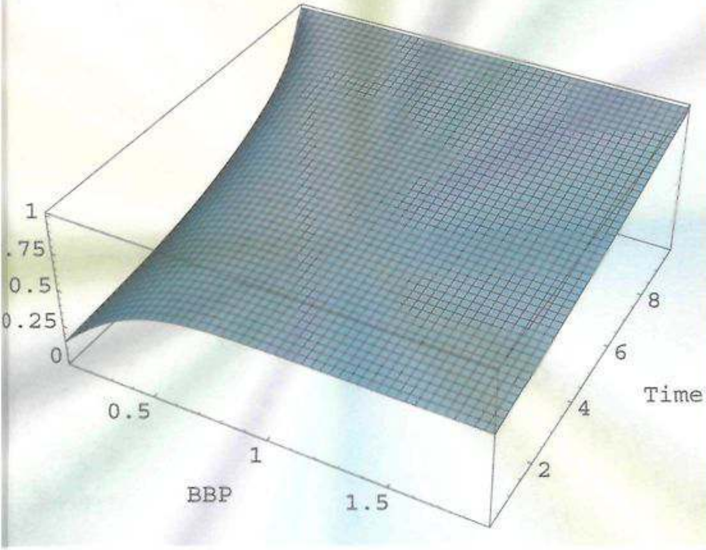


Editor

Jiongmin Yong

International Conference on Mathematical Finance

Recent Developments in Mathematical Finance

World Scientific

Recent Developments in Mathematical Finance

This page is intentionally left blank

Editor

Jiongmin Yong

Fudan University, China

International Conference on
Mathematical Finance

Recent Developments in Mathematical Finance

Shanghai, China 10 – 13 May 2001

World Scientific

New Jersey • London • Singapore • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

RECENT DEVELOPMENTS IN MATHEMATICAL FINANCE

ISBN 981-02-4797-4

Printed in Singapore by World Scientific Printers (S) Pte Ltd

Preface

The *International Conference on Mathematical Finance* was held at Fudan University, Shanghai, China, on May 10–13, 2001. Guests from 10 countries participated the conference and there were 29 invited speakers presented talks.

Mathematical finance is a very active area recently. The real world of finance and economics keep raising new and challenging problems, which make the subject very attractive. More and more advanced mathematical tools are found useful in the area. Some researches in mathematical finance lead to the invention of new mathematical theory. In this conference, the talks involved contingent claim pricing, optimal investment, interest term structure, insurance, risk analysis, numerical methods in finance, as well as backward and forward-backward stochastic differential equations, etc. The talks presented at the conference were of very high level and it is worthy of having a record. This volume contains 22 written version of the talks presented at the conference. The editor would like to thank the authors for their careful preparation of the contribution, which makes the publication of this proceedings possible.

The conference was financially sponsored by the Education Ministry of China, the National Natural Science Foundation of China, China Securites Regulatory Commission–Shanghai Regional Office, and Fudan University. On behalf of the organizing committee, the editor would like to acknowledge their supports.

Jiongmin Yong
Fudan University
October, 2001

Conference Sponsors:

The Education Ministry of China

The National Science Foundation of China

China Securites Regulatory Commission–Shanghai Regional Office

Fudan University

Organizing Committee:

Shuping Chen (Zhejiang University)

Lishang Jiang (Tongji University)

Shige Peng (Shandong University)

Jia-an Yan (Academy of Science Sinica)

Jiongmin Yong (Fudan University, Chairman)

Xun Yu Zhou (The Chinese University of Hong Kong)

Advisory Committee:

Darrell Duffie (Stanford University)

Ionnis Karatzas (Columbia University)

Xunjing Li (Fudan University)

Stanley Pliska (University of Illinois at Chicago)

Contents

Dynamic Asset Management: Risk Sensitive Criterion with Nonnegative Factors Constraints

Arunabha Bagchi
Faculty of Mathematical Sciences University Twente,
P.O.Box 217, 7500AE Enschede, The Netherlands
E-mail:bagchi@math.utwente.nl

K. Suresh Kumar
Department of Mathematics,
Indian Institute of Technology Bombay, Mumbai, India
E-mail:suresh@math.iitb.ac.in

Abstract

We extend an important recent work on risk-sensitive dynamic asset allocation to include nonnegativity constraints on the economic factors in the model. This is done in two steps. We first convert the dynamic asset allocation problem into an equivalent stochastic differential game. We then impose nonnegativity constraints on the game problem. We solve this new problem using some recent general results on such constrained stochastic differential games.

