}$$

$$y_i \geq \sum_{j=0}^{m} \beta_j X_{ji} - D_i \qquad \forall i \in F \tag{4.11}$$

$$y_i \geq \sum_{j=0}^{m} \beta_j X_{ji} - \frac{H_{k_i i}(K_{k_i} + K_{k_i-1})}{2} \qquad \forall i \in C \tag{4.12}$$

$$s_i, y_i \geq 0 \tag{4.13}$$

$$\beta_j \in \Re \tag{4.14}$$

The objective function (4.7) consists of inventory holding costs for all observations $i = 1, \ldots, N$ and penalty costs for unmet demand. The penalty costs for unmet demand is split up into complete ($i = 1, \ldots, c$) and censored ($i = c + 1, \ldots, N$) demand observations. In the case of censored demand observations, an estimator based on (4.6) accounts for the unobservable lost sales.

For days in set F (full demand observations), we introduce a constraint (4.8) that the newsvendor can only meet incoming demand, i.e., sales s_i may not exceed demand D_i. Constraint (4.9) ensures that sales may not exceed the lost sales estimate. The newsvendor cannot sell more than the order quantity (4.10).

The leftover inventory y_i, which is discarded at the end of the day, is the difference between the order quantity and the incoming demand (4.11). As part of the objective function, the retailer has to consider potential penalty costs for the estimated unobservable lost sales in set C. The lost sales estimate reduces the leftover inventory as in (4.12).

4.3.2 Benchmark Approaches

Nahmias (1994) assumes that demand is normally distributed and an order-up-to policy with known constant S in place. There are $i = 1, .., c$ complete demands D_i sorted in increasing order. Given that a normal distribution is symmetric around its mean, the following equations derive the right tail of the distribution. Maximum likelihood estimators for mean ($\hat{\mu}$) and standard deviation ($\hat{\sigma}$) are:

$$\hat{\mu} = S - z\hat{\sigma} \tag{4.15}$$

$$\hat{\sigma} = 0.5(S - \overline{D}_c)(-z + \sqrt{z^2 + V^2}) \text{ with} \tag{4.16}$$

$$\overline{D}_c = (1/c) \sum_{i=1}^{c} D_i \tag{4.17}$$

Therefore, the authors derive an estimator for V and implicitly for z according to the following equations:

$$V^2 = 4\left(1 + \frac{\sum_{i=1}^{r}(D_i - \overline{D}_c)^2}{r(S - \overline{D}_c)^2}\right) \tag{4.18}$$

$$\frac{(1/\sqrt{2\pi})e^{-0.5z^2}}{(1/\sqrt{2\pi})\int_z^{\infty} e^{-0.5D^2} dD} = \frac{-c}{N-c}z + \frac{2c}{N-c}\frac{z+\sqrt{z^2+V^2}}{V^2} \tag{4.19}$$

For a more detailed description and proofs on how to derive these parameters, see the maximum likelihood estimators for normal demand in Nahmias (1994) and Halperin (1952).

Agrawal and Smith (1996) investigate empirical retail data and claim that the negative binomial distribution provides a better fit. They argue that the negative binomial distribution is appropriate for retail data that shows high variability due to external influences such as weather or promotions. The observed frequency values for each sales quantity smaller than the order up-to-level are unbiased estimators of the true frequency values. The observed frequency of demand j is denoted f_j^{obs}. The censored sample mean $\bar{x}_s$ is calculated according to:

$$\bar{x}_s = \sum_{j=0}^{S-1}(j-S)f_j^{obs} + S. \tag{4.20}$$

The aim is then to match the censored sample mean to the mean computed from the negative binomial distribution with unknown $\hat{r}$ and $\hat{q}$:

$$\mu(\hat{r},\hat{q}) = \sum_{j=0}^{S-1}(j-S)f_j(\hat{r},\hat{q}) + S. \tag{4.21}$$

This can be achieved by matching observed frequency F_L^{obs} and expected frequency of observing demand smaller or equal to L for any value $L < S$. For each r,

$$F_L^{obs} = \sum_{j=0}^{L} f_j(\hat{r},\hat{q}) \tag{4.22}$$

assigns a unique success probability $q(r)$ to r for any value $L < S$. According to the procedure by Agrawal and Smith (1996), L should be chosen such that $F_L^{obs} \geq 0.1$.

In a non-parametric approach, Lau and Lau (1996) first determine the fractiles of the left-hand tail of the demand distribution according to Kaplan-Meier's Product Limit method. Based on complete demand observations, daily-sales patterns with

$$R_t = R_{t+1} - \frac{\bar{h}_{t+1}}{\bar{H}_{t+1}}\frac{\bar{H}_{t+1}}{\bar{D}_{t+1}} \tag{4.23}$$

are established and censored observations are replaced by demand estimates

$$D_i^k = \frac{2S}{(R_{k_i} + R_{k_i - 1})} \tag{4.24}$$

which allow to calculate the fractiles of the right-hand tail of the distribution. Finally, a Tocher-curve with parameters (a, b, c, d, e) is fitted to the fractiles q using regression analysis. Fractiles are selected according to subjective elicitation sets. The number of fractiles depends on the portion of the demand distribution included in the data. Given the inverse cumulative demand distribution function with:

$$F_T^{-1}(q) = a + bq + cq^2 + d(1-q)^2 \ln(q) + eq^2 \ln(1-q), \tag{4.25}$$

optimal order quantities can be determined (Lau and Lau 1997). Additionally, we correct for external factors by adding a term for the coefficients g_j of the independent variables when fitting the Tocher curve:

$$F_T^{-1}(q) = a + bq + cq^2 + d(1-q)^2 \ln(q) + eq^2 \ln(1-q) + \sum_{j=1}^{m} g_j X_j \tag{4.26}$$

We then calculate the optimal inventory level from the Tocher-curve. This adaptation allows for a fair comparison with lower inventories in a setting where demand dependency on external factors can be observed.

4.4 Numerical Examples

We compare our approach with the above benchmark approaches from the existing literature on lost sales estimation in a controlled simulation experiment. First, we generate data from the normal and negative binomial demand distributions. We then censor the observations at the order-up-to level S. Holding costs are $h = 1$ and we compare four settings with different penalty costs ($v_1 = 1, v_2 = 3, v_3 = 9, v_4 = 19$). The ratio of holding and penalty costs corresponds to the service level measured as in-stock probability which we use as censoring level. The service levels of the respective order-up-to levels (or censoring levels) S are: $S_1 = 50\,\%, S_2 = 75\,\%, S_3 = 90\,\%$, and $S_4 = 95\,\%$. In the sequel, we pretend the true demand to be unknown to the newsvendor. We further assume that mean demand is a linear function of price

$$\mu_i = b_0 - b_1 p_i. \tag{4.27}$$

We compare two settings concerning price p_i. In the first setting, price is constant at $p_i = 0.5$ and in the second one, price is uniformly distributed on the interval $[0; 1]$.

For each type of distribution, we choose a common parametric estimation approach (*Par*) from the literature and the nonparametric approach (*Non*) according to Lau and Lau (1996) as a comparison. For the normal distribution, we follow the approach of Nahmias (1994), and for the negative binomial Agrawal and Smith

(1996). The column *Known* contains the results for a decision maker who has full information on the true parameters of the distribution. The results of the Linear Programming model if demand was fully observable are contained in *LP*. Since we artificially censor the demands, we know the true level of demand and use these values in the model named *LP*, because the model in Chap. 3 works only with full demand observations and does not account for any censoring. The model with censoring based on (4.7)–(4.14) is named *LPc* in the following.

To measure the performance, we draw a sample of N observations to estimate the parameters of the respective model and then calculate the resulting inventory and service level for an out-of-sample size of 100,000 observations. The following two sections deal with samples of $N = 200$ observations, whereas Sect. 4.4.3 compares the results for varying sample sizes. The experiment is repeated for 500 randomly generated instances.

4.4.1 The Normal Distribution

We assume that daily demand is a linear function of price as in (4.27) where b_0 as the market size is uniformly distributed on the interval [1000; 2000]. The slope b_1 is uniformly distributed on [500; 1000]. Each day, the store is open for 10 h. For simplicity of exposition, we assume that demand is equally spread, i.e., demand per hour is 1/10 of the daily demand. We generate demand variability by adding an error term that is normally distributed with mean 0 and a standard deviation chosen such that the coefficient of variation ($cv = \sigma/\mu$) equals 0.3 at mean price $p = 0.5$.

Using the parameters estimated on the basis of 200 historical observations, we compute the total cost as stated in the objective function for the out-of-sample observations. Table 4.1 contains the results.

Since existing approaches make use of the structure of the data, e.g., the shape of the normal distribution in Nahmias (1994), these approaches yield slightly lower costs (by less than 0.01 %) compared to the *LP* and *LPc* models for highly censored data if price is held constant. For a varying price (i.e., $p \sim [0; 1]$), the *LP* and *LPc* models show that capturing the price effect results in cost savings of up to 15 % compared to the parametric and the adjusted non-parametric approach.

In the following, we will analyze these results in more detail at a more disaggregate level by calculating the average service level and corresponding inventory levels for different coefficients of variation (0.3 and 0.5). The average service and inventory levels resulting from 500 replications for a constant price are shown Table 4.2. For low service and high censoring levels, all models underachieve the target service level. With increasing service and decreasing censoring levels, all models achieve service levels close to the target. The parametric approach remains slightly below the service target, but at an average inventory level of 8 % below the results with known parameters.

Table 4.1 Cost comparison for the normal distribution

cv	S (%)	Average cost for $p = 0.5$ (standard deviation)				
		Known	LP	LPc	Par	Non
0.3	50	269.03 (70.59)	271.26 (71.19)	271.26 (71.19)	271.05 (71.15)	269.54 (70.72)
0.3	75	428.58 (112.47)	430.70 (112.99)	432.76 (113.46)	431.17 (113.20)	429.82 (112.70)
0.3	90	591.77 (155.30)	595.77 (156.46)	601.18 (157.69)	595.82 (156.65)	596.81 (156.29)
0.3	95	695.57 (182.52)	703.38 (184.45)	710.87 (186.82)	702.05 (184.68)	704.17 (184.24)
cv	S (%)	Average cost for $p \sim [0; 1]$ (standard deviation)				
		Known	LP	LPc	Par	Non
0.3	50	272.75 (69.32)	274.51 (69.84)	274.58 (69.88)	322.85 (57.85)	324.85 (58.37)
0.3	75	434.52 (110.30)	437.39 (110.89)	437.91 (110.98)	512.88 (92.81)	517.89 (94.38)
0.3	90	600.10 (152.19)	606.90 (153.95)	607.32 (153.95)	704.36 (130.71)	711.04 (133.21)
0.3	95	705.47 (178.88)	717.40 (181.48)	718.83 (182.38)	824.83 (155.61)	831.56 (158.51)

The fewer demands are censored (and the larger the tail of the distribution that can be observed), the more does the parametric approach outperform the other models in terms of inventory, since estimators of mean and standard deviation can be calculated from a greater number of observations. However, with increasing demand variability ($cv = 0.5$), the parametric approach has more difficulty in achieving the target service level.

The non-parametric approach is closest to the target service level for all censoring levels, which is compensated by higher overall inventory levels (up to 3.2 % higher than the *Known* for S_3). Comparing coefficients of variation, the average inventory levels increase due to the additional amount of safety stock required to hedge against increasing demand uncertainty.

For samples with uniformly distributed price on the interval [0; 1] (see Table 4.3), demand variability can be partly explained by price changes. The parametric approach does not take the price information into account and adjusts for the additional variability by holding higher inventory. The adapted non-parametric approach still incurs higher inventory levels than the *LPc* model, but at the same time overachieves the required service levels of 75 %, 90 %, and 95 %. The results show that the *LPc* model copes very well with all censoring levels and adjusts the optimal order quantity according to the selling price.

Table 4.2 Numerical results for the normal distribution ($p = 0.5$)

cv	S (%)	Average service levels (standard deviation)				
		Known	LP	LPc	Par	Non
0.3	50	0.4991 (0.0004)	0.4508 (0.0146)	0.4508 (0.0146)	0.4521 (0.0100)	0.4801 (0.0146)
0.3	75	0.7497 (0.0012)	0.7280 (0.0244)	0.7398 (0.0430)	0.7186 (0.0176)	0.7540 (0.0238)
0.3	90	0.8999 (0.0010)	0.8934 (0.0200)	0.9026 (0.0297)	0.8841 (0.0148)	0.9061 (0.0211)
0.3	95	0.9499 (0.0007)	0.9460 (0.0156)	0.9513 (0.0203)	0.9400 (0.0119)	0.9522 (0.0157)
0.5	50	0.4991 (0.0004)	0.4508 (0.0146)	0.4508 (0.0146)	0.4493 (0.0096)	0.4744 (0.0141)
0.5	75	0.7497 (0.0012)	0.7280 (0.0244)	0.7354 (0.0355)	0.7136 (0.0172)	0.7460 (0.0228)
0.5	90	0.8999 (0.0010)	0.8934 (0.0200)	0.8990 (0.0254)	0.8797 (0.0147)	0.8985 (0.0212)
0.5	95	0.9499 (0.0007)	0.9460 (0.0156)	0.9492 (0.0181)	0.9367 (0.0122)	0.9481 (0.0159)
cv	S (%)	Average inventory levels (standard deviation)				
		Known	LP	LPc	Par	Non
0.3	50	134.23 (35.23)	115.00 (30.66)	115.00 (30.66)	115.47 (30.48)	126.54 (33.72)
0.3	75	277.40 (72.80)	261.44 (70.60)	272.19 (79.48)	254.20 (67.60)	281.59 (75.85)
0.3	90	447.85 (117.53)	439.02 (118.29)	459.34 (130.94)	423.11 (112.52)	462.05 (125.63)
0.3	95	561.50 (147.36)	554.80 (150.42)	577.92 (165.50)	535.05 (142.39)	574.92 (156.36)
0.5	50	219.01 (57.48)	186.97 (49.89)	186.97 (49.89)	185.93 (49.10)	202.36 (53.89)
0.5	75	457.64 (120.10)	431.04 (116.47)	441.85 (124.61)	413.02 (109.87)	453.75 (121.89)
0.5	90	741.71 (194.65)	727.00 (195.96)	746.24 (207.02)	689.52 (183.34)	742.43 (201.60)
0.5	95	931.14 (244.36)	919.97 (249.53)	940.94 (262.87)	872.67 (232.31)	931.06 (252.82)

Table 4.3 Numerical results for the normal distribution ($p \sim [0; 1]$)

cv	S (%)	Average service levels (standard deviation) Known	LP	LPc	Par	Non
0.3	50	0.4985 (0.0004)	0.5149 (0.0149)	0.5174 (0.0168)	0.5086 (0.0282)	0.5433 (0.0303)
0.3	75	0.7496 (0.0009)	0.7637 (0.0188)	0.7672 (0.0202)	0.7567 (0.0226)	0.7912 (0.0240)
0.3	90	0.8998 (0.0009)	0.9029 (0.0161)	0.9059 (0.0161)	0.9043 (0.0147)	0.9197 (0.0169)
0.3	95	0.9499 (0.0008)	0.9495 (0.0134)	0.9509 (0.0132)	0.9531 (0.0097)	0.9595 (0.0115)
0.5	50	0.4985 (0.0004)	0.5149 (0.0149)	0.5162 (0.0151)	0.5046 (0.0222)	0.5403 (0.0226)
0.5	75	0.7496 (0.0009)	0.7637 (0.0188)	0.7660 (0.0197)	0.7551 (0.0177)	0.7884 (0.0198)
0.5	90	0.8998 (0.0009)	0.9029 (0.0162)	0.9053 (0.0162)	0.9011 (0.0123)	0.9141 (0.0160)
0.5	95	0.9499 (0.0008)	0.9495 (0.0134)	0.9509 (0.0134)	0.9497 (0.0087)	0.9560 (0.0117)
cv	S (%)	Average inventory levels (standard deviation) Known	LP	LPc	Par	Non
0.3	50	135.90 (34.68)	143.86 (37.87)	144.89 (37.91)	165.89 (33.80)	184.63 (38.26)
0.3	75	281.10 (71.47)	294.22 (76.44)	297.50 (77.80)	341.60 (65.75)	377.56 (74.35)
0.3	90	454.00 (115.28)	464.31 (119.59)	469.65 (120.71)	546.82 (104.61)	582.95 (117.01)
0.3	95	569.30 (144.50)	577.98 (155.46)	583.95 (158.30)	682.49 (130.36)	710.96 (143.88)
0.5	50	219.78 (57.13)	233.07 (62.50)	233.97 (62.49)	239.77 (56.19)	268.44 (63.28)
0.5	75	461.77 (118.46)	483.54 (126.75)	487.55 (129.12)	503.42 (117.18)	555.13 (128.99)
0.5	90	749.94 (191.47)	767.19 (198.70)	775.04 (202.00)	807.72 (185.57)	853.13 (200.75)
0.5	95	942.11 (240.16)	956.62 (258.68)	966.73 (265.45)	1008.56 (232.23)	1049.51 (248.19)

4.4.2 The Negative Binomial Distribution

Assuming that retail demand might be overdispersed, we choose the negative binomial distribution. In our example, the mean of the negative binomial data varies across draws due to price changes and the variance is a function of the mean to ensure the property of overdispersion (Cameron and Trivedi 2005). We generate daily mean of demand as a function of price with $b_0 \sim U[5;10]$ and $b_1 \sim U[2;4]$. We assume the coefficient of variation equals 1. μ and σ can be transformed into the parameters of the negative binomial distribution $NB(\hat{q},\hat{r})$ as:

$$\hat{q} = \mu/\sigma^2 \tag{4.28}$$

$$\hat{r} = \frac{\mu^2}{\sigma^2 - \mu}. \tag{4.29}$$

$\hat{r}$ denotes the rth success and $\hat{q}$ the probability of success on a single trial. We then draw the demands and perform the censoring. A cost comparison for negative binomial data with $cv = 1$ in Table 4.4 shows that total costs for highly censored data are always lower (0.3–2.2 %) with the *LP* and *LPc* model, even for constant prices.

Table 4.4 Cost comparison for the negative binomial distribution

		Average cost for $p = 0.5$ (standard deviation)				
cv	S (%)	Known	LP	LPc	Par	Non
1	50	4.2425 (1.0212)	4.2586 (1.0271)	4.2588 (1.0266)	4.2756 (1.0338)	4.2643 (1.0244)
1	75	8.3241 (2.0411)	8.3701 (2.0503)	8.3712 (2.0512)	8.3898 (2.0639)	8.3791 (2.0528)
1	90	13.6195 (3.3912)	13.7527 (3.4228)	13.7643 (3.4294)	13.7761 (3.4466)	13.7719 (3.4288)
1	95	17.5831 (4.4112)	17.8604 (4.4887)	17.8829 (4.4978)	17.8404 (4.4961)	17.8857 (4.4958)
		Average cost for $p \sim [0;1]$ (standard deviation)				
cv	S (%)	Known	LP	LPc	Par	Non
1	50	4.2249 (1.0403)	4.2598 (1.0443)	4.2597 (1.0445)	4.3574 (1.0077)	4.3528 (1.0045)
1	75	8.3323 (2.0625)	8.4180 (2.0719)	8.4195 (2.0736)	8.5376 (2.0271)	8.5174 (2.0276)
1	90	13.6966 (3.3997)	13.9107 (3.4433)	13.9170 (3.4477)	13.9621 (3.3868)	13.9002 (3.4014)
1	95	17.7292 (4.4071)	18.1298 (4.5043)	18.1398 (4.5147)	18.0505 (4.4137)	17.9740 (4.4486)

Tables 4.5 and 4.6 contain data at the less aggregate level, displaying average service and inventory levels for constant and varying prices, respectively. We investigate whether increasing the level of dispersion has an effect on the accuracy of the estimates by using coefficients of variation of 1, 1.5, and 2.