1 Introduction

Recently, Bielecki and Pliska [1] made an elaborate study on a continuous time portfolio optimization problem where the mean returns of individual securities are explicitly affected by underlying economic factors. For the purpose of detailed analysis, they considered linear (although quite general within this class) factor models with Brownian motion as the disturbance term. This means that we can never guarantee the factors to be always nonnegative. In this paper we study the same problem as in [1] under the constraint that the economic factors are always nonnegative.

The major difficulty of the optimization problem studied in [1] is that the criterion function is exponential of the state variable (the so-called risk sensitive criterion). This makes the study of the constrained problem hard. In this paper, we therefore proceed in two steps. We first convert the optimization problem into an equivalent stochastic game problem. The pay-off function of the game problem does not contain the exponential term. We then consider the constrained version of this equivalent game problem. General models of this sort have been recently studied by Ghosh and Suresh Kumar [2]. We use their results to solve the dynamic risk sensitive asset allocation problem with non-negativity constraints on the economic factors.

2 Risk Sensitive Dynamic Asset Allocation: The Unconstrained Case

Let us briefly outline the risk sensitive asset allocation problem as formulated and solved by Bielecki and Pliska [1]. Let $(\Omega, \mathcal{J}, \mathcal{P}, \mathcal{J}_{\sqcup})$ be the real world probability space. Our market consists of $m \geq 2$ securities and $n \geq 1$ economic factors. Let $S(t) = (S_1(t), \cdots, S_m(t))$ denote the price vector of the securities and $X(t) = (X_1(t), \cdots, X_n(t))'$ denote the value vector of the factors at time t. The classical Merton problem [3] deals with the best investment strategy for the investor in order that his/her utility of the terminal wealth is maximized.

Let $h(t) = (h_1(t), \cdots, h_m(t))'$ denote the fraction of the investment on securities at time t. A strategy $h(\cdot)$ is called admissible if:

(i) $h(t) \in U \subseteq \mathbb{R}^m$,

(ii) $h(t)$ is adapted to the information until time t; that is, $h(t) \in \mathcal{J}_{\sqcup}$,

(iii) $\mathcal{P}\left(\int_{\prime}^{\sqcup} |\langle(\int)|^{\in} \lceil \int < \infty\right) = \infty$ for all t.

We denote the class of admissible strategies by U_1. In [1] it is assumed that $S(t)$ and $X(t)$ evolve as

$$\frac{dS_i(t)}{S_i(t)} = (a + AX(t))_i dt + \sum_{k=1}^{m+n} \sigma_{i_k} dW_k(t), S_i(0) = s_i, \quad i = 1, \cdots, m \tag{1}$$

$$dX(t) = (b + BX(t))dt + \Lambda dW(t), \quad X(0) = x \tag{2}$$

where $\Lambda\Lambda'$ is positive definite and $\Sigma\Lambda' = 0$. The last assumption basically means that the "noises" affecting S and X are independent. Here $\{W(t)\}$ is $\mathbb{R}^{n+m}$ standard Brownian motion. We do not impose any constraints at present on the components of X. It is easy to see that the wealth process $V(t)$ corresponding to the investment strategy $h(\cdot)$ evolves as

$$\begin{aligned} dV(t) &= h(t)'(a + AX(t))V(t)dt + V(t)h(t)'\Sigma dW(t) \\ V(0) &= v, \end{aligned} \tag{3}$$

where Σ is the matrix with σ_{i_k} its i_k-th component. Together with the equation (2) for $X(t)$, we obtain the dynamics of the state of our stochastic system. It is important to note that the wealth process above makes the strategy $h(\cdot)$ self financing.

The risk sensitive cost criterion differs from the traditional cost criterion in a fundamental way. For the finite horizon case, the cost criterion is given by

$$J_T = \left(-\frac{2}{\theta}\right) \log E^{h(\cdot)} \left[e^{-\left(\frac{\theta}{2}\right)\log V(T)} | V(0) = v, X(0) = x\right] \tag{4}$$

while for the infinite horizon case, the cost criterion is given by

$$J = \varliminf_{T\to\infty} \frac{1}{T}\left(-\frac{2}{\theta}\right) \log E^{h(\cdot)} \left[e^{-\left(\frac{\theta}{2}\right)\log V(T)} | V(0) = v, X(0) = x\right] \tag{5}$$

Both the finite and infinite horizon problems can be solved using the method of dynamic problem. In the infinite horizon case, Bielecki and Pliska [1] solved the problem under the following assumption:

$$\lim_{\|x\|\to\infty} K_\theta(x) = -\infty \tag{6}$$

where

$$K_\theta(x) \triangleq \inf_{h\in U, \Sigma h_i = 1} \left[\frac{1}{2}\left(\frac{\theta}{2}+1\right) h'\Sigma\Sigma' h - h'(a + Ax)\right], x \in \mathbb{R} \tag{7}$$

This optimization problem, in the finite horizon case, has a long history in control theory. Jacobson [4] first solved this problem in the case of a quadratic criterion. The general linear-quadratic-exponential problem has been studied in the book of Whitle [5]. For the partial observation case, see Bensoussan and Van Schuppen [6]. It soon became apparent that this optimization problem is equivalent to a stochastic differential game without the exponentiation of the cost criterion [7]. It turned out to be also equivalent to some robust control problem. We focus here on the equivalent stochastic differential game and show that one of its optimal strategies is an optimal investment strategy for the investor.

Consider the following dynamics for the wealth process $V(\cdot)$, which is controlled by the investor's strategy $h(\cdot)$ and another player's strategy $w(\cdot)$ which may be thought of as "an uncertainty imposed by nature":

$$\begin{aligned} dV(t) &= h(t)'\left(a + AX(t) + h(t)\Sigma w(t)\right)V(t)dt \\ &\quad + h(t)'\Sigma V(t)dW(t), V(0) = v \end{aligned} \tag{8}$$

$$dX(t) = (b + BX(t))dt + \Lambda dW(t), X(0) = x \tag{9}$$

Note that we have made the drift term in the dynamics of the wealth process being controlled by the "uncertainty of nature". We may now think of a game played by the investor and "nature", where the investor wants to choose a strategy to maximize his benefit even in the worst case when the "nature" is trying to minimize it. We have thus a zero-sum stochastic differential game.

Let us make our pay-offs explicit. For the finite-horizon case, the pay-off is

$$J_\theta(x, v, h(\cdot)) = \varliminf_{T\to\infty} \frac{1}{T} E^{h(\cdot),w(\cdot)}[\log V(T) + \frac{1}{\theta}\int_0^T |w(t)|^2 dt| \quad X(0) = x, V(0) = v] \tag{10}$$

3 Equivalence of the Game Problem with the Asset Allocation Problem

We start with the finite horizon case with the pay-off (10). Let us write down the Hamilton-Jacobi-Issac (HJI) equation corresponding to this problem:

$$\begin{aligned}\frac{\partial\psi}{\partial t} &= \inf_{w\in\mathbf{R}^{n+m}} \sup_{h\in U_1} [\frac{\partial\psi}{\partial v} h'(a + Ax + \Sigma w) + \Lambda w \nabla_x \psi \\ &+ \frac{1}{2}\frac{\partial^2\psi}{\partial v^2} h'\Sigma\Sigma' h v^2 + \frac{1}{\theta}|w|^2] + (b + Bx)'\nabla_x\psi \\ &+ \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2\psi}{\partial x_i \partial x_j} \sum_{k=1}^{n+m} \lambda_{ik}\lambda_{jk} \\ &= \sup_{h\in U_1} \cdot \inf_{w\in\mathbf{R}^{n+m}} \cdot [\text{ the same expression as above }]\end{aligned} \tag{11}$$

with the terminal condition

$$\psi(T, x, v) = \log v \tag{12}$$

Then we have the following:

Theorem 3.1 *The value function ψ of the finite horizon game posed at the end of section 2 exists and is the unique solution to (11)-(12) in $C^{1+\alpha/2,2+\alpha}([0,T]\times \mathbb{R}^{n+1})$, for some α with $0 < \alpha < 1$.*

Proof The proof follows by an application of Fan's minimax Theorem [8] and the measurable selection Theorem [9]. We omit the proof.