For constant prices, the non-parametric approach incurs the lowest average inventory levels together with a general service underachievement. The other models tend to overestimate demand for high censoring. Other than for normally distributed demand, increasing the coefficient of variation greater than one yields lower inventory levels for highly censored data. Ridder et al. (1998) prove that larger variances may also result in lower costs in the newsvendor model. In our setting, the effect only occurs for low and is reversed for high service levels. The data reveals that the number of days with no demand increases strongly with higher coefficient of variations. Consequently, low service levels can easily be achieved with smaller order quantities which results in lower average inventory levels. Compared to the

Table 4.5 Numerical results for the negative binomial distribution ($p = 0.5$)

cv	S (%)	Average service levels (standard deviation)				
		Known	LP	LPc	Par	Non
1	50	0.5408 (0.0256)	0.5393 (0.0416)	0.5365 (0.0409)	0.5457 (0.0527)	0.4848 (0.0429)
1	75	0.7712 (0.0138)	0.7714 (0.0316)	0.7686 (0.0323)	0.7710 (0.0358)	0.7393 (0.0347)
1	90	0.9086 (0.0055)	0.9073 (0.0206)	0.9069 (0.0207)	0.9077 (0.0220)	0.8912 (0.0227)
1	95	0.9544 (0.0029)	0.9537 (0.0147)	0.9535 (0.0146)	0.9526 (0.0144)	0.9423 (0.0165)
1.5	50	0.5374 (0.0246)	0.5373 (0.0372)	0.5354 (0.0369)	0.5408 (0.0411)	0.4852 (0.0416)
1.5	75	0.7642 (0.0090)	0.7651 (0.0306)	0.7615 (0.0314)	0.7644 (0.0333)	0.7385 (0.0333)
1.5	90	0.9050 (0.0034)	0.9033 (0.0212)	0.9031 (0.0212)	0.9020 (0.0221)	0.8905 (0.0228)
1.5	95	0.9524 (0.0017)	0.9515 (0.0149)	0.9516 (0.0150)	0.9470 (0.0158)	0.9421 (0.0163)
2	50	0.5497 (0.0315)	0.5480 (0.0396)	0.5480 (0.0396)	0.5471 (0.0394)	0.4768 (0.0441)
2	75	0.7620 (0.0077)	0.7627 (0.0293)	0.7596 (0.0296)	0.7615 (0.0327)	0.7425 (0.0330)
2	90	0.9035 (0.0024)	0.9021 (0.0210)	0.9018 (0.0212)	0.8977 (0.0228)	0.8913 (0.0225)
2	95	0.9516 (0.0012)	0.9508 (0.0152)	0.9509 (0.0154)	0.9401 (0.0176)	0.9423 (0.0162)

(continued)

Table 4.5 (continued)

cv	S (%)	Average inventory levels (standard deviation)				
		Known	LP	LPc	Par	Non
1	50	1.26 (0.32)	1.27 (0.41)	1.27 (0.40)	1.33 (0.49)	1.19 (0.36)
1	75	3.95 (0.96)	4.01 (1.13)	4.01 (1.14)	4.02 (1.24)	3.85 (1.10)
1	90	8.45 (2.08)	8.49 (2.33)	8.56 (2.40)	8.56 (2.50)	8.15 (2.25)
1	95	12.17 (3.02)	12.35 (3.49)	12.53 (3.62)	12.21 (3.49)	11.63 (3.22)
1.5	50	0.93 (0.23)	0.96 (0.35)	0.96 (0.35)	1.00 (0.40)	0.88 (0.31)
1.5	75	4.46 (1.08)	4.56 (1.39)	4.56 (1.41)	4.57 (1.49)	4.30 (1.34)
1.5	90	11.84 (2.89)	11.91 (3.45)	12.03 (3.56)	11.82 (3.53)	11.27 (3.31)
1.5	95	18.61 (4.57)	18.94 (5.64)	19.28 (5.81)	17.92 (5.12)	17.52 (5.13)
2	50	0.44 (0.05)	0.47 (0.25)	0.48 (0.25)	0.47 (0.24)	0.38 (0.21)
2	75	3.98 (0.97)	4.11 (1.42)	4.10 (1.42)	4.12 (1.59)	3.88 (1.36)
2	90	13.88 (3.37)	14.05 (4.38)	14.24 (4.52)	13.48 (4.29)	13.15 (4.15)
2	95	24.04 (5.88)	24.64 (7.83)	25.08 (8.07)	21.24 (6.12)	22.49 (6.97)

other approaches, the *LPc* model best matches the target service level for data with high variability.

The results for data with price variability in Table 4.6 show that the non-parametric approach matches the target service levels well. This comes at the cost of high inventories. The *LPc* model also suffers from this drawback for high variability that cannot be explained by price ($cv = 2, S_4$). For lower coefficients of variation, the *LPc* model achieves service levels similar to the non-parametric model while building up fewer inventories. Compared to the dataset with constant prices, inventory levels of the parametric and non-parametric model are higher. In the case of the parametric model this is due to the unexplained demand variability resulting from varying prices. In contrast, the *LP* and *LPc* model capture the price effects and thus incur less inventory than with constant price for $cv = 1$ and $cv = 1.5$.

4.4.3 Sample Size Effects

In order to determine whether the approaches are also capable of estimating valid parameters for small sample sizes or get better with larger samples, we compare the results of the negative binomial distribution with $cv = 1.5$ and price variation for samples with 20–500 observations.

For small samples ($n = 20$), all approaches except the parametric one provide reasonable service levels. For the *LP* and *LPc* model, increasing sample size from 20 to 50 observations results in better service level achievement and at the same time lower inventories. Increasing the sample size further has only a small impact for the *LP* and *LPc* model since estimates obtained with small samples are already quite accurate. The strongest effect of the sample size can be observed for the parametric

Table 4.6 Numerical results for the negative binomial distribution ($p \sim [0; 1]$)

cv	S (%)	Average service levels (standard deviation)				
		Known	LP	LPc	Par	Non
1	50	0.5409 (0.0108)	0.5021 (0.0338)	0.4984 (0.0340)	0.5505 (0.0456)	0.5018 (0.0410)
1	75	0.7709 (0.0055)	0.7486 (0.0305)	0.7477 (0.0311)	0.7803 (0.0361)	0.7535 (0.0321)
1	90	0.9086 (0.0024)	0.8989 (0.0212)	0.8987 (0.0216)	0.9157 (0.0209)	0.9041 (0.0179)
1	95	0.9544 (0.0013)	0.9487 (0.0152)	0.9487 (0.0152)	0.9586 (0.0129)	0.9515 (0.0121)
1.5	50	0.5417 (0.0112)	0.5067 (0.0333)	0.5028 (0.0331)	0.5381 (0.0436)	0.4936 (0.0374)
1.5	75	0.7653 (0.0044)	0.7488 (0.0307)	0.7479 (0.0308)	0.7643 (0.0327)	0.7469 (0.0322)
1.5	90	0.9051 (0.0038)	0.8978 (0.0211)	0.8980 (0.0212)	0.9011 (0.0212)	0.9044 (0.0172)
1.5	95	0.9522 (0.0040)	0.9479 (0.0152)	0.9480 (0.0151)	0.9448 (0.0147)	0.951 (0.0124)
2	50	0.5456 (0.0154)	0.5205 (0.0321)	0.5132 (0.0336)	0.5454 (0.0368)	0.4893 (0.0464)
2	75	0.7545 (0.0227)	0.7481 (0.0301)	0.7481 (0.0305)	0.7633 (0.0327)	0.7502 (0.0308)
2	90	0.8942 (0.0284)	0.8933 (0.0259)	0.8961 (0.0223)	0.8959 (0.0244)	0.9027 (0.0198)
2	95	0.9418 (0.0302)	0.9418 (0.0262)	0.9437 (0.0238)	0.9357 (0.0244)	0.9467 (0.0232)

(continued)

Table 4.6 (continued)

cv	S (%)	Average inventory levels (standard deviation)				
		Known	LP	LPc	Par	Non
1	50	1.23 (0.30)	1.27 (0.37)	1.24 (0.37)	1.40 (0.43)	1.33 (0.38)
1	75	3.91 (0.96)	3.93 (1.09)	3.92 (1.09)	4.31 (1.18)	4.17 (1.10)
1	90	8.45 (2.09)	8.48 (2.46)	8.50 (2.50)	9.13 (2.36)	8.92 (2.38)
1	95	12.22 (3.03)	12.29 (3.56)	12.34 (3.63)	13.02 (3.33)	12.66 (3.47)
1.5	50	0.92 (0.22)	0.98 (0.31)	0.95 (0.31)	1.01 (0.42)	0.92 (0.32)
1.5	75	4.37 (1.10)	4.45 (1.36)	4.43 (1.35)	4.71 (1.42)	4.56 (1.37)
1.5	90	11.74 (2.95)	11.87 (3.66)	11.96 (3.76)	12.20 (3.43)	12.60 (3.53)
1.5	95	18.57 (4.64)	18.77 (5.72)	19.22 (8.01)	18.42 (4.89)	19.56 (5.56)
2	50	0.46 (0.10)	0.55 (0.22)	0.51 (0.25)	0.50 (0.27)	0.44 (0.24)
2	75	3.89 (0.98)	4.20 (1.29)	4.28 (1.52)	4.51 (1.72)	4.46 (1.73)
2	90	13.62 (3.55)	14.46 (4.40)	17.72 (13.96)	14.70 (4.54)	17.16 (7.93)
2	95	23.74 (6.20)	25.31 (7.32)	34.18 (28.03)	23.31 (7.02)	29.64 (13.71)

approach. Service deviates from the target values by up to 21 % for small sample sizes, but results are reasonably accurate for 500 observations (Table 4.7).

4.4.4 Real Data

This work is motivated by observations made at a large retail chain. Out-of-stock situations are particularly prevalent in fresh product assortments. Since leftover inventories have to be discarded, it is not desired to have full availability until the store closes. We therefore investigate the performance of the models discussed above on data from 64 stores and three perishable products (vegetables). In addition to hourly sales data, we collected information on the time of the last sale, price and weather forecast for 3 years: from 10/2008 to 10/2011. Additionally, we include weekdays as independent variables. The stores are open Monday through Saturday.

Prices lie between 0.09 € and 1.79 € for two sorts of lettuce and 0.85 € and 1.79 € for mushrooms. Maximum temperatures range between −11° and 38°C.

We only include days with full availability in order to have information on the true demand values. Availability strongly varies with product and store. We then artificially censor demand at order-up-to levels S and pretend not to know hourly sales as recorded by the POS scanner. We then divide all observations into a sample

Table 4.7 Sample size effects for the negative binomial distribution ($p \sim [0; 1]$, $cv = 1.5$)

		Average service levels (standard deviation)				
n	S (%)	Known	LP	LPc	Par	Non
20	50	0.5417 (0.0112)	0.5235 (0.0975)	0.5168 (0.0987)	0.5330 (0.1023)	0.4743 (0.1064)
	75	0.7653 (0.0044)	0.7406 (0.0891)	0.7358 (0.0901)	0.6892 (0.0951)	0.7189 (0.0959)
	90	0.9051 (0.0038)	0.8786 (0.0690)	0.8775 (0.0688)	0.7410 (0.1036)	0.8643 (0.0683)
	95	0.9522 (0.0040)	0.9237 (0.0596)	0.9239 (0.0593)	0.7503 (0.1092)	0.9049 (0.0629)
50	50	0.5417 (0.0112)	0.5158 (0.0673)	0.5110 (0.0674)	0.5465 (0.0741)	0.4867 (0.0679)
	75	0.7653 (0.0044)	0.7419 (0.0625)	0.7408 (0.0630)	0.7551 (0.0654)	0.7400 (0.0635)
	90	0.9051 (0.0038)	0.8906 (0.0443)	0.8908 (0.0446)	0.8600 (0.0577)	0.8902 (0.0386)
	95	0.9522 (0.0040)	0.9363 (0.0341)	0.9365 (0.0341)	0.8919 (0.0594)	0.9346 (0.0314)
100	50	0.5417 (0.0112)	0.5096 (0.0507)	0.5055 (0.0505)	0.5386 (0.0531)	0.4888 (0.0532)
	75	0.7653 (0.0044)	0.7465 (0.0443)	0.7453 (0.0447)	0.7630 (0.0479)	0.7440 (0.0459)
	90	0.9051 (0.0038)	0.8937 (0.0308)	0.8940 (0.0308)	0.8911 (0.0332)	0.8991 (0.0256))
	95	0.9522 (0.0040)	0.9448 (0.0221)	0.9447 (0.0219)	0.9309 (0.0276)	0.9456 (0.0194)
500	50	0.5417 (0.0112)	0.5070 (0.0241)	0.5033 (0.0230)	0.5406 (0.0311)	0.4943 (0.0310)
	75	0.7653 (0.0044)	0.7507 (0.0198)	0.7498 (0.0199)	0.7651 (0.0225)	0.7487 (0.0216)
	90	0.9051 (0.0038)	0.9007 (0.0134)	0.9005 (0.0134)	0.9046 (0.0135)	0.9082 (0.0109)
	95	0.9522 (0.0040)	0.9502 (0.0096)	0.9503 (0.0096)	0.9500 (0.0087)	0.9559 (0.0077)

(continued)

Table 4.7 (continued)

n	S (%)	Average inventory levels (standard deviation)				
		Known	LP	LPc	Par	Non
20	50	0.92 (0.22)	1.35 (1.05)	1.32 (1.06)	1.20 (1.34)	0.97 (0.83)
	75	4.37 (1.10)	5.16 (3.21)	5.11 (3.21)	3.24 (1.88)	4.42 (2.62)
	90	11.74 (2.95)	13.37 (7.69)	13.57 (8.51)	4.64 (2.74)	10.70 (5.62)
	95	18.57 (4.64)	21.16 (13.14)	21.83 (14.96)	5.15 (3.57)	15.19 (8.52)
50	50	0.92 (0.22)	1.12 (0.62)	1.09 (0.63)	1.20 (1.01)	0.95 (0.53)
	75	4.37 (1.10)	4.60 (2.06)	4.57 (2.06)	4.72 (2.25)	4.59 (1.99)
	90	11.74 (2.95)	12.26 (5.20)	12.53 (6.08)	9.44 (3.70)	11.79 (4.36)
	95	18.57 (4.64)	18.45 (7.53)	18.95 (9.68)	12.29 (4.79)	17.37 (6.48)
100	50	0.92 (0.22)	1.02 (0.44)	0.99 (0.44)	1.04 (0.51)	0.92 (0.41)
	75	4.37 (1.10)	4.49 (1.62)	4.48 (1.63)	4.80 (1.87)	4.55 (1.63)
	90	11.74 (2.95)	11.82 (4.13)	12.01 (4.90)	11.47 (3.68)	12.26 (3.83)
	95	18.57 (4.64)	18.87 (6.57)	19.20 (7.82)	16.29 (4.86)	18.68 (5.99)
500	50	0.92 (0.22)	0.97 (0.27)	0.94 (0.27)	1.00 (0.32)	0.92 (0.28)
	75	4.37 (1.10)	4.45 (1.23)	4.43 (1.22)	4.68 (1.22)	4.55 (1.21)
	90	11.74 (2.95)	11.95 (3.31)	11.94 (3.33)	12.40 (3.14)	12.87 (3.31)
	95	18.57 (4.64)	18.79 (5.07)	19.02 (6.01)	19.32 (4.88)	20.27 (5.22)

of size 250 for parameter estimation and out-of-sample observations. The out-of-sample size depends on the individual availabilities of each product at a store.

Table 4.8 shows a summary of the results: the non-parametric approach captures low target service levels very well and has the lowest inventory of all. However, it accumulates more inventory to achieve high service levels. Both parametric approaches underachieve target service levels, especially for high censoring. The parametric approaches have a major advantage in the examples with simulated

Table 4.8 Numerical results for real data

S (%)	LP	Average service levels (standard deviation) LPc	Par Normal	Par Neg.Bin.	Non
50	0.4326 (0.1571)	0.4559 (0.1708)	0.4063 (0.1424)	0.3972 (0.1388)	0.4559 (0.1449)
75	0.6976 (0.1589)	0.7197 (0.1649)	0.6800 (0.1252)	0.6681 (0.1277)	0.6841 (0.1494)
90	0.8729 (0.1021)	0.8859 (0.0961)	0.8616 (0.0735)	0.8444 (0.0787)	0.9216 (0.0447)
95	0.9354 (0.0631)	0.9398 (0.0593)	0.9252 (0.0446)	0.9064 (0.0539)	0.9760 (0.0198)
S (%)	LP	Average inventory levels (standard deviation) LPc	Par Normal	Par Neg.Bin.	Non
50	4.04 (2.90)	4.51 (3.23)	3.75 (2.87)	3.19 (2.28)	4.40 (3.27)
75	10.27 (6.35)	11.41 (7.20)	10.73 (6.92)	9.52 (5.74)	11.42 (8.10)
90	19.89 (10.38)	22.13 (12.18)	21.68 (12.14)	18.93 (9.58)	28.71 (14.53)
95	27.92 (13.31)	30.42 (15.13)	29.74 (15.14)	25.46 (11.74)	43.46 (18.42)

data from a theoretical demand distribution as they are targeted at this specific type of distribution. According to the Kolmogorov-Smirnov test, the assumption that demand follows the normal distribution can be rejected for all datasets. The assumption of the negative binomial distribution can be rejected for 65 of the 192 datasets at the 0.1 level. Since the real data does not always follow a specific theoretical distribution, the parametric approaches make assumptions on the form of the distribution that might not be true. The *LPc* model performs well for all levels of censoring and service. The inventory level is only slightly higher than the results of the non-parametric approach for low service levels, but achieves high service levels of 90 % (95 %) with 23 % (30 %) less inventories.

4.5 Conclusions

In retailing, out-of-stock situations are a common phenomenon. By taking this effect—as well as additional information that explains parts of the demand variability—into account, we extended the data-driven newsvendor model to censored demand observations. This distribution-free approach estimates the parameters of a linear inventory function directly from the sales data. Additionally, we establish daily sales patterns from full demand observations which can be applied to extrapolate censored observations.

In a numerical study, we compare the performance of the *LP* model with existing parametric and non-parametric approaches for different censoring levels and two demand distributions commonly found to fit retail data. Compared to the other models, the *LP* model shows improved performance, especially for highly censored data and price effects. It copes well both with normally distributed data at different coefficients of variation and with overdispersed data from a negative binomial distribution.

Chapter 5
Data-Driven Order Policies with Censored Demand and Substitution in Retailing

We extend the data-driven inventory model with censored demand to the two-product case with stockout-based substitution. If one product stocks out, a fraction of the demand that cannot be satisfied is shifted to the substitute. As a result, sales of the substitute are inflated by the additional demand. The amount of substituted demand as well as unobservable lost sales are estimated based on the timing of stockout events. Similarly to the model in Chap. 4, we establish sales patterns based on the hourly sales observations before a stockout occurs. Our numerical study and data from a large European retail chain shows that the data-driven model achieves higher average profits than an existing approach from the literature. Investigating the trade-off between learning about substitution behavior from highly censored data versus learning about demand from little censored data, we find that more learning about substitution yields slightly better results in terms of profits.

5.1 Motivation

Sales and operations planning in retail can considerably benefit from the availability of big data provided by point-of-sale demand information. Nevertheless, the solution of the trade-off between ordering too many and too few is still a challenge for operations practice, in particular when managing perishable products. This challenge results in a considerable amount of product waste on the one hand (Gustavsson et al. 2011) and stockouts on the other hand (Aastrup and Kotzab 2010). The practical retail decision problem is further complicated by the fact that part of the required data for decision making is not readily available from the large amount of sales data, that is lost sales and customer substitution behavior is typically unobserved when products are out-of-stock. On average, 8.3 % of all products are typically found to be out-of-stock in retailing (Corsten and Gruen 2003). This number is usually even higher for perishable products where leftover inventories

A.-L. Sachs, *Retail Analytics*, Lecture Notes in Economics and Mathematical Systems 680, DOI 10.1007/978-3-319-13305-8_5

can only be carried for a short period of time (ECR 2003). If customers do not find their first choice available, they may be willing to compensate this lack by buying a substitute instead or otherwise do not make or postpone the purchase.

In the existing retail operations literature, these particular challenges are addressed, but usually distributional and prior assumptions on demands and substitution rates are made. In our research, we follow a more recent approach of data-driven optimization (Bertsimas and Thiele 2006; Huh et al. 2011; Beutel and Minner 2012) to provide an integrated forecasting and inventory level optimization framework. The general idea of this approach is to use the historical data (typically multivariate data including sales prices, weather and other demand influencing environmental variables) directly to optimize a (linear) target inventory level as a function of the respective environmental variables and then use this functional relationship to provide ordering decisions. The problem to optimize the inventory level function in a multi-product assortment is modeled and solved as a Mixed-Integer Linear Program and therefore can rely on the large-scale solution capabilities of standard solvers which have successfully been applied for many other operations problems in practice. Following this approach does not require the analysis and parameterization of distributional assumptions which usually will also vary between products and stores and therefore does not provide a robust framework for proposing inventory level decisions.

We collected 4 years of hourly sales data including prices of perishable products of 66 stores of a large European retail chain. This data set is used to illustrate the functionality of our data-driven framework and to quantify the benefits and robustness compared to existing approaches from the literature.

From a customer's point of view, willingness to substitute depends on the products' characteristics and availability. From a retail company's point of view, holding large amounts of stock is costly and a store manager often prefers not to satisfy all the demand but rather shift some demand to other products. In order to control demand, a store manager needs a clear understanding of the factors influencing demand as well as unobservable lost sales and substitution behavior.

5.2 Related Work

Substitution can either be controlled by the supplier or by the consumer. In the former case, the supplier decides to satisfy customer demand from another product that has the same or superior properties, such as a faster computer chip. Thereby, the supplier can hold lower levels of inventory. For a review on supplier-controlled substitution, the interested reader is referred to Hsu and Bassok (1999). The latter case is usually found in retailing where the consumers themselves choose which products they buy to satisfy demand if their first choice is out-of-stock.

One important aspect is to estimate the willingness of customers to substitute or not to purchase a product if it is out-of-stock. Existing approaches often assume that demand follows some theoretical distribution which can be used to estimate unobservable lost sales, such as the normal (Nahmias 1994) or the negative binomial distribution (Agrawal and Smith 1996). Substitution probabilities can be derived based on similar assumptions such as the negative binomial in Smith and Agrawal (2000) or the Poisson process in Anupindi et al. (1998).

Another important aspect is to include substitution into inventory optimization. One stream of research focuses on the decision which products to offer in an assortment (van Ryzin and Mahajan 1999; Mahajan and van Ryzin 2001a; Gaur and Honhon 2006; Caro and Gallien 2007; Hopp and Xu 2008; Ulu et al. 2012; Sauré and Zeevi 2013) while another stream of research including this work studies inventory optimization decisions given a fixed assortment (Nagarajan and Rajagopalan 2008), or a combination of both (Topaloglu 2013).

The literature on inventory optimization with consumer-driven demand substitution goes back to the works by McGillivray and Silver (1978) and Parlar and Goyal (1984). McGillivray and Silver (1978) solve the two-product problem with substitution by heuristic policies. Parlar and Goyal (1984) show that the profit function for two products with substitution is concave. Netessine and Rudi (2003) extend their work by analyzing the n-product case in a competitive and a non-competitive environment.

While Anupindi et al. (1998) focus on estimating substitution probabilities without inventory optimization, Smith and Agrawal (2000) also study a newsvendor setting with substitution. As opposed to our approach, demand substitution and inventory optimization are calculated sequentially. Chen and Plambeck (2008) suggest an inventory management approach with learning about the distribution function. Information about the demand distribution and substitution probabilities is continuously updated using Bayesian learning. Vulcano et al. (2012) estimate primary demand and lost sales from sales data by determining preference weights and customer arrival rates. Their approach requires, however, information on a retailer's aggregate market share.

Several approaches working with the multinomial logit choice model to analyze demand require data on customer arrivals (Kök and Fisher 2007; Karabati et al. 2009) which is not always available. A number of other approaches with focus on inventory optimization assume that substitution probabilities are known (Caro and Gallien 2007). Netessine and Rudi (2003) study the multiple product inventory control problem with stockout-based substitution for centralized and decentralized decision-making. The focus of their work lies on the inventory optimization problem. Demand substitution occurs with deterministic rates. Our approach integrates inventory optimization and estimation of demand including unobservable lost sales and substitution behavior.

Gilland and Heese (2013) analyze the importance of the sequence of customer arrivals. In a retail setting with two mutual substitutes and shortage penalty costs, the profit gained depends on whether a product satisfies a customer's first choice (no shortage penalty) or whether it serves as a substitute (shortage penalty for not satisfying the customer's first choice incurs).

Aydin and Porteus (2008) consider the joint inventory and pricing problem for multiple products with price-based, but no stockout-based substitution. Karakul and

Chan (2008) study the joint inventory and pricing problem for multiple products with stockout-based substitution. Even though they account for price-dependent demand behavior, they do not consider that substitution itself may be price-dependent as well. We study an inventory problem for multiple products with price-based and stockout-based substitution, but without price optimization.

For a comprehensive review on perishable inventory control with substitution and related aspects refer to Kök et al. (2008), Pentico (2008), and Karaesmen et al. (2011).

There is still a large gap between models existing in the literature and decision-making in practice which is mostly based on subjective judgements (Tiwari and Gavirneni 2007; Fisher 2009). This is partly due to the fact that store-level *POS* scanner data does not contain all information required to apply these models or that distributional assumptions cannot be generalized to all products and retail environments in practice. Additionally, demand and substitution estimation are often separated from inventory optimization which is essentially one task for a store manager.

As opposed to the above literature, we extend the inventory optimization problem by introducing stockout-based substitution rates in a data-driven multi-product newsvendor setting. Our model does not require any data on customer-specific shopping behavior such as individual departure times. Our methodology is based on the data-driven newsvendor model (Bertsimas and Thiele 2006; Beutel and Minner 2012; Sachs and Minner 2014) where optimal order quantities are calculated directly from the data and which is distribution-free.

5.3 Model

A retailer offers two products within a category in a fixed assortment. Depending on the personal preferences, a customer may choose to purchase a substitute product if the first choice product is out-of-stock. We assume that a fixed proportion (rate) of unsatisfied demand is shifted to the substitute product which is common in the literature, e.g., Parlar and Goyal (1984), Netessine and Rudi (2003), and Nagarajan and Rajagopalan (2008). We consider only single-attempt substitution. If the substitute product is also out-of-stock, the sale is lost. If a stockout situation occurs and the customer is not willing to purchase another product instead, lost sales occur which are not recorded. We assume no shortage penalty costs so that the sequence of customer arrivals does not affect the model (see Gilland and Heese 2013 for an analysis).

5.3.1 Data

A retailer has collected *POS* scanner data for two newsvendor-type products ($i = 1, 2$) over several periods $t = 1, \dots, T$. The respective substitute of product i is denoted as j with $j = 1, 2$. The store is open Z hours per day ($z = 1, \dots, Z$).

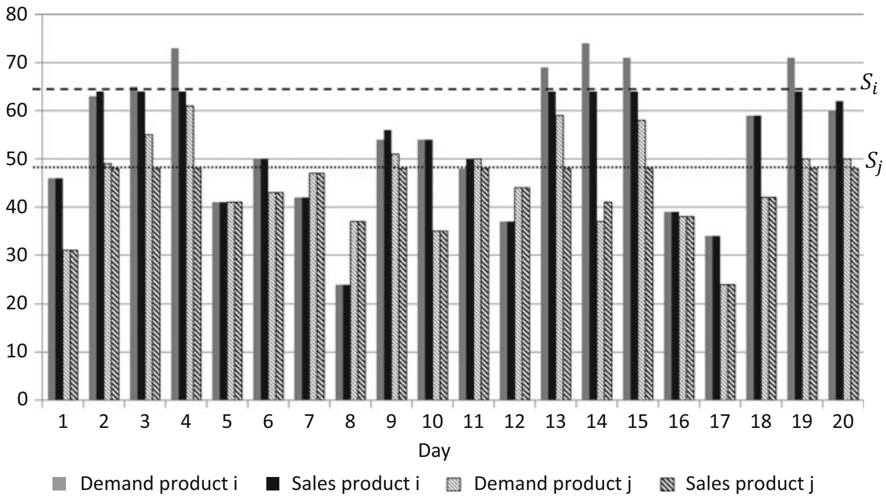


Fig. 5.1 Example of the data

There are two possible settings concerning the retailer's demand information: In one setting, the retailer observes only sales. He observes daily sales $D^c_{i,t}$ and hourly sales $h^i_{z,t}$. If a product is out-of-stock, demand is unknown and sales consist of demand censored at order-up-to level S_i. In the other setting, the retailer additionally has full demand information $D_{i,t}$ (in hindsight). If a product is out-of-stock, he knows the amount of lost sales. Figure 5.1 shows an example of the data for 20 days.

Furthermore, the retailer is aware of a causal relationship between price and demand and has collected data on the selling price $p_{i,t}$ as external variable.

5.3.2 Decisions

The retailer's objective is to maximize the profit of a given assortment by determining the optimal order quantities of two products that may be mutual substitutes. To determine the optimal order quantity, he takes the selling prices $p_{i,t}$ and unit purchasing costs $c_{i,t}$ into account. The selling price also explains a part of the demand variability. Therefore, he determines the optimal order quantity as a linear inventory function $B_{i,t}$ that depends on the price and reflects the price-demand relationship. An individual linear inventory function is determined per product i and substitute j. A linear relationship is postulated in order to enable solving the problem with Linear Programming. We obtain the following inventory function for product i:

$$B_{i,t} = a_i + b_{ii} p_{i,t} + b_{ji} p_{j,t} \qquad \forall t \in T; i = 1,2; j = 1,2; i \neq j. \tag{5.1}$$

The first part of the function determines the order quantity to satisfy primary demand and therefore depends on the selling price of the first-choice product. The second part of the function determines the order quantity to satisfy demand that is shifted from the out-of-stock product to its substitute and therefore depends on the selling price of the out-of-stock product. The coefficient a_i represents the y-intercept of the function. The influence of the external variables on the optimal inventory level is determined by the coefficients b_{ii} and b_{ij} which are set individually for each product based on the past sales observations. Setting the coefficients from past sales data requires that causal relationships from the past also hold for the future. b_{ii} is the coefficient that reflects the price-dependency of the demand and is negative for most products (exceptions due to snob effects are possible and would not affect the model). b_{ji} (and b_{ij}, respectively) is a coefficient for the substituted demand which is shifted from product j to i (and i to j, respectively).

We do not consider reference-price effects. The coefficients of the inventory function are decision variables. Further decision variables are the resulting sales from primary demand ($s_{i,t}$) and from stockout-based substitution ($s_{ji,t}$). These are determined according to the constraints that will be outlined in the following sections.

5.3.3 Objective Function

The retailer aims to optimize profits for both products as a centralized decision maker. The profit function (5.2) consists of revenues generated by sales of product i ($s_{i,t}$) and j ($s_{j,t}$) in period t. Additionally, demand that could not be satisfied by product j is shifted to product i with a certain substitution rate. This substituted demand is denoted as $s_{ji,t}$ and, in the opposite case, if demand is shifted from product i to j as $s_{ij,t}$.

$$\Pi = \sum_{t=1}^{T} \left(p_{i,t} \left(s_{i,t} + s_{ji,t} \right) - c_{i,t} B_{i,t} + p_{j,t} \left(s_{j,t} + s_{ij,t} \right) - c_{j,t} B_{j,t} \right) \tag{5.2}$$

5.3.4 Known Demand with Stockout Observations of One Product

The simplest case of the problem can be described by a setting where one product is always fully available and demand is known in hindsight. If one product is always available, the retailer can observe full substitution effects from the other product that may be out-of-stock.

Retailers often hold large inventories of a cheap, standard product and lower inventories resulting in potential stockouts of a more expensive, less frequently demanded product. A common example are conventional versus organic perishable products such as fruits and vegetables. Customers with a first-choice preference for organic products are less price-sensitive and willing to pay a premium. Therefore, the retailer tries to capture as much as possible of the premium the customers are willing to pay for organic products, but at the same time wants avoid costly excess inventories. If the organic product is sold out, customers often buy the conventional equivalent instead. In contrast, customers purchasing conventional products are more price-sensitive and do not buy the more expensive organic product if their first choice is out-of-stock.

Usually, the retailer only observes sales; demand is rarely known. However, one example of this rare case with demand observations is a supermarket that delivers online orders of groceries. The webshop does not show whether a product is in-stock, so customers place an order from a fixed assortment assuming that all products are available. Additionally, customers can note potential substitutes in a comments section in case that their primary choice is out-of-stock. We will use this assumption for illustrational purposes and relax it in the next section.

In the past, the retailer had large amounts of inventory of product i so that no out-of-stock situations were observed for product i. If product j stocked out, substituted demand was satisfied by product i. Fitting an inventory function to the historical observations, results in sales quantities as decision variables. The retailer's objective function (5.2) is subject to several constraints that will be outlined in the following.

The retailer can only sell as many units as he has available which is determined by the inventory function. Therefore, total sales of product i consisting of primary sales ($s_{i,t}$) and secondary sales due to substitution from j to i ($s_{ji,t}$) may not exceed the available inventory $B_{i,t}$ (5.3).

$$s_{i,t} + s_{ji,t} \leq B_{i,t} \qquad \forall t \in T \tag{5.3}$$

Since there was no substitution behavior from product i to j in the past, the retailer cannot learn about the customers' potential willingness to substitute ($s_{ij,t} = 0$). Consequently, primary sales of product j are limited by the inventory function without accounting for potential substitution according to:

$$s_{j,t} \leq B_{j,t} \qquad \forall t \in T \tag{5.4}$$

Further, the retailer can only sell as many units as the customer demands. Customer demand $D_{j,t}$ for product j is either satisfied from the first-choice product resulting in sales of product j or from the secondary choice $s_{ji,t}$ due to substitution (5.5). For product i only primary demand occurs (5.6).

$$s_{j,t} + s_{ji,t} \leq D_{j,t} \qquad \forall t \in T \tag{5.5}$$

$$s_{i,t} \leq D_{i,t} \qquad \forall t \in T \tag{5.6}$$

If substitution occurs, sales observations ($D^c_{i,t}$) increase by the amount of demand shifted from j to i. Therefore, the difference between sales observations and primary demand for i reflects the willingness of customers to substitute and is an upper limit (5.7).

$$s_{ji,t} \leq D^c_{i,t} - D_{i,t} \qquad \forall t \in T \tag{5.7}$$

Note that it is not possible to learn about substitution from historical observations if no substitution occurred. If the order quantities set by the *LP* model are lower than the historical inventory level with full availability and the product stocks out so that substitution could occur, the potential substitution effects are still not considered by the model due to the lack of information on the customers' willingness to substitute.

The following additional constraints for all decision variables apply:

$$s_{i,t}, s_{j,t}, s_{ij,t}, s_{ji,t}, B_{i,t}, B_{j,t} \geq 0 \tag{5.8}$$

$$a_i, b_{ii}, b_{ji}, a_j, b_{jj}, b_{ij} \in \Re. \tag{5.9}$$

5.3.5 Censored Demand

We will now turn to the more realistic case where the retailer only observes sales, i.e. censored demand observations if a product stocks out. This is common in most retail environments. In addition to recording total sales quantities, *POS* scanner systems record the time of the last sale and hourly sales quantities. We will use this information to estimate the total demand if one product stocks out.

Demand Estimation for Out-of-Stock Situations For products and days with censored demand observations C_i, i.e. stockouts, we need to calculate unobservable lost sales first. We estimate unobservable lost sales and demand substitution behavior without any distributional assumptions based on Lau and Lau (1996) and Sachs and Minner (2014). As the amount sold until a product stocks out is an unbiased estimator of the demand, we can use this information to estimate unobservable lost sales of the out-of-stock product.

Furthermore, sales of potential substitutes are inflated. As we can also observe original demand of this product up to the stockout of the other product, we can apply this information to determine an estimate of the additional demand and deflate sales of the substitute.

To determine the amount of original demand that occurred before product i stocked out (Lau and Lau 1996; Sachs and Minner 2014), we consider the hourly sales observations $h^i_{z,t}$. We denote the hourly time intervals by z. The store closes at Z. We consider that the ratio of cumulative demands until the end of the day $H^i_{Z,t}$

and the sum of all demand occurring during the day $D_{i,t}$ is one. This holds for the mean over all observations T, denoted by $\bar{H}^{iZ}$ and $\bar{D}^i$:

$$H^i_{Z,t} = \sum_{z=1}^{Z} h^i_{z,t} \tag{5.10}$$

$$\frac{\bar{H}^i_Z}{\bar{D}^i} = 1. \tag{5.11}$$

Our aim is to obtain a multiplier K^i_z that can be used to inflate unobservable lost sales and deflate sales with substitution. Recursively, we can calculate the ratios of mean cumulative sales and demand for all days with full availability based on mean hourly sales $\bar{h}^i_z$ by

$$\frac{\bar{H}^i_{Z-1}}{\bar{D}^i} = 1 - \frac{\bar{h}^i_Z}{\bar{D}^i} \tag{5.12}$$

$$\frac{\bar{H}^i_z}{\bar{D}^i} = \frac{\bar{H}^i_{z+1}}{\bar{D}^i} - \frac{\bar{H}^i_{z+1}}{\bar{D}^i}\frac{\bar{h}^i_{z+1}}{\bar{H}^i_{z+1}} = \frac{\bar{H}^i_{z+1}}{\bar{D}^i}\left(1 - \frac{\bar{h}^i_{z+1}}{\bar{H}^i_{z+1}}\right) = \frac{1}{K^i_z} \tag{5.13}$$

so that

$$K^i_z = \frac{K^i_{z+1}}{1 - \frac{\bar{h}^i_{z+1}}{\bar{H}^i_{z+1}}}. \tag{5.14}$$

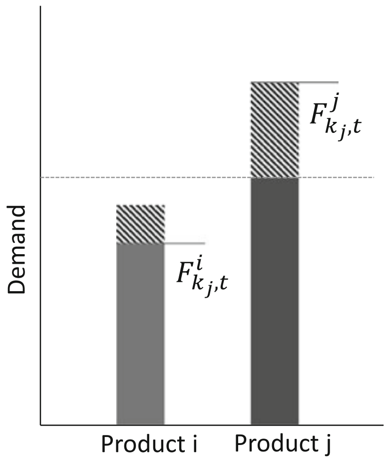
In order to estimate demand, we replace the hourly time intervals z with the stockout time $k_{i,t}$. If product i stocks out at time $k_{i,t}$ on day t, we have to correct for unobservable lost sales which is accomplished by

$$F^i_{k_i,t} = H^i_{k_{i,t}} \frac{K^i_{k_{i,t}} + K^i_{k_{i-1,t}}}{2}. \tag{5.15}$$

Replacing $k_{i,t}$ by $k_{j,t}$ in (5.15) allows to calculate the deflated demand of the substitute i if product j is out-of-stock.

Figure 5.2 illustrates demand estimation for the two-product case where product j is sold out. Demand for product j is filled until the order-up-to level (horizontal line) is reached. The time of the stockout is recorded as $k_{j,t}$. Any demand occurring after $k_{j,t}$ cannot be satisfied from product j's inventory and is unobserved. The shaded area of product j is an estimate for the unobserved demand and depends on $k_{j,t}$. The total demand for product j is estimated as $F^j_{k_j,t}$. Cumulative sales of product i before j stocks out corresponds to primary demand for i. After $k_{j,t}$, demand for i is inflated by substitution (shaded area). A fraction of the unsatisfied demand for product j is shifted to product i. $F^i_{k_j,t}$ is an estimate for the primary demand of product i deducting substitution.

Fig. 5.2 Demand estimation



Stockout Observations of One Product Constraints (5.3) and (5.4) are not affected by censoring and remain unchanged. Constraints (5.5) and (5.6) have to be adapted since demand cannot be observed. They are therefore replaced by an estimate of demand [see (5.15)] in constraints (5.16) and (5.17).

$$s_{i,t} \leq F^i_{k_j,t} \qquad \forall t \in T \tag{5.16}$$

$$s_{j,t} + s_{ji,t} \leq F^j_{k_j,t} \qquad \forall t \in T \tag{5.17}$$

Additionally, observations have to be separated into censored C_j and full F_j demand observations. Product i is always fully available, i.e. F_i. The amount of demand to be substituted may not exceed the past substitution quantity. Since the past substitution quantity cannot be directly observed, it is estimated from the total sales of product i and the estimate of primary demand of i $F^i_{k_j,t}$ in (5.18) . The total sales $D^c_{i,t}$ are inflated by the substitution from j to i so that by taking the difference we obtain an estimate of the substitution amount.

$$s_{ji,t} \leq D^c_{i,t} - F^i_{k_j,t} \qquad \forall t \in F_i \cap C_j \tag{5.18}$$

Since there are no out-of-stock observations for product i, no observations on the customers' willingness to substitute from i to j can be observed ($s_{ij,t} = 0$). This holds also for product j on all days with full availability of both products ($F_i \cap F_j$).

$$s_{ji,t} = 0 \qquad \forall t \in F_i \cap F_j \tag{5.19}$$

Further, the following constraints for all decision variables apply:

$$s_{i,t}, s_{j,t}, s_{ij,t}, s_{ji,t}, F^i_{k_t}, F^j_{k_t}, B_{i,t}, B_{j,t} \geq 0 \tag{5.20}$$

$$a_i, b_{ii}, b_{ji}, a_j, b_{jj}, b_{ij} \in \Re. \tag{5.21}$$

Stockout Observations of Both Products We conclude this chapter by outlining the model for censored demand and stockouts of both products. Constraint (5.4) has to be adapted to account for the substituted demand from i to j:

$$s_{j,t} + s_{ij,t} \leq B_{j,t} \qquad \forall t \in T \tag{5.22}$$

If both products stock out, the sequence of stockouts is important. Only sales observations before the first stockout occurs reflect primary demand. Therefore, the time of the first stockout k_t is obtained as the minimum of the stockout time of product i and j:

$$k_t = min(k_{i,t}; k_{j,t}). \tag{5.23}$$

In addition to constraints (5.3) and (5.16) to (5.19), the cases when product i or product i and j stock out have to be covered. If product i stocks out and j is fully available, constraints (5.24) ensures that substitution does not exceed its estimate. In this case, no substitution from j to i can be observed (5.25).

$$s_{ij,t} \leq D^c_{j,t} - F^j_{k_i,t} \qquad \forall t \in C_i \cap F_j \tag{5.24}$$

$$s_{ji,t} = 0 \qquad \forall t \in C_i \cap F_j \tag{5.25}$$

If both products stock out ($t \in C_i \cap C_j$), the demand and substitution estimates are calculated based on the first stockout time k_t in constraints (5.26) and (5.27).

$$s_{ji,t} \leq D^c_{i,t} - F^i_{k_t} \qquad \forall t \in C_i \cap C_j \tag{5.26}$$

$$s_{ij,t} \leq D^c_{j,t} - F^j_{k_t} \qquad \forall t \in C_i \cap C_j \tag{5.27}$$

Note that information on substitution behavior is limited by the historical inventory level and lowering $B_{i,t}$ does not result in larger substitution quantities.

The complete *LP* model for all possible cases with censored demand can be summarized as follows:

$$\Pi = \sum_{t=1}^{T} \Big(p_{i,t} \left(s_{i,t} + s_{ji,t} \right) - c_{i,t} B_{i,t} + p_{j,t} \left(s_{j,t} + s_{ij,t} \right) - c_{j,t} B_{j,t} \Big) \tag{5.28}$$

s.t.