We now state a series of theorems without proofs.

Theorem 3.2 *There exists a saddle point $(h^*(\cdot), w^*(\cdot))$ to the finite horizon game formulated above.*

Theorem 3.3 *The value function ψ is the optimal value of the asset allocation problem and $h^*(\cdot)$ of theorem 3.2 is an optimal strategy.*

Let us now turn to the infinite horizon case. Consider the following Cauchy problem:

$$\begin{aligned}
&\frac{\partial \bar{\phi}}{\partial t} = \sup_{h \in U_1} \cdot \inf_{w \in \mathbf{R}^{n+m}} [\frac{\partial \bar{\phi}}{\partial v} \left(h'(a + Ax) + h'\Sigma w\right) v \\
&+\Lambda w \nabla_x \bar{\phi} + \frac{1}{2} \frac{\partial^2 \bar{\phi}}{\partial v^2} h'\Sigma\Sigma' h v^2 + \frac{1}{\theta} |w|^2] \\
&+(b + Bx)' \nabla_x \bar{\phi} + \tfrac{1}{2} \textstyle\sum_{i,j=1}^{n} \frac{\partial^2 \bar{\phi}}{\partial x_i \partial x_j} \sum_{k=1}^{n+m} \lambda_{ik} \lambda_{jk} \\
&\bar{\phi}(0, x, v) = \log v
\end{aligned} \tag{13}$$

It is easy to see that $\bar{\psi}(t, x, v) \triangleq \bar{\phi}(T - t, x, v)$ satisfies (11)-(12).

Theorem 3.4 *Let $\bar{\phi}$ be the solution of (13), Then, as $t \to \infty$, $\bar{\phi}(t, x, v) - \bar{\phi}(t, 0, 1) \to \phi(x, v)$ in $W^{1,2}_{loc}$ and $\frac{\partial \bar{\phi}}{\partial t} \to \rho_\theta$, where (ρ_θ, ϕ) satisfies*

$$\begin{aligned}
&\rho_\theta = \sup_{h \in U_1} \cdot \inf_{w \in \mathbf{R}^{n+m}} [\frac{\partial \bar{\phi}}{\partial v} \left(h'(a + Ax) + h'\Sigma w\right) v \\
&+\Lambda w \nabla_x \phi + \frac{1}{2} \frac{\partial^2 \phi}{\partial v^2} h'\Sigma\Sigma' h v^2 + \frac{1}{\theta} |w|^2] \\
&+(b + Bx)' \nabla_x \phi + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \phi}{\partial x_i \partial x_j} \sum_{k=1}^{n+m} \lambda_{ik} \lambda_{jk} \\
&\phi \in C^2(\mathbb{R}^n \times \mathbb{R}), \lim_{\|x\| \to \infty} \phi(x) = \infty
\end{aligned} \tag{14}$$

Remark 3.5 *The proof is crucially based on some results of Nagai*[10].

Theorem 3.6 *ρ_θ is the value of the game and $(h^*(\cdot), w^*(\cdot))$ a saddle-point if the pair satisfies*

$$\begin{aligned}
&\sup_{h \in U_1} \cdot \inf_{w \in \mathbf{R}^{n+m}} [\frac{\partial \bar{\phi}}{\partial v} (h'(a + Ax) + h'\Sigma w) v + \Lambda w \nabla_x \bar{\phi} \\
&+\frac{1}{2} \frac{\partial^2 \bar{\phi}}{\partial v^2} h'\Sigma\Sigma' v^2 + \frac{1}{\theta} |w|^2] \\
&= \inf_{w \in \mathbf{R}^{n+m}} [\frac{\partial \bar{\phi}}{\partial v} (h^*)'(a + Ax) + (h^*)'\Sigma w) v + \Lambda w \nabla_x \bar{\phi} \\
&+\frac{1}{2} \frac{\partial^2 \bar{\phi}}{\partial v^2} (h^*)'\Sigma\Sigma' h^* v^2 + \frac{1}{\theta} |w|^2]
\end{aligned} \tag{15}$$