One product is out-of-stock:

$$s_{ji,t} \leq D^c_{i,t} - F^i_{k_j,t} \qquad \forall t \in F_i \cap C_j; i = 1,2; j = 1,2; i \neq j \tag{5.29}$$

$$s_{ij,t} = 0 \qquad \forall t \in F_i \cap C_j; i = 1,2; j = 1,2; i \neq j \tag{5.30}$$

No stockouts:

$$s_{ij,t} = 0 \qquad \forall t \in F_i \cap F_j; i = 1,2; j = 1,2; i \neq j \tag{5.31}$$

Both products are out-of-stock:

$$s_{ji,t} \leq D^c_{i,t} - F^i_{k_t} \qquad \forall t \in C_i \cap C_j; i = 1,2; j = 1,2; i \neq j \tag{5.32}$$

All observations:

$$s_{i,t} + s_{ji,t} \leq B_{i,t} \qquad \forall t \in T; i = 1,2; j = 1,2; i \neq j \tag{5.33}$$

$$s_{i,t} + s_{ij,t} \leq F^i_{k_j,t} \qquad \forall t \in T; i = 1,2; j = 1,2; i \neq j \tag{5.34}$$

$$s_{i,t}, s_{j,t}, s_{ij,t}, s_{ji,t}, F^i_{k_t}, F^j_{k_t}, B_{i,t}, B_{j,t} \geq 0 \tag{5.35}$$

$$a_i, b_{ii}, b_{ji}, a_j, b_{jj}, b_{ij} \in \Re \tag{5.36}$$

5.4 Numerical Study and Empirical Analysis

To analyze the performance of our model, we compare it to already existing models in a numerical study. Since models existing in the literature often make distributional assumptions, we compare our model with data generated from a Poisson process. Firstly, we compare it to the optimal solution with known parameters (Parlar and Goyal 1984). Secondly, we estimate the parameters according to Anupindi et al. (1998). Since the approach by Anupindi et al. (1998) only estimates the arrival rates and substitution probabilities without considering optimal order decisions, we determine the order quantities according to the optimality conditions in Parlar and Goyal (1984). Note that we use a simplification that is frequently used in the literature (Netessine and Rudi 2003; Nagarajan and Rajagopalan 2008): Parlar and Goyal (1984) determine the optimal order quantities assuming substitution rates instead of substitution probabilities.

5.4.1 Benchmark to Estimate Arrival Rates and Substitution Probabilities

The approach of Anupindi et al. (1998) estimates arrival rates and substitution probabilities of a Poisson process according to the following procedure.

Given that two products $i = 1, 2$ are replenished at the beginning of a period up to starting stock levels $S_{i,t}$, customers arrive according to a Poisson process and purchase the products according to their preferences which is reflected in purchase probabilities u_i. If product j stocks out, time of the stockout ($k_{j,t}$) and sales of the substitute product at $k_{j,t}$ are recorded. The purchase probability if product j stocks out is denoted as u_{ji}. Five cases can be distinguished:

Case 1 ($F_i \cap F_j$): no stockout
Case 2 ($C_i \cap F_j$): product i stocks out, j is available
Case 3 ($C_j \cap F_i$): product j stocks out, i is available
Case 4 ($C_i \cap C_j$): first product i stocks out, then product j
Case 5 ($C_j \cap C_i$): first product j stocks out, then product i

If a product stocks out, demand for the substitute increases and therefore, purchase probabilities with substitution are greater, too. Nevertheless, some demand is lost since not all customers are equally willing to substitute. Consequently, we obtain:

$$u_i \leq u_{ji} \leq u_i + u_j \tag{5.37}$$

$$u_j \leq u_{ij} \leq u_i + u_j \tag{5.38}$$

Customers arriving at each product can be calculated from the overall customer arrival rate λ as follows:

$$\lambda_i = u_i \lambda \qquad \forall i = 1, 2 \tag{5.39}$$

$$\lambda_{ji} = u_{ji} \lambda \qquad \forall i = 1, 2; j = 1, 2; i \neq j \tag{5.40}$$

As long as both products are available, demand for products i and j follows independent Poisson processes ($Pois$). In the presence of stockouts, the purchase process for the substitute is independent from pre-stockout purchases. Hence, inter-arrival times satisfy the i.i.d. (independent and identically distributed) criterium. Therefore, the likelihood function can be described as a product of the purchase processes for the five cases and transformed into a log-likelihood function:

$$\begin{aligned} L = & \prod_{t \in F_i \cap F_j} Pois_1(S_i, S_j) \prod_{t \in C_i \cap F_j} Pois_2(S_i, S_j) \prod_{t \in C_j \cap F_i} Pois_3(S_i, S_j) \cdot \\ & \prod_{t \in C_i \cap C_j} Pois_4(S_i, S_j) \prod_{t \in C_j \cap C_i} Pois_5(S_i, S_j) \end{aligned} \tag{5.41}$$

$$\begin{aligned} logL = & H_i log\lambda_i - W\lambda_i + H_j log\lambda_j - W\lambda_j + H_i(j) log\lambda_{ji} - W_i \lambda_{ji} \\ & + H_j(i) log\lambda_{ij} - W_j \lambda_{ij} \end{aligned} \tag{5.42}$$

Subscripts of Poisson processes $Pois$ refer to the five cases with numbers 1 to 5. W refers to the time when either both products are available (W) or W_i when only i is available and W_j analogously. H_i and H_j describe the cumulative sales of each product over all periods if they are both in stock. Cumulative sales of substitutes in the presence of stockouts are denoted as $H_i(j)$, which are the sales of product i if product j is out-of-stock. By maximizing $logL$ subject to the constraint (5.44), we obtain the following maximization problem:

$$\max logL \tag{5.43}$$

s.t.

$$\lambda_i \leq \lambda_{ji} \leq \lambda_i + \lambda_j \qquad \forall i = 1,2; j = 1,2; i \neq j \tag{5.44}$$

The solution results in the following estimator equations:

$$\hat{\lambda}_i = \frac{H_i}{W} \qquad \forall i = 1,2 \tag{5.45}$$

$$\hat{\lambda}_{ji} = \frac{H_i(j)}{W_i} \qquad \forall i = 1,2; j = 1,2; i \neq j \tag{5.46}$$

and substitution probabilities:

$$\hat{u}_{ji} = \frac{\lambda_{ji} - \lambda_i}{\lambda_j} \qquad \forall i = 1,2; j = 1,2; i \neq j. \tag{5.47}$$

5.4.2 Optimal Solution

Parlar and Goyal (1984) formulate the two-product inventory problem with demand substitution. The objective function (5.48) consists of the revenues generated due to primary demand (first and last line) and substitution (two middle lines). $q_{i,t}$ denotes the optimal order quantity of the two products $i = 1,2$. Note that $q_{i,t}$ corresponds to $B_{i,t}$ of the LP approach, but $q_{i,t}$ does not depend on external variables. The substitution rate from product i to j, i.e., when product i stocks out is a, respectively b from product j to i.

$$\Pi_t = p_{i,t} \sum_{x=0}^{q_{i,t}-1} x f(x) + p_{i,t} q_{i,t} \sum_{x=q_{i,t}}^{\infty} f(x) - c_{i,t} q_{i,t}$$

$$+ p_{i,t} \sum_{x=0}^{q_{i,t}-1} \sum_{y=q_{j,t}}^{B_t} b(y - q_{j,t}) g(y) f(x)$$

$$+ p_{i,t} \sum_{x=0}^{q_{i,t}-1} \sum_{y=B_t+1}^{\infty} (q_{i,t} - x) g(y) f(x)$$

$$+ p_{j,t} \sum_{y=0}^{q_{j,t}-1} \sum_{x=q_{i,t}}^{A_t} a(x - q_{i,t}) f(x) g(y)$$

$$+ p_{j,t} \sum_{y=0}^{q_{j,t}-1} \sum_{x=A_t+1}^{\infty} (q_{j,t} - y) f(x) g(y)$$

$$+ p_{j,t} \sum_{y=0}^{q_{j,t}-1} y g(y) + p_{j,t} q_{j,t} \sum_{y=q_{j,t}}^{\infty} g(y) - c_{j,t} q_{j,t} \tag{5.48}$$

with A_t and B_t:

$$A_t = \left\lfloor \frac{q_{j,t} - y}{a} + q_{i,t} \right\rfloor \tag{5.49}$$

$$B_t = \left\lfloor \frac{q_{i,t} - x}{b} + q_{j,t} \right\rfloor . \tag{5.50}$$

Parlar and Goyal (1984) show that the function is concave if the following relationship between selling prices and substitution rates holds:

$$b p_{i,t} \leq p_{j,t} \leq p_{i,t}/a.$$

We assume that customers arrive according to a Poisson process with arrival rates λ_i and λ_j:

$$f(x) = \frac{\lambda_i^x}{x!} e^{-\lambda_i} \tag{5.51}$$

$$g(y) = \frac{\lambda_j^y}{y!} e^{-\lambda_j} \tag{5.52}$$

Since we study a discrete problem, we determine optimal order quantities $q_{i,t}$ and $q_{j,t}$ such that (5.53)–(5.56) hold.

$$\Pi_t \left(q_{i,t}, q_{j,t}\right) \geq \Pi_t \left(q_{i,t} - 1, q_{j,t}\right) \tag{5.53}$$

$$\Pi_t \left(q_{i,t}, q_{j,t}\right) \geq \Pi_t \left(q_{i,t} + 1, q_{j,t}\right) \tag{5.54}$$

$$\Pi_t \left(q_{i,t}, q_{j,t}\right) \geq \Pi_t \left(q_{i,t}, q_{j,t} - 1\right) \tag{5.55}$$

$$\Pi_t \left(q_{i,t}, q_{j,t}\right) \geq \Pi_t \left(q_{i,t}, q_{j,t} + 1\right). \tag{5.56}$$

5.4.3 Data Generation

We generate demand from a Poisson process with price-dependent hourly customer arrival rates. The arrival rate λ_i is a linear function of price $p_{i,t}$ with $p_{1,t} \sim [0; 1]$ and $p_{2,t} = gp_{1,t}$. g varies with the setting. The hourly customer arrival rate is generated as follows:

$$\lambda_i = a_i - b_i p_{i,t} \qquad \forall i = 1, 2 \tag{5.57}$$

For each instance, we randomly draw the parameters from the uniform distribution according to:

$$a_1 \sim [7; 12],\ b_1 \sim [2; 5],\ a_2 \sim [5; 10] \text{ and } b_2 \sim [1; 3].$$

The daily arrival rate can be obtained from the sum of the hourly arrival rates since the sum of independent Poisson random variables is also Poisson distributed (Ross 2010). We assume that demand follows a homogeneous Poisson process. We artificially censor demand by assuming different order-up-to levels. In the first setting, we assume that the order-up-to level of product 1 is always high enough to cover demand, i.e. no stockouts. We set the order-up-to level of product 1 so that it can fill all the demand for product 1 and substituted demand from product 2 in hindsight. In the second setting, we set order-up-to levels at the 75th percentile of demand for product 1. The order-up-to levels of product 2 are at the median in both settings.

Any demand exceeding the order-up-to level is shifted to the substitute product with a certain rate. The substitution rate if product 1 is out-of-stock is $a = 0.8$ and if product 2 is out-of-stock $b = 0.4$.

We fit the parameters of the different models to a sample of 100 observations. Given the parameters, we compare the resulting profits of an out-of-sample size of 100 observations. The experiment was repeated for 100 randomly generated instances.

5.5 Results

5.5.1 Known Demand with Stockout Observations of One Product

We first analyze the simple setting from Sect. 5.3.4 where demand is known and product 1 is always available. We compare the results for the *data-driven approach* with the optimal solution (Parlar and Goyal 1984) given the customer arrival rates of the Poisson process and substitution rates for both products. We denote the optimal solution as *parametric approach*. The price difference between both products is fixed such that $p_{2,t} = gp_{1,t}$ with $g = 1.4$.

Table 5.1 Known demand with stockout observations of one product

	Mean (Standard deviation)	
	Data-driven approach	Known values
In-stock probability product 1	0.9141 (0.0441)	0.8834 (0.0342)
In-stock probability product 2	0.6247 (0.0687)	0.5476 (0.0501)
Total leftover inventory	18.4 (2.8)	16.0 (2.7)
Total profit	59.62 (11.4)	59.65 (11.5)

Table 5.1 shows the mean results in terms of the total profit, leftover inventory and in-stock probabilities of both products for the out-of-sample observations. Standard deviations are in parentheses. The mean total profits of the *data-driven approach* are slightly lower than those of the *parametric approach*. The *parametric approach* leads to slightly more demand substitution with 3.5 % of the total demand of product 2 being satisfied by product 1 compared to 3.3 % for the *data-driven approach*. According to t-tests, the differences between profits are not significant for 82 % of all instances ($p > 0.1$). The *parametric approach* makes use of the distributional assumptions and takes the customer arrival rates of each day as given. Therefore, this approach serves as a benchmark for the optimal solution in the following. The standard deviation is slightly lower and the mean in-stock probabilities of both products are higher for the *data-driven approach*. The higher in-stock probabilities result in higher revenues which are compensated by the costs of larger leftover inventories.

5.5.2 *Censored Demand with Stockout Observations of One Product*

We now drop the assumption of observing demand and study the more realistic problem with sales, i.e. censored demand observations. All other parameters from Sect. 5.5.1 remain the same.

In addition to the *parametric approach with known values* from the previous section, we estimate the parameters for the optimal solution according to the procedure by Anupindi et al. (1998). We denote this approach as *Parametric approach with estimated values*. The results of the *Parametric approach with known values* are repeated from Sect. 5.5.1 as benchmark.

Table 5.2 Censored demand with stockout observations of one product

	Mean (Standard deviation)		
	Data-driven approach	Parametric approach with	
		Known values	Estimated values
In-stock probability product 1	0.8064 (0.0690)	0.8834 (0.0342)	0.771 (0.0736)
In-stock probability product 2	0.4819 (0.0696)	0.5476 (0.0501)	0.3618 (0.0675)
Total leftover inventory	12.2 (2.4)	16.0 (2.7)	18.6 (3.5)
Total profit	59.59 (11.4)	59.65 (11.5)	58.7 (11.7)

From Table 5.2, we observe that the *data-driven approach* achieves slightly higher profits than the *parametric one with estimated values*, but achieves lower mean total profits and in-stock probabilities than with the known demand values. The slightly lower profits of the *parametric approach with estimated values* are due to lower in-stock probabilities resulting in fewer revenues and at the same time higher leftover inventories. Furthermore, more substitution from product 2 to 1 takes place with the *parametric one with estimated values* (8.3 % of the total demand for product 2) than with the *data-driven approach* (4.3 %). Conducting t-tests on the instances shows that the differences between the profits of the *data-driven approach* and the *parametric approach with estimated values* are significant ($p < 0.1$) for 93 % of the instances.

5.5.3 Censored Demand with Stockout Observations of Both Products

If product 1 is artificially censored at an order-up-to level of 75 % for the in-sample observations, substitution behavior in both directions can be observed, but there is also a loss of information for the demand that could not be satisfied if both products stock out at the same time. This is reflected in lower mean profits (see Table 5.3) and on average lower in-stock probabilities for the *data-driven approach*. This means that even though more learning on the substitution behavior by observing substituted demand from product 1 to 2 is enabled, the loss of information due to more censoring has a stronger effect on profits. The service level differentiation between product 1 and 2 is reduced for both approaches.

So far, we assumed that the price of product 2 is always higher than the one of product 1. We will now look at varying price differences so that $p_{2,t} = gp_{1,t}$ with g being uniformly distributed between 0.8 and 1.4. Furthermore, we investigate

Table 5.3 Censored demand with stockout observations of both products

	Mean (Standard deviation)		
	Data-driven approach	Parametric approach with	
		Known values	Estimated values
In-stock probability product 1	0.6952 (0.0679)	0.8834 (0.0342)	0.6801 (0.0797)
In-stock probability product 2	0.4741 (0.0663)	0.5476 (0.0501)	0.5411 (0.0829)
Total leftover inventory	9.3 (1.5)	16.0 (2.7)	15.9 (3.5)
Total profit	59.5 (11.4)	59.6 (11.5)	58.5 (11.7)

sample size effects. The results are shown in Table 5.4. The difference between the profits from the *parametric approach with known values* and from the *data-driven approach* increases and is significant ($p < 0.1$) for all instances. The *parametric approach with estimated values* achieves already similar results with a small sample of only 20 observations compared to the one with 100 observations, but slightly lower profits than the *data-driven approach*. However, from a customer's point of view, the *parametric approach with estimated values* should be preferred since the in-stock probabilities of the two products are higher. The difference increases with larger sample sizes where the *data-driven approach* improves profits. But overall, the increase of profits is small. The number of instances for which the difference is significant at the 10 % level increases as well from 58 to 78 % with larger sample sizes.

In the previous analyses we assumed that the in-sample observations of product 2 are highly censored. There is a trade-off associated with censoring: If demand is highly censored, less can be learnt about the true demand values, but at the same time, more can be learnt about the substitution behavior. We now vary the censoring level by adding percentiles from 75 to 95 %.

For both approaches, we observe that the effect of learning about substitution if censoring is high, results in slightly higher profits than if there is only few censoring and more can be learnt about demand. The effect is due to the fact that we are able to estimate unobservable lost sales if demand is censored, but we are not able to learn about substitution behavior if only few substitution occurs (Table 5.5).

5.5.4 Real Data

We now test the two models on a set of real data. We collect data from 66 stores of a retail chain for two sorts of lettuce that are mutual substitutes. The time horizon considered is from October 2008 to July 2012. We divide the dataset in a

Table 5.4 Sample size effects—censored demand with stockout observations of both products and varying price differences

Sample size		Mean (Standard deviation) Data-driven approach	Parametric approach with Known values	Parametric approach with Estimated values
20	In-stock probability product 1	0.6167 (0.1380)	0.8964 (0.0284)	0.7161 (0.0916)
	In-stock probability product 2	0.4958 (0.0613)	0.2917 (0.0468)	0.5357 (0.1338)
	Total leftover inventory	10.4 (3.1)	15.6 (2.6)	16.9 (4.0)
	Total profit	54.5 (9.9)	55.9 (10.0)	54.4 (10.0)
50	In-stock probability product 1	0.6706 (0.0947)	See above	0.7051 (0.0890)
	In-stock probability product 2	0.5477 (0.0554)		0.5363 (0.1205)
	Total leftover inventory	11.2 (2.2)		16.4 (4.0)
	Total profit	54.7 (9.9)		54.4 (10.0)
100	In-stock probability product 1	0.681 (0.0764)	See above	0.6996 (0.0761)
	In-stock probability product 2	0.4745 (0.0641)		0.541 (0.0887)
	Total leftover inventory	10.1 (1.9)		16.2 (3.5)
	Total profit	54.9 (9.9)		54.4 (10.0)

sample of 250 days per store on which we fit the parameters of the models. We then compare the performance on the remaining observations as an out-of-sample. Since the retailer only observes sales if a product is out-of-stock, we include only days with no stockouts in the out-of-sample so that we have full demand information. The size of the out-of-sample therefore varies with the frequency of stockouts and contains between 404 and 610 days (on average 522 days).

In addition to daily and hourly sales, we collect information on prices and daily temperatures. The price of product 1 (2) lies between 0.29 € (0.09 €) and 1.79 € (1.49 €). Since we have no information on costs, we assume that $c_{1,t} = 0.1p_{1,t}$ and $c_{2,t} = 0.25p_{2,t}$. The substitution rates are unknown. We assume that $a = 0.1$ and $b = 0.7$ according to expert opinions from the managers of the retail chain. We add information on weekdays. $d_{1,t}$ is a binary variable if it is a

Table 5.5 Censoring effects—censored demand with stockout observations of both products and varying price differences

Censoring level of product 2		Mean (Standard deviation) Data-driven approach	Mean (Standard deviation) Parametric approach with estimated values
75	In-stock probability product 1	0.6865 (0.0835)	0.7128 (0.0765)
	In-stock probability product 2	0.5343 (0.0632)	0.6018 (0.0879)
	Total leftover inventory	10.9 (2.2)	17.3 (3.5)
	Total profit	54.8 (9.9)	54.3 (10.0)
90	In-stock probability product 1	0.6861 (0.0810)	0.7062 (0.0852)
	In-stock probability product 2	0.5398 (0.0629)	0.6104 (0.1561)
	Total leftover inventory	11.0 (2.1)	17.9 (4.2)
	Total profit	54.8 (9.9)	54.2 (10.0)
95	In-stock probability product 1	0.6872 (0.0779)	0.6966 (0.0826)
	In-stock probability product 2	0.5406 (0.0589)	0.6334 (0.1466)
	Total leftover inventory	11.0 (2.0)	17.8 (4.0)
	Total profit	54.8 (9.9)	54.2 (10.0)

Tuesday or Wednesday when demand is typically lowest and $d_{2,t}$ for Fridays and Saturdays when peak demand occurs. Mondays and Thursdays are not included to avoid collinearity. The temperatures (w_t) during the time horizon considered range between $-11.1\,°C$ and $36.4\,°C$. We include price, weekday and temperature as external factors in the inventory functions (5.58 and 5.59) to determine the optimal order quantities according to the *data-driven approach*. The coefficient for Tuesdays/Wednesday is denoted $b_i^{d_1}$, for Fridays/Saturdays $b_i^{d_2}$ and temperature b_i^{w}.

$$B_{1,t} = a_1 + b_{11}p_{1,t} + b_{21}p_{2,t} + b_1^{d_1}d_{1,t} + b_1^{d_2}d_{2,t} + b_1^{w}w_t \qquad \forall t \in T \tag{5.58}$$

$$B_{2,t} = a_2 + b_{22}p_{2,t} + b_{12}p_{1,t} + b_2^{d_1}d_{1,t} + b_2^{d_2}d_{2,t} + b_2^{w}w_t \qquad \forall t \in T. \tag{5.59}$$

Table 5.6 Real data

	Mean (Standard deviation)	
	Data-driven approach	Estimated values
In-stock probability product 1	0.8241 (0.0620)	0.7065 (0.0761)
In-stock probability product 2	0.7004 (0.1042)	0.6158 (0.1192)
Total leftover inventory	45.9 (14.7)	34.3 (9.3)
Total profit	51.0 (11.4)	48.5 (10.5)

The results in Table 5.6 show that the *data-driven approach* also outperforms the *parametric approach with estimated values* for real data. According to t-tests, the differences between profits are significant ($p < 0.1$) for the majority of stores (91 %). One reason for the lower profit of the *parametric approach with estimated values* is that the approach assumes that the data is Poisson distributed. But the Chi-Square Goodness of Fit and the Kolmogorov-Smirnov Test reject that the data is Poisson distributed ($p = 0.0000$). In contrast, the *data-driven approach* is distribution-free. Another reason is that the *parametric approach with estimated values* does not take the external variables which are considered by the *data-driven approach*.

5.6 Conclusions

We suggest a novel approach based on data-driven optimization which integrates demand forecasting, substitution and inventory optimization. There are no prior distribution assumptions and it only requires *POS* scanner data as it is readily available in most stores. Observations obtained from store-level scanner data are directly incorporated in inventory optimization, thus taking substitution effects and unobservable lost sales into account. Additionally, we account for other variables affecting demand such as price, weekdays or weather. These factors also influence the willingness of customers to substitute which is reflected in our model. The model fits a linear function of external variables to the available data. We build our model based on data for perishable products from a large European retail chain.

The model is limited in that it is only able to account for substitution behavior that was observed in the past. It does not reflect additional substitution behavior if inventory levels set by the *LP* approach are lower than the historical ones. In addition, shortage penalty costs are not considered. Another limitation of the *LP* approach is the assumption that the relationship between the inventory level and the external variables is linear. In both cases, the model would otherwise become non-linear.

Chapter 6
Empirical Newsvendor Decisions Under a Service Level Contract

We analyze the order decisions of a manufacturer who supplies several retail stores with bakery products in a newsvendor context. According to a service level contract between the manufacturer and the retailer, the manufacturer has to achieve an in-stock service level measured over multiple products. We derive the optimal inventory policy for an aggregate service level contract and compare the optimal solution to the decisions made by the manufacturer. Further, we address the question whether the results of experimental studies from the laboratory can be transferred to real-world decisions. Our findings indicate that real decision makers show similar decision biases as students in laboratory environments.

6.1 Introduction

In recent studies, many researchers show that decision making in newsvendor settings does not follow profit-maximizing predictions if human decision makers are involved: In various lab experiments, students and managers face the newsvendor problem, but their behavior consistently differs from normative profit-maximizing textbook solutions. But how do decision makers facing newsvendor problems behave in real-world situations? Are they also biased in their decisions? Do they show similar deviations for expected profit-maximizing behavior? In this chapter, we will address this issue. We analyze empirical data from newsvendor-like decisions in order to compare the observed deviations in the lab with real-world decisions.

To analyze real-world newsvendor-type decisions, we investigate data from a company selling bakery items at several supermarkets. Bakery items are typical newsvendor products, since all leftovers are discarded at the end of the day. In contrast to most of the recent behavioral newsvendor studies, our decision maker is not facing a single-product cost-minimization problem, but sells various products.

A.-L. Sachs, *Retail Analytics*, Lecture Notes in Economics and Mathematical Systems 680, DOI 10.1007/978-3-319-13305-8_6

In addition to that, he is facing an aggregated service level constraint, i.e., he has to achieve an α service level over all products, measured as the fraction of products in-stock at the end of the day. This setting is also interesting from a normative perspective, as different product characteristics have to be taken into account. We analyze the setting and derive normative solutions for the profit-maximizing order quantities for different products.

In recent studies, subjects in laboratory environments made newsvendor decisions in different settings and framings and placed orders in newsvendor situations (e.g., Schweitzer and Cachon 2000; Bolton and Katok 2008; Becker-Peth et al. 2013; and many more). The studies show consistent ordering biases which cannot be explained by traditional or extended newsvendor theories. The orders are biased towards the mean demand, also referred to as the pull-to-center effect (see Bendoly 2006; Benzion et al. 2008; Bostian et al. 2008). This bias is hardly reduced over time; even long experiments with many decisions (up to 100) show no or only little learning (Bolton and Katok 2008) and also managers show similar behavior (Bolton et al. 2012). Many explanations have been ruled out by different researchers. There are five remaining possible explanations for the pull-to-center effect: random errors, anchoring and adjustment, demand chasing, and ex-post inventory error and psychological cost for overage and underage.

The settings analyzed so far are mainly focusing on the single product newsvendor. In our empirical setting, the decision maker is focusing on an aggregated service level contract. Therefore, many theories, such as risk aversion, are not applicable to our setting. In the remainder of this chapter, we will concentrate on those biases which can be transferred to the service level setting such as anchoring and demand chasing.

Our main contribution is threefold. Firstly, we derive the optimal inventory policy for a manufacturer who operates under an aggregated service level contract. Secondly, we compare the decision making of an actual manufacturer who operates under such a contract, and compare it with the optimal policy. We identify the elements of the optimal policy that the manufacturer has implemented and those that he has ignored. Thirdly, we use recent insights from behavioral operations management to explain the behavior of the actual manufacturer and to develop a behavioral model to describe actual decision making under an aggregated service level contract. We find evidence that some biases can also be observed in the service level settings, e.g., demand chasing (although we have auto-correlated demand in our setting), and others have a different effect in our setting, e.g., the anchoring heuristic, where the anchor is not mean demand any more, but the service level target.

The results of our research also have important implications for behavioral operations management research in general and for practitioners. To our best knowledge, we are the first to use actual data to analyze decision making in a newsvendor-type setting. The literature on newsvendor-type decision making has so far focused on laboratory experiments and the behavioral operations management community discusses extensively how well the laboratory experiments translate to practice. The results of our analyses show that the main effects that have been observed in the laboratory, such as the pull-to-center effect and demand chasing, can also be observed in practice. This implies that the concepts that have been developed by behavioral operations researchers using laboratory experiments, are likely to generalize to practice.

For practitioners, the results of our research provide the means to optimally manage inventory under aggregated service level constraints. Our results indicate that substantial profit improvements can be achieved by differentiating service levels by unit revenue, unit cost, and demand uncertainty. The manufacturer who motivated our research, uses elements of the optimal inventory policy and differentiates the service levels between make and buy products. This policy allows to achieve a profit that is 0.9 % higher than the profit he would achieve under a policy that does not differentiate service levels. However, if the full optimal policy were implemented, i.e., if service levels were differentiated based on the unit cost and unit revenue of individual products and based on demand uncertainty by product and location, then the manufacturer could improve profits by an additional 7.1 %.

The remainder of this chapter is structured as follows. In Sect. 6.2, we describe the setting we analyze. In Sect. 6.4, we present our model and derive the optimal solution. The optimal solution determines the service levels (strategic solution) and daily order quantities (operational solution) by product and location. In Sect. 6.5, we compare the strategic solution of the model with the solution the manufacturer has implemented. We show that the manufacturer differentiates service levels between make and buy products, but ignores unit revenue and unit cost differences within each of these segments. He also differentiates service levels based on demand uncertainties. We build a model based on this observation and show that the model captures actual behavior well. In Sect. 6.6, we compare the operational solution of the model with the solution the manufacturer has implemented. We show that the manufacturer over-reacts to changes in the demand, an effect that has been observed in many laboratory experiments on newsvendor-type decision making. The literature on behavioral operations management offers various explanations for this behavior. We analyze the most popular explanations and show that a demand-chasing model describes actual operational ordering better than the other models we analyze. In Sect. 6.8, we summarize our results and discuss the implications for theory and practice.

6.2 The Setting

The decision maker in our empirical setting is a manufacturer who is selling $N = 23$ bakery products at a retail chain on his own bill. Delivering single-day products every morning, he faces a newsvendor situation. The demand for each product is unknown, the unit revenue r is fixed and determined by the retailer. Each day, he determines the quantity to stock in each store s, supplying $n_s = 66$ stores of the retailer. Unmet demand is lost (no refill possible) and lost sales are unobservable, i.e., we have censored demand data.

The sequence of events is as follows: Late in the morning, the decision maker observes sales of the previous day and analyzes historical sales data to forecast demand for each product in each store for the next day. The decision maker then determines the production quantity for each product. During the day, items are produced and production is completed early in the morning on the next day. The finished items are then delivered to each store before the stores open. Upon delivery, the decision maker collects leftover items from the previous day and discards them. Given that the decision maker knows the production quantities of the previous day, he can draw conclusions on how many items were sold and which products stocked out from the amount of leftover inventory. The decision maker then returns to his production facility to determine the forecast and the production quantities for the next upcoming day.

The decision maker has a service level contract with the retailer for the bakery products and has to achieve a service level of 70 % (in-stock probability) across all products in each store. Since the decision maker has to achieve the service level target at each store, the stores are treated individually in the subsequent calculations.

6.2.1 Data Overview

Our dataset consists of almost two years (12/14/2010 to 12/07/2012), with in total about 600 selling days. Each day contains: number of delivered units, number of units sold per hour and time of last item sold for each store and product. The stores are open Monday through Saturday from 8 am to 8 pm. Stores are closed on public holidays, which affects sales on the day before and after a holiday. We therefore exclude these days from our dataset as well as the week after Christmas and New Year's Eve. We also had to exclude data from two stores due to technical problems. In total, we analyzed data from 64 stores. There were no promotions for bakery products.

The decision maker purchases some of the products from an external supplier (external products, $i = 17, .., 23$) but produces most of them himself ($i = 1, .., 16$). We name the latter ones internal products for the remainder of the chapter. In general, the contribution margin on internal products is higher than on external products, as external products include the supplier's fixed costs and profit. Therefore, we group the products into high and low profit categories.

The data are summarized in disguised form in Table 6.1, where c is the production/purchasing costs, r is the selling price for the manufacturer and CR the resulting critical ratio. The critical ratio is the profit-maximizing in-stock probability the decision would target, if there were no service level constraints. The table contains means and standard deviations of sales and demand, including weekday-specific values (Monday–Saturday). Since POS scanner data only records sales, we have to estimate demand. To allow for a fair comparison with the decision maker's performance, we use two different approaches requiring different levels

Table 6.1 Data summary (disguised)

	Product	c	r	CR	Mean sales (Standard deviation)							Mean demand (Standard deviation)						
					Avg.	Mo	Tu	We	Th	Fr	Sa	Avg.	Mo	Tu	We	Th	Fr	Sa
Internal products	1	0.49	0.71	0.31	21.4	22.1	18.2	18.5	22.3	21.8	25.6	22.1	22.8	18.7	19.1	23.0	22.5	26.9
					(10.5)	(10.2)	(8.9)	(9.3)	(10.6)	(10.0)	(12.0)	(11.4)	(10.9)	(9.4)	(10.0)	(11.3)	(10.6)	(13.5)
	2	0.33	0.70	0.52	10.1	12.6	8.9	8.6	9.7	9.3	11.5	10.5	13.0	9.2	8.8	10.1	9.7	12.1
					(5.3)	(6.7)	(4.9)	(4.4)	(4.6)	(4.4)	(5.6)	(5.7)	(7.0)	(5.1)	(4.6)	(4.9)	(4.7)	(6.2)
	3	0.41	0.96	0.58	10.8	11.8	9.0	9.2	10.9	10.8	13.1	11.2	12.3	9.3	9.5	11.3	11.2	13.9
					(4.8)	(5.1)	(3.8)	(4.2)	(4.5)	(4.3)	(5.5)	(5.3)	(5.7)	(4.1)	(4.5)	(4.9)	(4.8)	(6.3)
	4	0.26	0.79	0.67	8.5	10.2	7.5	7.3	8.1	7.8	10.1	8.9	10.7	7.7	7.6	8.5	8.2	10.7
					(4.5)	(5.3)	(3.9)	(4.0)	(3.7)	(3.9)	(5.0)	(4.8)	(5.6)	(4.1)	(4.2)	(4.0)	(4.2)	(5.6)
	5	0.30	1.00	0.70	5.7	6.8	4.9	4.8	5.7	5.7	6.3	6.0	7.1	5.1	5.0	6.0	6.0	6.7
					(3.5)	(4.2)	(3.1)	(3.0)	(3.3)	(3.2)	(3.8)	(3.8)	(4.5)	(3.2)	(3.1)	(3.5)	(3.5)	(4.3)
	6	0.55	1.18	0.53	5.5	6.3	4.8	4.8	5.7	5.5	6.2	5.8	6.6	4.9	5.0	6.0	5.7	6.6
					(3.1)	(3.5)	(2.7)	(2.7)	(3.1)	(3.0)	(3.4)	(3.4)	(3.8)	(2.9)	(2.9)	(3.3)	(3.2)	(3.8)
	7	0.55	1.19	0.54	12.7	15.0	10.8	10.9	12.4	12.4	14.6	13.1	15.5	11.1	11.3	12.8	12.7	15.1
					(6.6)	(7.2)	(5.8)	(6.0)	(6.1)	(6.2)	(6.8)	(6.9)	(7.5)	(6.0)	(6.4)	(6.4)	(6.4)	(7.2)
	8	0.31	0.68	0.55	5.4	4.6	4.1	4.6	5.1	5.8	7.9	5.7	4.9	4.3	4.9	5.4	6.1	8.5
					(3.7)	(2.8)	(2.5)	(3.9)	(3.1)	(3.5)	(4.7)	(4.1)	(3.1)	(2.6)	(4.2)	(3.3)	(3.7)	(5.6)
	9	0.64	0.72	0.11	11.1	11.3	9.4	9.7	12.2	11.6	12.4	11.5	11.7	9.7	10.0	12.7	12.0	13.0
					(5.5)	(5.5)	(4.7)	(5.0)	(5.6)	(5.2)	(6.0)	(5.9)	(5.9)	(5.0)	(5.3)	(6.0)	(5.7)	(6.8)
	10	0.37	0.42	0.12	17.0	19.6	15.4	15.2	18.9	16.5	16.6	17.6	20.3	15.8	15.7	19.6	17.1	17.5
					(8.1)	(8.9)	(7.5)	(7.5)	(8.3)	(7.6)	(7.7)	(8.7)	(9.6)	(7.9)	(7.9)	(8.9)	(8.3)	(8.7)
	11	0.60	0.68	0.11	6.7	7.9	5.8	5.8	7.4	6.5	6.8	7.0	8.2	6.0	6.2	7.7	6.8	7.1
					(4.0)	(4.5)	(3.5)	(3.6)	(4.1)	(3.6)	(4.0)	(4.3)	(4.8)	(3.7)	(4.0)	(4.6)	(3.9)	(4.3)
	12	0.59	0.75	0.21	4.8	3.6	3.5	3.8	5.2	6.2	6.2	5.0	3.8	3.6	4.0	5.4	6.4	6.5
					(3.1)	(2.0)	(2.0)	(2.6)	(3.2)	(3.4)	(3.3)	(3.3)	(2.2)	(2.1)	(2.7)	(3.4)	(3.8)	(3.7)
	13	0.59	0.81	0.27	7.9	5.3	5.0	5.8	8.2	10.5	12.5	8.3	5.6	5.2	6.1	8.5	10.9	13.2
					(5.1)	(2.9)	(2.6)	(4.0)	(4.4)	(4.9)	(5.6)	(5.5)	(3.3)	(2.8)	(4.2)	(4.7)	(5.4)	(6.5)
	14	0.26	0.63	0.58	9.7	9.5	8.0	8.2	10.2	9.9	12.5	10.2	10.0	8.3	8.5	10.6	10.4	13.2
					(4.8)	(4.3)	(3.7)	(3.9)	(4.6)	(4.6)	(5.8)	(5.2)	(4.8)	(3.9)	(4.2)	(5.0)	(5.1)	(6.5)
	15	0.45	0.51	0.12	11.0	9.7	8.6	9.4	12.5	12.5	13.2	11.5	10.2	9.0	9.8	13.1	13.0	14.0
					(5.1)	(4.2)	(3.8)	(4.4)	(5.2)	(5.4)	(5.5)	(5.7)	(4.6)	(4.1)	(4.7)	(5.8)	(5.8)	(6.5)
	16	1.00	1.35	0.26	6.9	5.4	4.8	5.2	7.0	8.9	9.8	7.2	5.6	5.0	5.4	7.3	9.2	10.3
					(3.9)	(2.7)	(2.4)	(3.3)	(3.5)	(4.1)	(4.2)	(4.3)	(2.9)	(2.6)	(3.4)	(3.7)	(4.5)	(4.8)
External products	17	1.00	1.52	0.34	9.5	6.4	6.1	7.0	8.6	12.0	16.8	9.9	6.7	6.3	7.3	9.0	12.4	17.7
					(7.8)	(4.5)	(4.0)	(5.7)	(6.9)	(7.9)	(9.7)	(8.3)	(4.8)	(4.1)	(5.8)	(7.0)	(8.2)	(10.9)
	18	0.69	1.16	0.40	17.6	19.8	14.8	15.0	16.5	17.0	22.8	18.4	20.6	15.4	15.5	17.2	17.7	24.1
					(11.0)	(13.4)	(11.1)	(8.7)	(8.2)	(9.9)	(11.5)	(12.1)	(16.1)	(11.4)	(9.4)	(8.7)	(10.4)	(12.9)
	19	0.60	1.04	0.42	15.0	17.3	12.7	12.7	15.0	14.8	17.9	15.6	17.9	13.1	13.1	15.6	15.4	18.8
					(7.8)	(8.6)	(6.7)	(6.7)	(6.8)	(7.0)	(9.1)	(8.3)	(9.1)	(7.0)	(7.0)	(7.2)	(7.5)	(10.0)
	20	0.75	1.19	0.37	38.7	38.2	32.6	34.3	39.6	38.4	49.3	40.1	39.5	33.6	35.3	41.0	39.5	51.6
					(20.4)	(19.9)	(16.7)	(18.4)	(19.0)	(18.2)	(25.0)	(21.7)	(21.0)	(17.6)	(19.2)	(19.9)	(19.2)	(27.0)
	21	0.73	0.82	0.11	23.2	22.3	19.1	19.8	25.8	26.0	26.7	24.1	23.1	19.7	20.5	26.8	26.8	27.9
					(10.7)	(9.7)	(9.0)	(9.6)	(10.8)	(10.9)	(11.6)	(11.6)	(10.5)	(9.6)	(10.2)	(11.7)	(11.7)	(12.7)
	22	0.90	1.09	0.18	11.0	10.4	8.6	8.3	11.6	12.5	14.7	11.4	10.8	8.8	8.6	12.1	12.9	15.2
					(9.3)	(10.4)	(7.5)	(5.7)	(9.4)	(9.5)	(10.8)	(9.7)	(10.6)	(7.6)	(6.0)	(10.1)	(9.9)	(11.3)
	23	0.89	1.08	0.18	9.1	8.9	7.4	7.2	9.9	10.3	10.9	9.4	9.1	7.6	7.4	10.2	10.6	11.3
					(6.7)	(7.2)	(5.5)	(4.4)	(7.2)	(7.3)	(7.2)	(6.9)	(7.3)	(5.6)	(4.5)	(7.5)	(7.5)	(7.5)

of information, i.e. daily versus hourly sales data. For our forecast and inventory optimization we use the approach by Bell (1981) and Wecker (1978) which makes distributional assumptions, but works with daily sales observations which the decision maker is able to observe. For the analysis of the performance of the different models and the decision maker, we use the approach by Lau and Lau (1996). This approach takes hourly sales into account which improves the forecast accuracy, but the decision maker is not able to observe. Furthermore, it does not require any assumptions on the demand distribution, thus avoiding any potential misspecification issues. Table 6.1 contains the demand estimated according to Lau and Lau (1996).

Figure 6.1 shows the achieved service level for all 23 products per store. Each boxplot spans over 64 stores per product.

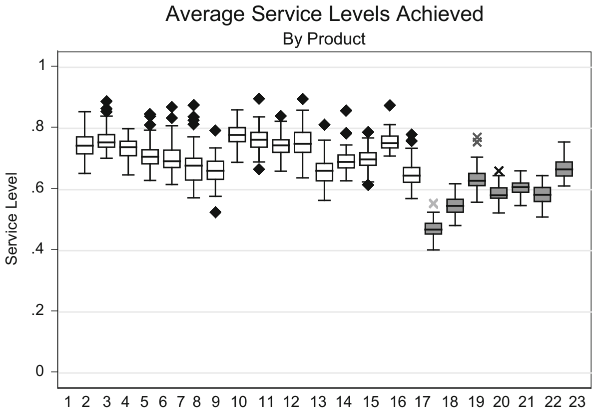


Fig. 6.1 Empirical service level per product over all stores

6.3 Modeling Demand

Analyzing the demand for our products, we find a strong seasonality within a week, but a constant pattern between weeks, which is typical for grocery retailers (van Donselaar et al. 2006, 2010). Figure 6.2 shows the average auto-correlation of demand ρ_l with lag l for our 23 products. We see the highest correlation for lag=6. As our week has six working days (Monday to Saturday), ρ_6 determines the correlation of demand for similar weekdays, e.g., the demand correlation between Monday and Monday of the previous week. Note that the decision maker is not able to take lag= 1 into consideration, as he observes the number of units sold at the beginning of the next day, so after the order decision has been made. Testing the best lag-order also results in the best Schwarz's Bayesian information criterion (BIC) for a lag= 6 in 97.63 % of all store and product combinations (lag= 0 with 1.65 %, lag= 1 with 0.46 %, lag= 2 with 0.07 %, lag= 3 with 0.07 %, lag= 4 with 0.13 % and lag= 5 with 0 %).

Modeling the demand for our setting, we have to account for two important effects: firstly the decision maker faces POS data (censored demand data); secondly we have auto-correlated demand.

In the case of POS data, we usually observe the problem of unobservable lost sales. In each period t, we have an order quantity $q_{i,t}$ stocked at the POS. If demand exceeds supply, we only observe sales $S_{i,t}$, which are limited by the actual number of units stocked $S_{i,t} = \min(y_{i,t}, q_{i,t})$. Taking this into account, we cannot use sales to forecast future demand. The retailer (where the items are actually sold) has more information, as he also knows the time of the last item sold.

To estimate the unobservable lost sales, we use the approach by Bell (1981) and Wecker (1978) who assume that demand per product i is normally distributed with density function $f(x)$. For situations where $S_{i,t} < q_{i,t}$, we can set $Y_{i,t} = S_{i,t}$ as we observe complete demand.

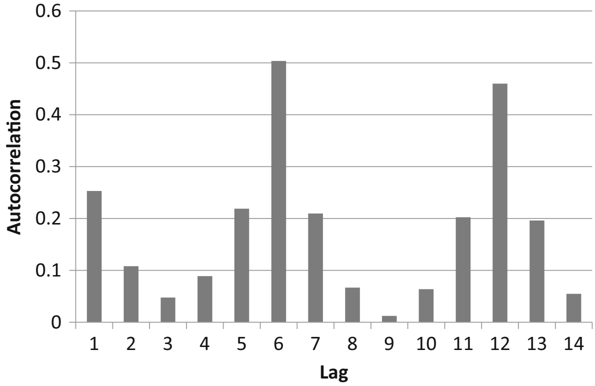


Fig. 6.2 Autocorrelation coefficients of demand

For $S_{i,t} = q_{i,t}$, we can only observe censored demand, but conclude that $y_{i,t} \geq q_{i,t}$. Furthermore, assumptions about the distribution of complete demand observations also hold for censored demands, but we have to estimate the truncated part of the distribution function. Therefore, we can estimate the expected demand conditional on the out-of-stock situation $E(Y_{i,t}|Y_{i,t} \geq q_{i,t})$ based on $\hat{\mu}_{i,t}$ and $\hat{\sigma}_{i,t}$ as

$$E(Y_{i,t}|Y_{i,t} \geq q_{i,t}) = \frac{\int_{q_{i,t}}^{\infty} xf(x|\hat{\mu}_{i,t}, \hat{\sigma}_{i,t})dx}{\int_{q_{i,t}}^{\infty} f(x|\hat{\mu}_{i,t}, \hat{\sigma}_{i,t})dx}. \tag{6.1}$$

We use $y_{i,t} = E(Y_{i,t}|Y_{i,t} \geq q_{i,t})$ as the demand estimate in out-of-stock situations for forecasting future demand.

Having (estimated) demand observations for all days now, we can forecast future demand. We denote demand as $Y_{i,t}$ for a specific product i, in period t, with mean $E(Y_{i,t}) = \mu_{i,t}$ and standard deviation $\sigma_{i,t}$. Taking the auto-correlation over time into account, we model demand as a time series using an *ARIMA*(0,1,1) model:

$$Y_{i,t} = Y_{i,t-1} + \epsilon_{i,t} - \theta\epsilon_{i,t-1}, \tag{6.2}$$

where $\epsilon_{i,t}$ is normally distributed white noise, with $E(\epsilon_{i,t}) = 0$ and $Var(\epsilon_{i,t}) = \sigma_i^2$.

Single exponential smoothing is the mean squared error (*MSE*) minimizing forecast method for *ARIMA*(0,1,1) (Chatfield 2001).Using the single exponential smoothing forecast with a smoothing factor η, we optimally update our mean demand forecast $\hat{\mu}$ with each new demand observation:

$$\hat{\mu}_{i,t} = \eta_i Y_{i,t-1} + (1 - \eta_i)\hat{\mu}_{i,t-1} \tag{6.3}$$

We determine the store- and product-specific smoothing parameter η_i such that the sum-of-squared forecast errors is minimized. According to our findings on demand autocorrelations, and as we are looking for a parsimonious model, we focus on a weekly time-series, where $Y_{i,t-1}$ denotes the demand observation of the same weekday in the previous week in our dataset (lag= 6). We choose product-weekday specific η in our analysis since they have the best out-of-sample predictive power. We use the first 10 weeks to initialize our model. Each week, we update the smoothing parameter.

We calculate the standard deviation of the forecast as root mean square error (*RMSE*):

$$\hat{\sigma}_i = \sqrt{\frac{1}{T_i - df_i} \sum_{t=1}^{T} (Y_{i,t} - \hat{\mu}_{i,t})^2} \tag{6.4}$$

with df being the degrees of freedom, i.e. $df = 1$ in the case of exponential smoothing.

Next, we explain the theoretical normative decision model for our empirical retailer.

6.4 Normative Decision Model

Our setting deals with a manufacturer selling N one-day products in multiple stores. Unit revenue for product i is r_i, production cost is c_i. Demand Y_i is stochastic with cumulative distribution function $F_i(y_i)$. For each day, the manufacturer has to determine the order quantity to produce and stock in each store. The goal of our decision maker is to maximize his expected profit, achieving an α service level $\hat{\alpha} \geq 70\,\%$ set by the retail stores. This service level is measured for each store separately each day as the fraction of non-stockout products.

We will now analyze how the product-specific service level target can be determined optimally.

6.4.1 Product-Specific Service Level

The service level is defined as the fraction of products in a store with positive inventory at the end of the day. The service level is measured as average over all products in a store. We define the product-specific service level target α_i as the in-stock probability of product i:

$$\alpha_i = Pr(\mathrm{IS}_i = 1) = F_i(q_i), \tag{6.5}$$

where q_i is the stocking quantity of product i. Therefore, the expected non-demand-weighted service level for a store $\bar{\alpha}$ with N products is:

$$\bar{\alpha} = \frac{1}{N} \sum_{i=1}^{N} \alpha_i \tag{6.6}$$

The optimization model can therefore be written as:

$$Z^* = \max_{q_i} Z(q_i) = \max_{q_i} \sum_{i=1}^{N} E_{Y_i}[\Pi_i(q_i)] \qquad s.t. \bar{\alpha} \geq \hat{\alpha} \tag{6.7}$$

with

$$E[\Pi_i(q_i)] = \int_0^\infty (r_i \ \min(q_i, y_i) - c_i \ q_i) f_i(y_i) dy_i, \tag{6.8}$$

being the expected profit of the stocking quantity q_i for the demand y_i.

A simple approach for achieving the service level constraint (based on an item approach) is to stock at least as many units of each product that $Pr(\text{IS}_i = 1) \geq \hat{\alpha}$, so $F_i(q_i) = \alpha_i \geq \hat{\alpha}$ for all products. It can be shown that this solution is not generally maximizing expected profits. A differentiation of service levels between products is favorable, e.g., it is beneficial to achieve a higher service level for cheap products, and lower the service level of expensive products, keeping the aggregated service level constant (Thonemann et al. 2002).

Solving the model in Eq. (6.7) we can derive the following Proposition concerning the optimal ordering decision:

Theorem 6.1 (Optimal Order Quantities Under System Approach) *The optimal product-specific service level targets α_i^* fulfill the following first order condition:*

$$\frac{\delta Z_i}{\delta \bar{\alpha}_i} = \frac{\delta Z_j}{\delta \bar{\alpha}_j} \qquad \forall i, j \tag{6.9}$$

$$\frac{c_i^o \ F_i(q_i) - c_i^u \ (1 - F_i(q_i))}{f_i(q_i)} = \frac{c_j^o \ F_j(q_j) - c_j^u \ (1 - F_j(q_j))}{f_j(q_j)} \qquad \forall i, j \tag{6.10}$$

where $c_i^o = c_i$ are product-specific overage costs, $c_i^u = r_i - c_i$ are product-specific underage costs, $\delta Z_i = \frac{dZ}{dq_i}$, and $\delta \bar{\alpha}_i = \frac{d\bar{\alpha}}{dq_i}$. In the optimum, the expected costs for a marginal service level increase are equal for all products.

Proof Lagrange function with λ representing the multiplier to the single service level constraint.

$$L = \sum_{i=1}^{N} ((r_i - c_i) q_i - r_i \int_0^{q_i} F_i(x) dx) + \frac{\lambda}{N} \sum_{i=1}^{N} F_i(q_i) \tag{6.11}$$

First-order condition (assuming that non-negativity constraints are never binding and all products are profitable).

$$\begin{aligned} \frac{\partial L}{\partial q_i} &= (r_i - c_i) - r_i F_i(q_i) + \frac{\lambda}{N} f_i(q_i) \qquad i = 1, \ldots, N \\ &= 0 \end{aligned} \tag{6.12}$$

$$0 = \lambda(\alpha - \sum_{i=1}^{N} F_i(q_i)) \tag{6.13}$$

If the service level constraint is not binding, i.e., $\lambda^* = 0$, then all products achieve their profit optimal level, i.e. $F_i(q_i) = \frac{r_i - c_i}{r_i}$. Otherwise,

$$\frac{\lambda}{N} = -\frac{r_i - c_i - r_i F_i(q_i)}{f_i(q_i)} \qquad \forall i = 1, \ldots, N \tag{6.14}$$

$$\lambda^* = \left(N\alpha - \sum_{i=1}^{N} \frac{r_i - c_i}{r_i}\right) \frac{N}{\sum_{i=1}^{N} f_i(q_i)} \tag{6.15}$$

which proves the proposition for the case of a binding constraint. □

From Theorem 6.1 we can derive the following corollary concerning the optimal service level targets:

Corollary 6.1 (Service Level Differentiation) *Keeping all other factors constant,*

(a) The service level of a product i decreases in its overage costs c_i^o.
(b) The service level of a product i increases in its underage costs c_i^u.
(c) The service level of a product i decreases in its demand variability σ_i.
(d) The service level of a product i is not effected by its mean demand μ_i.

We note that this corollary is a ceteris paribus analysis just changing one factor of one product. Changing attributes of multiple products, the relative change determines the change of SL per attribute of a product. Changing multiple attributes of one product, the final direction of the change is not clear, as the attributes interfere with each other.

We have analyzed the optimal service level determination for given demand distributions. Using this model, we can model the expected profit-maximizing behavior of a decision maker.

Finding the normative predictions of the setting, we will analyze the empirical data next.

6.5 Empirical Analysis

In this section, we analyze our empirical dataset to test the normative decision model. The analysis is divided into two parts. We test the described normative decision model by comparing the empirical data with the normative predictions and we analyze whether the decision maker is using a simpler model to determine his target service levels.

6.5.1 Expected Profit Maximization

Analyzing whether the order quantities of the decision maker are expected profit maximizing, we compare the decisions of the empirical decision maker with the predictions of the theoretical model. To determine the profit-maximizing benchmark, we calculate the optimal service level targets for each product and store for each day using Eq. (6.10). For each observation in our data set, we calculate the optimal forecast based on historical demand data according to Eqs. (6.1) and (6.2), using exponential smoothing. Given this forecast, we can optimize the daily stocking decisions for each store, using Eq. (6.10).

Figure 6.3 compares the service levels achieved by the decision maker with the optimal solution for all products. The rational profit-maximizing decision maker should achieve product-specific service levels along the 45 degree line. As we see in Fig. 6.3, this is hardly the case. We see a systematic bias in the observed service levels. For products with high normative predictions, the actual service levels are too low, for products with (relatively) low normative predictions, the empirical service levels are too high.

We take this observation as a first hint that the decision maker is making suboptimal decisions. Supporting this result we can compare the actual achieved profit of the decision maker to the profit of the optimal solution. We therefore simulate the performance of the optimal model for our data set, computing optimal orders for each store and product for all days. We then compute the average profit generated by the optimal model. As a result we find that our empirical decision maker is only achieving an efficiency of 92.4 % compared to the optimal solution. So we can conclude that the empirical decisions are indeed suboptimal. This motivates us to look deeper into the decisions to find the reasons for this performance gap.

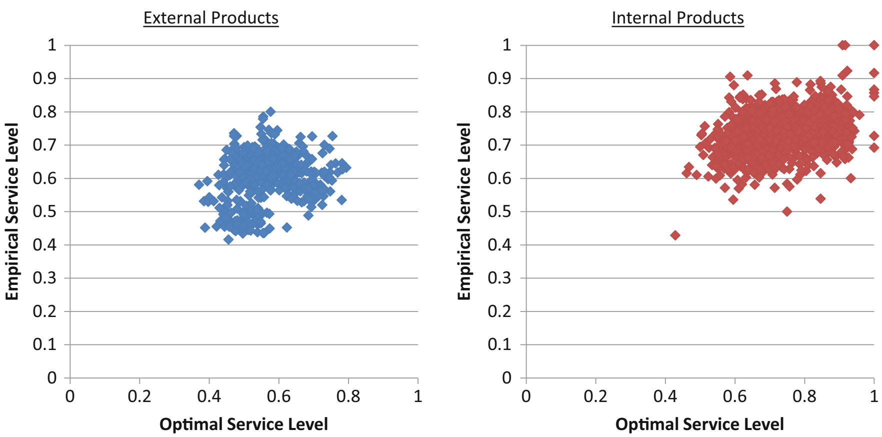


Fig. 6.3 Empirical service level vs. normative service level

Looking at the achieved service levels in more detail, we observe that the service levels differ between internal and external products (Fig. 6.3). But the differentiation within the groups seems to be too small. A regression analysis reveals that the decision maker does not systematically differentiate within product categories: For internal products, the differentiation is significant, but very small ($\beta_{SL_{Opt}} = 0.052$, $p = 0.005$). For external products, there is also a significant differentiation ($p = 0.000$), but in the opposite direction ($\beta_{SL_{Opt}} = -0.263$).

We take this as a first hint for a simpler optimization model the decision maker might be using. The question remains why actual service levels do not vary enough compared to the normative benchmark. According to Theorem 6.1, the service level should decrease in demand variability, and decrease in overage cost, but increase in underage costs. Therefore, we test whether the achieved empirical service levels are in line with this Theorem.

We use a random-intercept model (Eq. 6.16) to analyze the effect of different factors on empirical service levels. We model the products and stores as random effects:

$$\alpha_{i,s} = \beta_0 + \beta_1 Internal_i + \beta_2 \sigma_{i,s} + \beta_3 \mu_{i,s} + \beta_4 c_i^o + \beta_5 c_i^u + \upsilon_i + \epsilon_{i,s} \tag{6.16}$$

where *Internal* is an indicator which is 1 if the product is an internal product, and zero otherwise, υ are random errors for products. ϵ is the standard error term and β_k are the regression coefficients with $k = 0, \dots, 5$ and β_0 being the constant.

We also analyze the effect of these factors on the optimal service levels calculated by our aggregated service level approach. This is useful as Theorem 6.1 is only valid for ceteris paribus analyses. In our empirical setting we are not able to control for single parameter changes, e.g., demand variability σ_i might vary for multiple products between days and stores, and costs and revenues are constant over time and stores, but differ between products. Table 6.2 shows the results of our analyses. First, Table 6.2 shows that our general theoretical predictions of Theorem 6.1 also hold for the solution of our data set—the product-specific levels decrease in demand variability and overage costs, but increase in underage costs—mean demand does not affect service levels.

Table 6.2 also reveals insights into the decisions of our empirical decision maker. The second column shows that the empirical service levels are negatively influenced by demand variability, as predicted by the optimal model. But the decision maker does not take the actual overage and underage costs into consideration correctly. Overage costs do not significantly influence empirical service levels at all, underage costs are only slightly significant. Testing the size of the coefficients, we find that the empirical decision maker reacts significantly less to overage and underage costs than the optimal model predicts ($p < 0.01$ for both costs).

In the third model we included the dummy variable *Internal*, and we find that the classification as an internal product significantly increases service levels, which can be seen as an aggregated consideration of cost, as demand variability stays significant and similar in size (remember that internal products are more profitable than external products). Testing the three models, Model α_{EMP}^{ADJ} has the best fit

Table 6.2 Regression analysis: drivers of empirical service levels, standard errors in parenthesis

Factor	α_{OPT}	α_{EMP}	α_{EMP}^{ADJ}	α_{EMP}^{ADJ2}
Constant	0.876*** (0.033)	0.743*** (0.054)	0.648*** (0.021)	0.699*** (0.038)
σ_{FC}	−0.020*** (0.001)	−0.016*** (0.003)	−0.016*** (0.003)	−0.016*** (0.003)
μ_{FC}	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
c_i^o	−0.101*** (0.011)	−0.024 (0.018)	–	−0.014 (0.013)
c_i^u	0.081*** (0.01)	0.037* (0.020)	–	−0.039* (0.021)
Internal	–	–	0.128*** (0.024)	0.170*** (0.03)
# obs	1472	1472	1472	1472
# groups	23	23	23	23
Log-likelihood	3450	2377	2386	2382
AIC	−6886	−4740	−4761	−4749
BIC	−6849	−4703	−4729	−4707

p-value: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

(*LR*-test: $p < 0.01$) and also the smallest *AIC*/*BIC* for the empirical models (controlling for the additional parameter). In the last two columns we separated our analyses for internal and external products. There is a differentiation within external products, but partly in the wrong direction, i.e., underage costs have a significant negative impact.

Summarizing the results in this section, we find empirical evidence for suboptimal decision making. The empirical service levels differ between internal and external products, but not between actual cost differences within the groups, and too little for different demand variabilities. Using internal and external products for differentiation is actually a proxy for the correct cost-differentiation. But the cost differences within categories and the impact of demand variability are not considered adequately.

We use these observations to derive alternative (simpler) decision models in the next section.

6.5.2 Alternative Decision Models

To analyze the effect of these aspects, we set up four simplified decision models and apply them to our data set, varying the included details of cost-differentiation and demand variability.

Firstly, we calculate the naive or trivial decision model (Model 0) which is the standard newsvendor solution for a service level constraint for each product (ignoring costs and demand variability). *Note:* although we do not consider demand variability in the calculation of the optimal service level target, we use the demand variability to calculate the actual order quantity using the classical equation $q = \mu + z\,\sigma$.

Secondly, we determine the optimal order quantities if the decision maker only differentiates based on the variability of demand (Model 1).

Thirdly, we calculate the prediction of a decision process where the decision maker is adding cost differences into the service level optimization. Model 2 assumes that the decision maker only differentiates products based on cost differences between internal and external products. To do so, we calculate the average production/purchasing cost and revenues for internal and external products, and optimize the system approach ignoring demand variability.

We adapt our model of Eq. (6.8) by aggregating cost and revenue for our two groups (internal/external) using

$$E[\Pi_i(q_i)] = \int_0^\infty (R_i^{Group} \min(q_i, y_i) - C_i^{Group}\, q_i) f_i(y_i) dy_i, \tag{6.17}$$

where R_i^{Group} is the average revenue for internal product, for products $1, \ldots, 16$ and the average revenue for external products for products $16, \ldots, 23$. The same holds for production costs concerning C_i^{Group}.

Ignoring demand variability differences the first order condition for the optimal target service level is:

$$\frac{C^o_{Int}\,\alpha_{Int} - C^u_{Int}\,(1-\alpha_{Int})}{f_{N(0,1)}(z(\alpha_{Int}))} = \frac{C^o_{Ext}\,\alpha_{Ext} - C^u_{Ext}\,(1-\alpha_{Ext})}{f_{N(0,1)}(z(\alpha_{Ext}))}, \tag{6.18}$$

where $z(\alpha_{Int})$ is the value of the inverse cumulative distribution function of the standard normal distribution, and $f(\cdot)$ is the corresponding probability density function.

Using the second constraint

$$\hat{\alpha} \geq \sum_{p=1}^{16} \alpha_{Int} + \sum_{p=17}^{23} \alpha_{Ext} \tag{6.19}$$

we can derive the optimal service level for internal and external products.

Note: The target service levels are similar for all internal (external) products, as we use identical cost parameters within the groups and standard deviations even between groups (mean demand is irrelevant for optimal service level targets).

Using detailed, i.e. product-specific, cost information for all products results in Model 3. The different models are summarized in Table 6.3.

Table 6.3 Overview of decision models and *MSE* compared to empirical data

	Cost differentiation		
Demand variability	None	Group level	Individual
Individual	Model 1	**Model 2**	Model 3
	9.42	**9.25**	9.94

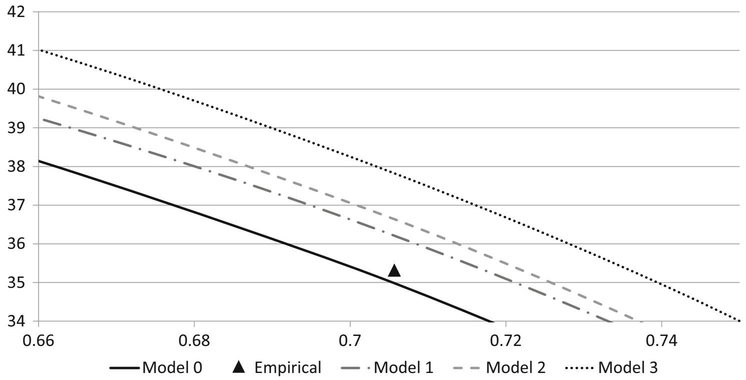


Fig. 6.4 Profits of Models 0, 1, 2, and 3 compared to empirical decision maker

6.5.3 Comparison of Alternative Decision Models with the Empirical Retailer

The regression analysis in Table 6.2 shows that the empirical differentiation is comparable to Model 2, i.e., he is differentiating between the groups, but not enough within the groups. For a detailed analysis we compare the predictive accuracy of our models with the empirical decision maker. Table 6.3 shows the MSE for the deviation of the order quantities of the different models compared to the empirical decision maker. We see that the MSE is smallest for Model 2.

To compare the statistical differences between the models, we use the Model Confidence Set (MCS) by Hansen et al. (2011). We determine the set of models performing significantly better than the other models at a significance level of 5 %. As a result, for our data set the MCS consists of Model 2 (bold in Table 6.3), concluding that Model 2 has a significantly better fit to the empirical decision maker (p-value: $p < 0.001$), but is not significantly different ($p > 0.9$). This is also in line with the regression results in Table 6.2, where demand variability was significant on top of the distinction between the groups.

We have seen so far that the decision maker seems to follow the rational behind Model 2 and differentiates between internal and external products. This is in some way in-line with the proxy-preferences. Figure 6.4 shows the profits of the four models compared to the empirical decision maker. Unsurprisingly, we see that Model 3 achieves a higher profit than Model 2 and Model 0 is the lowest. We also see that the empirical decision maker's profits lie between Model 0 and Model 2 meaning that he is better than the naive newsvendor, but not as good as our best fitting alternative decision models suggest. The question is why this is the case. Recent behavioral operations literature found additional aspects of decision making, such as demand chasing, mean chasing, and others. We will test these factors in the next section to analyze whether these aspects can be found in our empirical data and might be an explanation for the suboptimal performance of the decision maker compared to our simplified Model 2.

6.6 Additional Behavioral Aspects of Decision Making

Analyzing the strategic level of the decision maker in the previous sections, we now analyze additional behavioral aspects of the decision maker. Firstly, we test whether the decision maker is anchoring his ordering decision on certain anchors. Secondly, we test minimization of ex-post inventory errors. Thirdly, we analyze demand chasing, where the decision maker is anchoring his order quantities on previous demand realizations.

6.6.1 Anchoring and Adjustment

Many recent studies in the behavioral operations literature found evidence for anchoring of decision makers on mean demand. The order decisions are biased away from profit-maximizing quantities towards mean demand.

A natural anchor in this setting is the aggregated service level target. In the normative model the decision maker should differentiate service levels between products. Anchoring on the overall service level target leads to a smaller differentiation of the individual service level targets.

To test this observation analytically, we conduct a random-coefficient-regression (with products as the group variable) on the empirical order quantities compared to the predictions of the best fitting model (Model 2) and the prediction of the naive 70 % prediction:

$$q_{i,s,t} = \alpha_{70}\, q_{i,s,t}^{70\%} + (1-\alpha_{70}) q_{i,s,t}^{Model2} + \epsilon_{i,s}. \tag{6.20}$$

As we have two possible anchors in our setting, we also include anchoring on mean demand in the regressions:

$$q_{i,s,t} = \alpha_{70}\, q_{i,s,t}^{70\%} + \alpha_{\mu}\, \mu + (1-\alpha_{70}-\alpha_{\mu}) q_{i,s,t}^{Model2} + \epsilon_{i,s}. \tag{6.21}$$

The results of the regression are shown in Table 6.4.

We find evidence for an anchoring on the aggregated service level target, but just small anchoring on mean demand. Modeling an anchoring factor on mean demand

Table 6.4 Results of anchoring estimates

Parameters	(1)	(2)	(3)
α_{70}	0.582***		0.405***
α_{μ}		0.532***	0.268***
Log-likelihood	−1,000,799	−1,003,703	−1,000,275
BIC	2,001,662	2,007,471	2,000,665

***: $p < 0.001$

results in a worse fit than anchoring on the overall SL target. In general, both types of anchoring are possible, as service level targets can be below mean demand in extreme cases. Including both anchoring parameters into the regression shows similar results as the separate analyses.

6.6.2 *Minimizing Ex-Post Inventory Error*

As a second behavioral aspect we analyze the minimization of ex-post inventory error. This behavioral aspect can explain a pull-to-center bias (Schweitzer and Cachon 2000). Ho et al. (2010) model an asymmetric inventory error aversion allowing for a stronger aversion towards leftovers or stockouts. As the symmetric model is a special case of the asymmetric one, we use the latter one. Similar to Ho et al. (2010), we add an additional cost parameter to the overage and underage costs. If it is a true aversion against inventory errors, these factors should be the same over products. In our setting we use δ_u for the psychological underage cost and δ_o for the corresponding overage costs.

The optimal decision for a decision maker facing (asymmetric) ex-post inventory error has the following first order condition:

$$\frac{(C^o_{Int} + \delta_o)\,\alpha_{Int} - (C^u_{Int} + \delta_u)\,(1 - \alpha_{Int})}{16/23 f_{N(0,1)}(z(\alpha_{Int}))}$$
$$= \frac{(C^o_{Ext} + \delta_o)\,\alpha_{Ext} - (C^u_{Ext} + \delta_u)\,(1 - \alpha_{Ext})}{7/23 f_{N(0,1)}(z(\alpha_{Ext}))}. \tag{6.22}$$

Note that we use the aggregated Model 2 as a baseline model, as we found evidence for this in the previous section. As the optimal *SL*-target changes based on the psychological costs and the *SL*-target is not available in closed form solution we conduct an extensive numerical analysis. We calculate the optimal *SL*-targets for $0 \le \delta_o \le 2$ and $0 \le \delta_u \le 2$ to cover a feasible region of possible values. We then calculate the deviations of the resulting order quantities compared to our empirical decision maker.

The model achieves the best fit for $\delta_o = 0; \delta_u = 0.95$. The differences in the forecast errors between Model 2 and Model 2 with $\delta_o = 0; \delta_u = 0.95$ are significant ($p < 0.0001$, paired t-test of product-specific *MSE*). As a result we conclude that ex-post inventory error minimization is observable in our multi-product setting with SL targets.

6.6.3 Order Adaptation and Demand Chasing

As we do not have independent demand over time, decision makers observing demand realizations should adapt the forecast in order to account for autocorrelated time series. Based on the above forecasting (Sect. 6.3), we can analyze the normative effect of demand realizations on the production decisions of the decision maker. Comparing order quantities between weeks, we can derive Theorem 6.2 regarding rational changes in order quantities based on demand realizations.

Theorem 6.2 (Order Adaptation) *Orders are adapted as*

$$\Delta q_{i,t} = q_{i,t} - q_{i,t-1} = \Delta\mu + z\Delta\sigma, \tag{6.23}$$

where

$$\Delta\mu_{i,t} = \mu_{i,t} - \mu_{i,t-1}, \text{ and } \Delta\sigma_{i,t} = \sigma_{i,t} - \sigma_{i,t-1}.$$

In Theorem 6.2, $z = F_i^{-1}(\alpha_i)$ denotes the z-value of the normal distribution for the optimal product-specific target service level from Eq. (6.10). As a result, we see that the decision maker might react to demand realizations based on the forecast updating, as we do not have independent demand over time.

As a first analysis on reactions to demand realizations Fig. 6.5 shows the changes in order quantities between weeks. Adjustment towards demand realizations might be a good strategy according to forecast updating.

If the decision maker is subject to demand chasing, he over-adapts his order quantities. Therefore, we have to take *optimal* order adjustments into account. To test the demand chasing effect for our data we conduct a regression for Eq. (6.24). Formally, we estimate

$$\Delta q_{t,p,s} = a_0\, \Delta q^*_{t,p,s} + \sum \alpha_\kappa \left(d_{t-\kappa,p,s} - q_{t-\kappa,p,s}\right) + \epsilon_{t,p,s} + u_{s*p} \tag{6.24}$$

using a fixed-effect regression, controlling for each store-product combination $s * p$. To calculate the order quantity $q^*_{t,p,s}$ for each product p at time t in store s and its Δ, we use z_p corresponding to the optimal target service level according to the best-fitting Model 2. α_k is the chasing factor for the lag of k days. We model the demand chasing effect over the last $\kappa = 6$ working-days.

We use the same approach as in Sect. 6.5.1, namely Model 2. Table 6.5 shows the estimation results.

Fig. 6.5 Order adjustments

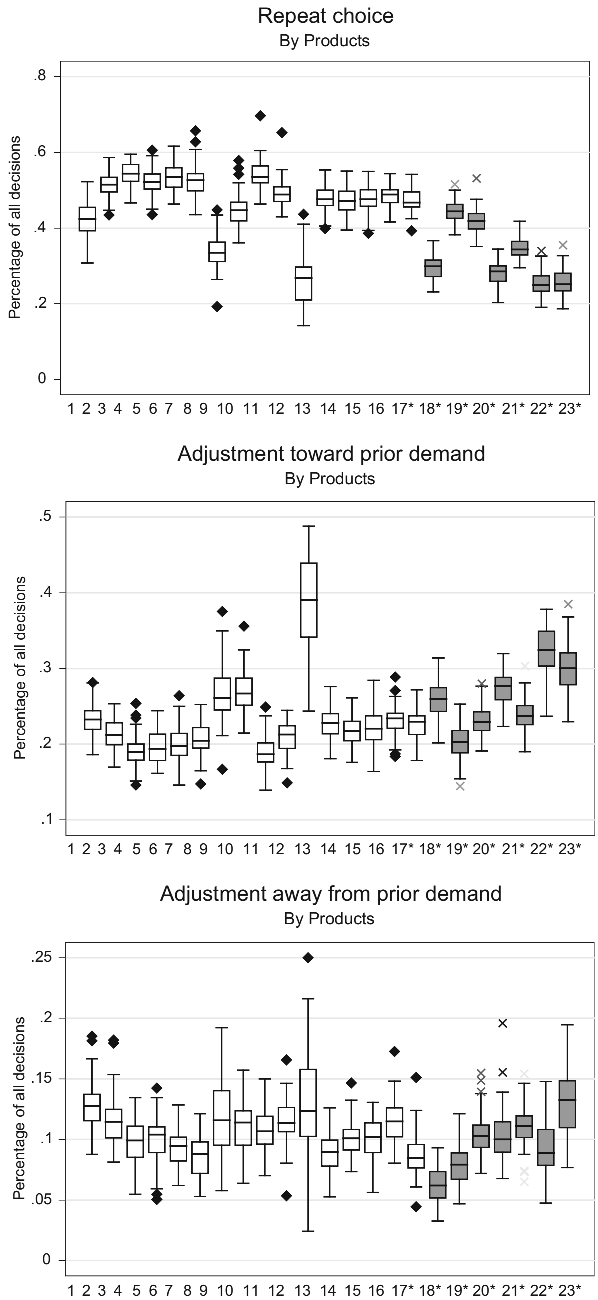


Table 6.5 Estimation results for demand chasing effects, standard errors in parentheses, all parameters significant at 0.01

Parameter	Opt M2	Baker	Baker M2
$\Delta q^*_{t,p}$	–	0	1
$(d-q)_{t-6}$	0.189	0.695	0.480
	(0.001)	(0.002)	(0.002)
Ind. stockout$_{t-6}$	0.265	0.406	0.274
	(0.007)	(0.013)	(0.016)
Log-likelihood	−506,803	−727,842	−732,136

Note that $\left(d_{t-\kappa,p} - q_{t-\kappa,p}\right) \leq 0$, as we have censored demand data. Therefore, we additionally test the effect of a stockout indicator which is 1 for a stockout and zero otherwise. Our model generally also enables a normative increasing of order quantities, e.g., in cases of stockouts, $\Delta q^*_{t,p}$ is most likely greater than zero, as we have an average $SL \approx 70\,\% > 50\,\%$ and the updated forecast tends to increase in stockout situations and so most likely $\Delta q^*_{t,p} > 0$.

We analyzed different models to investigate the effect of demand realizations. As a baseline we regress the optimal order adaptations using the stockouts and the leftovers of the previous week. As we use these factors to calculate the optimal order quantities, this yields a simplified model. We see that the decision maker should adopt his order quantities downwards in case of left-over inventory (0.189) and upwards in cases of stockouts (+0.265). Model *Opt M2* shows the overall effect of inventory and stockouts on order quantities. Using $a_0 = 0$ in Model *Baker*, we see that the decision maker decreases the order quantity of a product by 0.695 units for each item left over in the previous week and increases his order quantity if a product is out-of-stock in the previous week (+0.406). Comparing this to our optimal baseline model we see that the decision maker adapts stronger than expected in both cases. Model *Baker M2* analyzed this in detail and tests whether these adaptations are rational, setting $a_0 = 1$. We also tested the other lags (1–5) which partly have a significant impact, but the size of these effects is negligible.

We see that α_6 is significantly different from 0 for all models and we conclude that the decision maker is actually chasing demand with respect to the same weekday of the previous week. This chasing is not based on rational changes (Model *Baker M2*) and we conclude that the decision maker is chasing demand.

In a recent study Lau and Bearden (2013) analyze which method to use in order to test demand chasing. They show that several methods are prone to false detection of demand chasing, showing demand chasing where there is actually no chasing. They show that a simple correlation analysis is not prone to this failure. Therefore, we also conduct a correlation analysis between last period demands and recent order quantities. In order to account for the autocorrelated demand series we do not compare the correlation coefficients ρ against 0, but against the demand correlation. Table 6.6 shows the results of this analysis. We find that the correlation between last period demand and recent orders are significantly higher than the

Table 6.6 Correlation analysis to detect demand chasing

Product	$\rho(d_t, d_{t-6})$	$\rho(q_t, d_{t-6})$
1	0.55	0.75
2	0.51	0.71
3	0.48	0.67
4	0.41	0.66
5	0.32	0.54
6	0.32	0.56
7	0.50	0.74
8	0.58	0.76
9	0.61	0.77
10	0.55	0.71
11	0.62	0.73
12	0.47	0.68
13	0.56	0.74
14	0.29	0.69
15	0.66	0.75
16	0.30	0.52
17	0.63	0.75
18	0.53	0.70
19	0.53	0.72
20	0.45	0.61
21	0.73	0.83
22	0.53	0.81
23	0.47	0.70

demand correlation (paired t-test, $p < 0.0001$). This also shows the demand chasing behavior of our empirical decision maker.

As a third test of demand chasing we analyze the correlation of last period's inventory and the change of order quantities between periods (Rudi and Drake 2014). Figure 6.6 shows the changes of empirical and optimal orders and sales depending on inventories of the last week. We see a strong positive correlation between inventories of the last week and the changes between weeks ($\rho = 0.63$). Opposed to that optimal changes have a significant lower correlation ($\rho = 0.32$) and even slightly negative for sales ($\rho = -0.19$).

Our results show that demand chasing is not only a laboratory artifact, but also occurs in the real world. But the chasing is asymmetric, as the decision maker reacts more strongly on inventory than on stockouts.

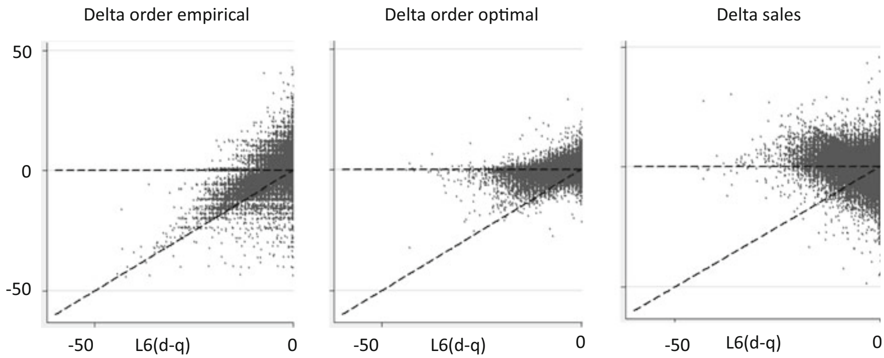


Fig. 6.6 Changes in quantities compared to inventories

6.7 Value of Product Characteristics: Managerial Insights

Applying these models to our data set, we can analyze the benefit these models offer with respect to the actual performance of the company.

To calculate the profits of our decision models, we simulate the actual performance of the different ordering policies with our real data. Due to unobservable lost sales, we do not know demand if the decision maker had a lower order quantity than our model. To estimate unobservable lost sales, we now choose the approach by Lau and Lau (1996). By taking a different approach than the one we used for the demand forecast, we allow for a fair comparison by avoiding any potential biases from again using the same estimation procedure.

Using our system approach, we achieve an average profit per store and day of 37.8 € for our data set with a similar service level as the empirical decision maker. The empirical decision maker only achieves 35.3 €. The optimal profit is 7.1 % higher than his profit. This supports our model and shows that the decision maker is not fully expected profit maximizing.

The profits for one specific setting of our models are shown in Fig. 6.7, including the empirical profit. For the profits shown in Fig. 6.7, we simulated the ordering decisions for each model for our data set, and calculated the profit the model would have achieved. We chose those models that achieved a service level similar to the empirical decision maker, as we have a clear service level profit trade-off.

Figure 6.7 also shows the value of integrating different product characteristics into the decision model. Taking demand variability into account increases profits by 3.5 % compared to the trivial model. Including cost information is the main driver in our model. While adding group-specific cost data to the model with demand variability increases the profit by 1.2 % compared to the trivial model, detailed cost data increases profits by further 3.4 % to a total value of 4.6 % for the cost information.

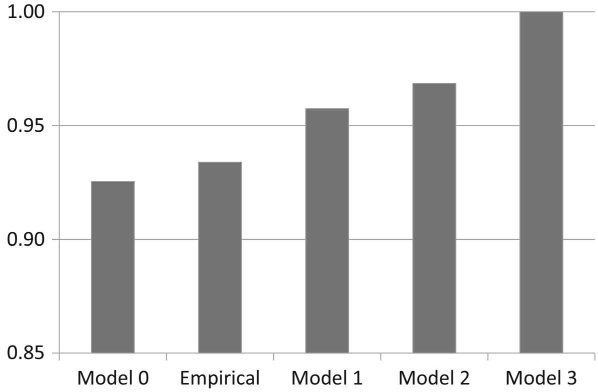


Fig. 6.7 Profits of different decision models (interpolated to achieve empirical SL of decision maker)

6.8 Conclusions

In this chapter, we study the situation of a newsvendor-like decision maker facing an aggregated service level contract. For this setting, we analyze the relevant cost drivers and derive the profit-maximizing order policy integrating product-specific costs and demand variability. We show that expected profits can be increased significantly if these factors are incorporated compared to a simple item approach. Additionally, we analyze the decision-making process in a real setting. While most of the recent behavioral operations studies focus on the cost-minimizing single product newsvendor in laboratory settings, we extend the research in two directions. Firstly, we analyze field data from a real decision maker. Secondly, the decision maker focuses on an aggregated service level contract.

Our findings show that the decision maker strategically differentiates service levels between internal and external products, although not perfectly maximizing his expected profit. While this differentiation is based on differences in profitability between products, the decision maker ignores additional aspects such as demand variability, and product-specific cost differences within these categories.

On the operational level, we find some additional explanations for the suboptimal performance: The decision maker is chasing demand. Adapting the order quantities to demand realizations is correct, because we do not necessarily have stationary demand. But our analyses show that the decision maker is over-reacting on demand realizations compared to normative changes predicted by the newsvendor model.

A limitation of our study is that our analyses use a simple exponential smoothing forecast ignoring additional external factors and substitution. This could be an area for possible extensions for future research.

Chapter 7
Conclusions

7.1 Summary

Point-of-sale scanner systems collect large amounts of data that can be used to make better informed decisions. The present work shows how available information such as selling prices and the timing of sales occurrences can be leveraged to better align supply and demand with retail analytics. For this purpose, we have collected data from a large European retail chain. We develop several models to improve stocking decisions and to analyze decision-making in the real world.

We suggest a novel approach that works directly with the data by analyzing causal relationships between demand and external variables. This data-driven approach integrates forecasting and inventory optimization. It is distribution-free and can also be applied if the underlying assumptions for other methods such as *OLS* regression analysis are violated. The problems are solved with Linear Programming. As a result, the decision maker obtains optimal order quantities by fitting a linear inventory function to historical demand observations and external variables affecting demand such as price, weather and weekdays.

The first model determines order decisions where the retail manager has full demand observations. We formulate the model for a cost-minimization objective and service level targets. Comparing the results of different approaches shows that the approaches taking causal relationships into account outperform the time-series forecasting method. The regression analysis yields better estimates for small sample sizes and in-stock target service levels. This holds only if the data meets the *OLS* assumptions, otherwise the *LP* approach achieves more robust inventory levels than regression analysis.

In the next model, the retailer is not able to observe demand, only sales. If a stockout occurs, demand is censored at the order-up-to level and the lost sales are unobserved. We establish sales patterns from days with full demand observations to estimate the unobservable lost sales. We integrate this aspect into our data-driven model. Based on demand data that follow the normal and the negative

A.-L. Sachs, *Retail Analytics*, Lecture Notes in Economics and Mathematical Systems 680, DOI 10.1007/978-3-319-13305-8_7

binomial distribution, we show that the *LP* model outperforms other parametric and non-parametric approaches for highly censored and price-dependent demand. Furthermore, we find that the non-parametric approaches achieve better results for real data where the true demand distribution is not known.

Additionally, we extend the data-driven model to problems with two products where stockout-based substitution takes place. If a product stocks out, sales of the product itself are censored and sales of the substitute are inflated by the additional demand that is shifted from the out-of-stock product to the substitute. We estimate the amount of unobservable lost sales and substituted demand based on the daily sales patterns. We compare the *LP* model to a parametric approach that estimates the parameters of a Poisson process with demand substitution. The data-driven model achieves higher profits than the parametric approach both for data with Poisson customer arrival rates and for real data where the demand distribution is unknown. We find that fitting the parameters of the model on highly censored data (censored at the median) yields slightly higher profits since more can be learnt about substitution behavior than based on data with almost no censoring (censored at 95 %).

Finally, we analyze empirical decisions in a real-world newsvendor setting. A manufacturer supplies several stores of a retail chain with bakery products. He aims to achieve an in-stock probability as service level target in each store over several products. We develop an aggregated service level model that optimally differentiates between the products by taking the demand variability, under- and overage costs into account. Comparing the results of our model to the decisions of the manufacturer shows that the manufacturer also considers some elements of the optimal policy, but his order quantities are not profit maximizing. A further analysis reveals that his decisions are subject to several behavioral biases: demand chasing, anchoring and ex-post inventory error minimization.

To sum up, the key contributions of this work are:

- Developing a data-driven model that integrates forecasting and inventory optimization in a single-period environment. It copes well with large amounts of data and is distribution-free. By fitting an inventory function to historical demands that depends on external factors such as price, it uses already available information from *POS* scanner systems to improve order decisions.
- Improving estimates of unobservable lost sales and substitution by establishing sales patterns based on timing information of sales.
- Validating the performance of the model based on real data that we collected from a large European retail chain.
- Formulation of a normative model to satisfy an aggregated service level constraint.
- Analyzing real-data for behavioral biases to explain the pull-to-center effect and showing that biases found in laboratory experiments are also present in real-world decisions.

7.2 Limitations and Future Research Directions

This research investigates only newsvendor-type decisions. The retail characteristics addressed are also found in multi-period problems for products with limited shelf-lives.

Using additional information and applying the data-driven model as a distribution-free approach could improve multi-period order decisions. If a product can be sold over several periods, order decisions should take the current inventory level into account. Additional questions such as whether customers make their purchases according to a first in first out (*FIFO*) or last in first out (*LIFO*) policy have to be discussed in this context.

The multi-period setting would also be an interesting area for further research on behavioral effects and the aggregated service level. If a decision maker knows that leftover inventory can be sold in future periods, it might affect his psychological costs of over- and underages.

The aggregated service level model discussed in this work determines the order quantities of several products in a store, but ignores potential substitution effects that may exist between the products. Taking this into account could provide additional profit gains compared to the model without substitution and the empirical decision maker.

Concerning substitution behavior, the data-driven model could be extended by allowing multiple substitution attempts and price-based substitution. Customers finding a potential substitute product offered at a much lower price than their first choice, are usually more willing to substitute than if the substitute is more expensive. This relationship could also be included into the model and might interfere with stockout-based substitution. Another interesting area for future research would be to include shortage penalty costs and to analyze the sequence of customer arrivals which would become relevant in this case. A challenging question is then how to estimate the shortage penalty costs to reflect the loss of customer goodwill in practice.

Bibliography

Aastrup, J., & Kotzab, H. (2010). Forty years of out-of-stock research - and shelves are still empty. *The International Review of Retail, Distribution and Consumer Research*, *20*(1), 147–164.

Agrawal, N., & Smith, S. A. (1996). Estimating negative binomial demand for retail inventory management with unobservable lost sales. *Naval Research Logistics*, *43*(6), 839–861.

Akkerman, R., Farahani, P., & Grunow, M. (2010). Quality, safety and sustainability in food distribution: A review of quantitative operations management approaches and challenges. *OR Spectrum*, *32*(4), 863–904.

Anderson, E. T., Fitzsimons, G. J., & Simester, D. (2006). Measuring and mitigating the costs of stockouts. *Management Science*, *52*(11), 1751–1763.

Anupindi, R., Dada, M., & Gupta, S. (1998). Estimation of consumer demand with stock-out based substitution: An application to vending machine products. *Marketing Science*, *17*(4), 406–423.

Aydin, G., & Porteus, E. L. (2008). Joint inventory and pricing decisions for an assortment. *Operations Research*, *56*(5), 1247–1255.

Becker-Peth, M., Katok, E., & Thonemann, U. (2013). Designing buyback contracts for irrational but predictable newsvendors. *Management Science*, *59*(8), 1800–1816.

Bell, P. C. (1981). Adaptive sales forecasting with many stockouts. *Journal of the Operational Research Society*, *32*(10), 865–873.

Bendoly, E. (2006). Incorporating behavioral theory in OM empirical models. *Journal of Operations Management*, *24*(6), 735–736.

Bendoly, E., Croson, R., Goncalves, P., & Schultz, K. (2010). Bodies of knowledge for research in behavioral operations. *Production and Operations Management*, *19*(4), 434–452.

Benzion, U., Cohen, Y., Peled, R., & Shavit, T. (2008). Decision-making and the newsvendor problem: An experimental study. *Journal of the Operational Research Society*, *59*(9), 1281–1287.

Berk, E., Gürler, U., & Levine, R. A. (2007). Bayesian demand updating in the lost sales newsvendor problem: A two-moment approximation. *European Journal of Operational Research*, *182*(1), 256–281.

Bertsimas, D., & Thiele, A. (2005). *A data-driven approach to newsvendor problems*. Working Paper, Massachusetts Institute of Technology, Cambridge, MA.

Bertsimas, D., & Thiele, A. (2006). A robust optimization approach to inventory theory. *Operations Research*, *54*(1), 150–168.

Beutel, A.-L., & Minner, S. (2012). Safety stock planning under causal demand forecasting. *International Journal of Production Economics*, *140*(2), 637–645.

A.-L. Sachs, *Retail Analytics*, Lecture Notes in Economics and Mathematical Systems 680, DOI 10.1007/978-3-319-13305-8

Bolton, G. E., & Katok, E. (2008). Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing & Service Operations Management*, *10*(3), 519–538.

Bolton, G. E., Ockenfels, A., & Thonemann, U. W. (2012). Managers and students as newsvendors. *Management Science*, *58*(12), 2225–2233.

Bostian, A., Holt, C., & Smith, A. (2008). Newsvendor pull-to-center effect: Adaptive learning in a laboratory experiment. *Manufacturing & Service Operations Management*, *10*(4), 590–608.

Burgin, T. (1975). The gamma distribution and inventory control. *Operational Research Quarterly*, *26*(3), 507–525.

Buzby, J. C., Wells, H. F., Axtman, B., & Mickey, J. (2009). *Supermarket loss estimates for fresh fruit, vegetables, meat, poultry, and seafood and their use in the ERS loss-adjusted food availability data*. Technical report, United States Department of Agriculture, Economic Research Service.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge: Cambridge university press.

Caro, F., & Gallien, J. (2007). Dynamic assortment with demand learning for seasonal consumer goods. *Management Science*, *53*(2), 276–292.

Chatfield, C. (2001). *Time-series forecasting*. London: Chapman & Hall/CRC.

Chen, L., & Plambeck, E. L. (2008). Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing & Service Operations Management*, *10*(2), 236–256.

Conrad, S. A. (1976). Sales data and estimation of demand. *Operations Research Quarterly*, *27*(1), 123–127.

Corsten, D., & Gruen, T. (2003). Desperately seeking shelf availability: An examination of the extent, the causes, and the efforts to address retail out-of-stocks. *International Journal of Retail & Distribution Management*, *31*(12), 605–617.

ECR (2003). *Optimal shelf availability: Increasing shopper satisfaction at the moment of truth*. Kontich: ECR Europe and Roland Berger.

Feiler, D. C., Tong, J. D., & Larrick, R. P. (2013). Biased judgment in censored environments. *Management Science*, *59*(3), 573–591.

Fildes, R., Nikolopoulos, K., Crone, S., & Syntetos, A. (2008). Forecasting and operational research: A review. *Journal of the Operational Research Society*, *59*(9), 1150–1172.

Fisher, M. L. (2009). Rocket science retailing: The 2006 Philip McCord Morse Lecture. *Operations Research*, *57*(3), 527–540.

Fisher, M. L., & Raman, A. (2010). *The new science of retailing: How analytics are transforming the supply chain and improving performance*. Boston: Harvard Business Press.

Gallego, G., & Moon, I. (1993). The distribution free newsboy problem: Review and extensions. *Journal of the Operational Research Society*, *44*(8), 825–834.

Gaur, V., & Honhon, D. (2006). Assortment planning and inventory decisions under a locational choice model. *Management Science*, *52*(10), 1528–1543.

Gilland, W. G., & Heese, H. S. (2013). Sequence matters: Shelf-space allocation under dynamic customer-driven substitution. *Production and Operations Management*, *22*(4), 875–887.

Gujarati, D., & Porter, D. (2009). *Basic econometrics* (5th ed.). New York: McGraw-Hill.

Gustavsson, J., Cederberg, C., Sonesson, U., van Otterdijk, R., & Meybeck, A. (2011). *Global food losses and food waste*. Technical report, Food and Agriculture Organization of the United Nations.

Halperin, M. (1952). Maximum likelihood estimation in truncated samples. *The Annals of Mathematical Statistics*, *23*(2), 226–238.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497.

Ho, T.-H., Lim, N., & Cui, T. H. (2010). Reference dependence in multilocation newsvendor models: A structural analysis. *Management Science*, *56*(11), 1891–1910.

Honhon, D., Gaur, V., Seshadri, S., & Rajagopalan, S. (2010). Assortment planning and inventory decisions under stockout-based substitution. *Operations Research*, *58*(5), 1364–1379.

Hopp, W. J., & Xu, X. (2008). A static approximation for dynamic demand substitution with applications in a competitive market. *Operations Research*, *56*(3), 630–645.

Hosoda, T., & Disney, S. (2009). Impact of market demand misspecification on a two-level supply chain. *International Journal of Production Economics*, *121*(2), 739–751.

Hsu, A., & Bassok, Y. (1999). Random yield and random demand in a production system with downward substitution. *Operations Research*, *47*(2), 277–290.

Hübner, A. H., Kuhn, H., & Sternbeck, M. G. (2013). Demand and supply chain planning in grocery retail: An operations planning framework. *International Journal of Retail & Distribution Management*, *41*(7), 512–530.

Huh, W. T., Levi, R., Rusmevichientong, P., & Orlin, J. B. (2011). Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator. *Operations Research*, *59*(4), 929–941.

Iyer, A., & Schrage, L. (1992). Analysis of the deterministic (s,S) inventory problem. *Management Science*, *38*(9), 1299–1313.

Jain, A., Rudi, N., & Wang, T. (2013). *Demand estimation and ordering under censoring: Stock-out timing is (almost) all you need*. Operations Research, forthcoming.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–291.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, *53*(282), 457–481.

Karabati, S., Tan, B., & Ozturk, O. C. (2009). A method for estimating stock-out-based substitution rates by using point-of-sale data. *IIE Transactions*, *41*(5), 408–420.

Karaesmen, I. Z., Scheller-Wolf, A., & Deniz, B. (2011). Managing perishable and aging inventories: Review and future research directions. In K. Kempf, P. Keskinocak, & R. Uzsoy (Eds.), *Handbook of production planning* (pp. 393–436). Wiesbaden: Kluwer.

Karakul, M., & Chan, L. M. A. (2008). Analytical and managerial implications of integrating product substitutability in the joint pricing and procurement problem. *European Journal of Operational Research*, *190*(1), 179–204.

Kässmann, G., Kühn, M., & Schneeweiß, C. (1986). Spicher's SB-Algorithmus Revisited - Feedback versus Feedforeward - Steuerung in der Lagerhaltung. *OR Spektrum*, *8*(2), 89–98.

Katok, E. (2011). Using laboratory experiments to build better operations management models. *Foundations and trends in technology, information and operations management*, *5*(1), 1–86.

Khouja, M. J. (2000). Optimal ordering, discounting and pricing in the single-period problem. *International Journal of Production Economics*, *65*(2), 201–216.

Kök, A. G., & Fisher, M. L. (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research*, *55*(6), 1001–1021.

Kök, A. G., Fisher, M. L., & Vaidyanathan, R. (2008). Assortment planning: Review of literature and industry practice. In N. Agrawal & S. Smith (Eds.), *Retail supply chain management* (pp. 99–153). Berlin: Springer.

Kranert, M., Hafner, G., Barabosz, J., Schuller, H., Leverenz, D., Kölbig, A., et al. (2012). Ermittlung der weggeworfenen Lebensmittelmengen und Vorschläge zur Verminderung der Wegwerfrate bei Lebensmitteln in Deutschland. *Kurzfassung. Institut für Siedlungswasserbau, Wassergüte-und Abfallwirtschaft*, *13*, 2012.

Kremer, M., Minner, S., & van Wassenhove, L. N. (2010). Do random errors explain newsvendor behavior? *Manufacturing & Service Operations Management*, *12*(4), 673–681.

Kremer, M., Minner, S., & van Wassenhove, L. N. (2014). On the preference to avoid ex-post inventory errors. *Production and Operations Management*, *23*, 773–787.

Kremer, M., Moritz, B., & Siemsen, E. (2011). Demand forecasting behavior: System neglect and change detection. *Management Science*, *57*(10), 1827–1843.

Krupp, J. A. (1997). Safety stock management. *Production and Inventory Management Journal*, *38*, 11–18.

Lau, H.-S., & Lau, A. H.-L. (1996). Estimating the demand distributions of single-period items having frequent stockouts. *European Journal of Operational Research*, *92*(2), 254–265.

Lau, H.-S., & Lau, A. H.-L. (1997). Some results on implementing a multi-item multi-constraint single-period inventory model. *International Journal of Production Economics*, *48*(2), 121–128.

Lau, N., & Bearden, J. N. (2013). Newsvendor demand chasing revisited. *Management Science*, *59*(5), 1245–1249.

Lee, H., & Billington, C. (1992). Managing supply chain inventory: Pitfalls and opportunities. *Sloan Management Review*, *33*(3), 65–73.

Lu, X., Song, J.-S., & Zhu, K. (2006). *Inventory control with unobservable lost sales and Bayesian updates*. Working Paper, University of California.

Lu, X., Song, J.-S., & Zhu, K. (2008). Analysis of perishable-inventory systems with censored demand data. *Operations Research*, *56*(4), 1034–1038.

Maddah, B., & Bish, E. K. (2007). Joint pricing, assortment, and inventory decisions for a retailer's product line. *Naval Research Logistics*, *54*(3), 315–330.

Mahajan, S., & van Ryzin, G. (2001a). Inventory competition under dynamic consumer choice. *Operations Research*, *49*(5), 646–657.

Mahajan, S., & van Ryzin, G. (2001b). Stocking retail assortments under dynamic consumer substitution. *Operations Research*, *49*(3), 334–351.

McGillivray, A. R., & Silver, E. A. (1978). Some concepts for inventory control under substitutable demand. *INFOR*, *16*(1), 47–63

Miller, C. M., Smith, S. A., McIntyre, S. H., & Achabal, D. D. (2010). Optimizing and evaluating retail assortments for infrequently purchased products. *Journal of Retailing*, *86*(2), 159–171.

Nagarajan, M., & Rajagopalan, S. (2008). Inventory models for substitutable products: Optimal policies and heuristics. *Management Science*, *54*(8), 1453–1466.

Nahmias, S. (1994). Demand estimation in lost sales inventory systems. *Naval Research Logistics*, *41*(6), 739–757.

Netessine, S., & Rudi, N. (2003). Centralized and competitive inventory models with demand substitution. *Operations Research*, *51*(2), 329–335.

Park, R. (1966). Estimation with heteroscedastic error terms. *Econometrica*, *34*(4), 888.

Parlar, M., & Goyal, S. K. (1984). Optimal ordering decisions for two substitutable products with stochastic demands. *OPSEARCH*, *21*(1), 1–15.

Pentico, D. W. (1974). The assortment problem with probabilistic demands. *Management Science*, *21*(3), 286–290.

Pentico, D. W. (2008). The assortment problem: A survey. *European Journal of Operational Research*, *190*(2), 295–309.

Petruzzi, N., & Dada, M. (1999). Pricing and the newsvendor problem: A review with extensions. *Operations Research*, *47*(2), 183–194.

Ren, Y., & Croson, R. (2013). Overconfidence in newsvendor orders: An experimental study. *Management Science*, *59*(11), 2502–2517.

Ridder, A., van Der Laan, E., & Salomon, M. (1998). How larger demand variability may lead to lower costs in the newsvendor problem. *Operations Research*, *46*(6), 934–936.

Ritchken, P., & Sankar, R. (1984). The effect of estimation risk in establishing safety stock levels in an inventory model. *Journal of the Operational Research Society, 35*(12), 1091–1099.

Ross, S. M. (2010). *Introduction to probability models* (10th ed.). New York: Academic.

Rudi, N., & Drake, D. (2014). Observation bias: The impact of demand censoring on newsvendor level and adjustment behavior. *Management Science*, *60*(5), 1334–1345.

Sachs, A.-L., & Minner, S. (2014). The data-driven newsvendor with censored demand observations. *International Journal of Production Economics*, *149*, 28–36.

Sauré, D., & Zeevi, A. (2013). Optimal dynamic assortment planning with demand learning. *Manufacturing & Service Operations Management*, *15*(3), 387–404.

Scarf, H. (1958). A min-max solution of an inventory problem. In K. J. Arrow, S. Karlin, & H. Scarf (Eds.), *Studies in the mathematical theory of inventory and production* (pp. 201–209). Stanford: Stanford University Press.

Schiffels, S., Fügener, A., Kolisch, R., & Brunner, O. J. (2014). On the assessment of costs in a newsvendor environment: Insights from an experimental study. *Omega*, *43*, 1–8.

Schweitzer, M. E., & Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science*, *46*(3), 404–420.

Silver, E., Pyke, D., & Peterson, R. (1998). *Inventory management and production planning and scheduling* (3rd ed.). New York: Wiley.

Smith, S. A., & Agrawal, N. (2000). Management of multi-item retail inventory systems with demand substitution. *Operations Research*, *48*(1), 50–64.

Spicher, K. (1975). Der SB1-Algorithmus. Eine Methode zur Beschreibung des Zusammenhangs zwischen Ziel-Lieferbereitschaft und Sicherheitsbestand. *Zeitschrift für Operations Research*, *19*(2), B1–B12.

Springael, J., & van Nieuwenhuyse, I. (2005). *A lost sales inventory model with a compound poisson demand pattern*. Working Papers 2005017, University of Antwerp, Antwerpen.

Strijbosch, L., & Moors, J. (1999). *Simple expressions for safety factors in inventory control*. Center Discussion Paper, No. 99112, Center for Economic Research, Tilburg University, Tilburg.

Su, X. (2008). Bounded rationality in newsvendor models. *Manufacturing & Service Operations Management*, *10*(4), 566–589.

Tan, B., & Karabati, S. (2004). Can the desired service level be achieved when the demand and lost sales are unobserved. *IIE Transactions*, *36*(4), 345–358.

Thonemann, U., Brown, A., & Hausman, W. (2002). Easy quantification of improved spare parts inventory policies. *Management Science*, *48*(9), 1213–1225.

Tiwari, V., & Gavirneni, S. (2007). ASP, the art and science of practice: Recoupling inventory control research and practice: Guidelines for achieving synergy. *Interfaces*, *37*(2), 176–186.

Topaloglu, H. (2013). Joint stocking and product offer decisions under the multinomial logit model. *Production and Operations Management*, *22*(5), 1182–1199.

Ulu, C., Honhon, D., & Alptekinoğlu, A. (2012). Learning consumer tastes through dynamic assortments. *Operations Research*, *60*(4), 833–849.

van Donselaar, K., Gaur, V., van Woensel, T., Broekmeulen, R., & Fransoo, J. (2010). Ordering behavior in retail stores and implications for automated replenishment. *Management Science*, *56*(5), 766–784.

van Donselaar, K., van Woensel, T., Broekmeulen, R., & Fransoo, J. (2006). Inventory control of perishables in supermarkets. *International Journal of Production Economics*, *104*(2), 462–472.

van Ryzin, G., & Mahajan, S. (1999). On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*, *45*(11), 1496–1509.

Vulcano, G., van Ryzin, G., & Ratliff, R. (2012). Estimating primary demand for substitutable products from sales transaction data. *Operations Research*, *60*(2), 313–334.

Wagner, H. (2002). And then there were none. *Operations Research*, *50*(1), 217–226.

Wecker, W. (1978). Predicting demand from sales data in the presence of stockouts. *Management Science*, 24(10), 1043–1054.

Yücel, E., Karaesmen, F., Salman, F. S., & Türkay, M. (2009). Optimizing product assortment under customer-driven demand substitution. *European Journal of Operational Research*, *199*(3), 759–768.

Zhao, X., & Atkins, D. R. (2008). Newsvendors under simultaneous price and inventory competition. *Manufacturing & Service Operations Management*, *10*(3), 539–546.

Zinn, W., & Marmorstein, H. (1990). Comparing two alternative methods of determining safety stock: The demand and the forecast systems. *Journal of Business Logistics*, *11*(1), 95–110.