and

$$
\begin{aligned}
&\inf_{w\in\mathbf{R}^{n+m}} \cdot \sup_{h\in U_1} \cdot [\frac{\partial\bar{\phi}}{\partial v}\left(h'(a+Ax)+h'\Sigma w\right)v+\Lambda w\nabla_x\bar{\phi} \\
&+\frac{1}{2}\frac{\partial^2\bar{\phi}}{\partial v^2}h'\Sigma\Sigma' h v^2+\frac{1}{\theta}|w|^2] \\
&=\sup_{h\in U_1} \cdot [\frac{\partial\bar{\phi}}{\partial v}\cdot\left(h'(a+Ax)+h'\Sigma w^*\right)v+\Lambda w^*\nabla_x\bar{\phi} \\
&+\frac{1}{2}\frac{\partial^2\bar{\phi}}{\partial v^2}h'\Sigma\Sigma' h v^2+\frac{1}{\theta}|w^*|^2]
\end{aligned}
\tag{16}
$$

where $\bar{\phi}$ is the solution of (13).

Theorem 3.7 *ρ_θ is the value of the optimal allocation problem and $h^*(\cdot)$ is an optimal strategy.*

4 Risk Sensitive Dynamic Asset Allocation: The Constrained Case

We have now arrived at the main part of the paper. So far we did not impose any constraints on the components of X. If we look at equation (2), we see that the economic factors are modelled as mean-reverting to make it unlikely for the factors to be negative. But there is every possibility that some of the economic factors for some time intervals (however small) will be negative. This is often not realistic from economic considerations. We study here the risk sensitive asset allocation problem under the constraint that $X_i(t), i=1,\cdots,n$, are nonnegative a.s. for all $t\geq 0$. To handle this constraint, we model the process $X(\cdot)=(X_1(\cdot),\cdots,X_n(\cdot))'$ as a reflecting diffusion [11]. More precisely, in the game theoretic framework studied in the preceding section, we take the dynamics of the wealth process $V(\cdot)$ to be given by

$$
\begin{aligned}
dV(t) &= h(t)'(a+AX(t)+\Sigma w(t))V(t)dt+h(t)'\Sigma V(t)dW(t) \\
dX(t) &= (b+BX(t)+\Lambda w(t))dt+\Lambda dW(t)-\gamma_1(X(t))d\xi_1(t) \\
d\xi_1(t) &= I\{X(t)\epsilon\partial D_1\}d\xi_t \\
V(0) &= v,\ \ X(0)=x,\ \ \xi_1(0)=0
\end{aligned}
\tag{17}
$$

where D_1 is the positive orthrant of $\mathbb{R}^n$ and $\gamma_1:\mathbb{R}^n\to\mathbb{R}^n$ is the direction of reflection. Note that the case of $n\geq 2$ is very difficult mathematically because of the nonsmooth nature of the boundary. In this paper, we address only the case where $n=1$; that is, our domain $D_1=\mathbb{R}_+$.

Set $D=\mathbb{R}^2_+$, the upper half plane, and $\gamma:\mathbb{R}^2\to\mathbb{R}^2$ be such that

$$
\gamma(v,0)=(0,\gamma_1(v)),\ \text{for a suitable}\ \gamma_1:\mathbb{R}\to\mathbb{R}. \tag{18}
$$

Consider the controlled stochastic differential equation

$$d\begin{pmatrix} V(t) \\ X(t) \end{pmatrix} = \begin{pmatrix} h(t)'(a + AX(t) + \Sigma w(t))V(t) \\ b + BX(t) + \Lambda w(t) \end{pmatrix} dt$$

$$+ \begin{pmatrix} h(t)'\Sigma V(t) \\ \Lambda \end{pmatrix} dW(t) - \gamma \begin{pmatrix} V(t) \\ X(t) \end{pmatrix} d\xi(t)$$

$$d\xi(t) = I\left\{\begin{pmatrix} V(t) \\ X(t) \end{pmatrix} \epsilon \partial D\right\} d\xi(t) \tag{19}$$

$$\begin{pmatrix} V(0) \\ X(0) \end{pmatrix} = \begin{pmatrix} v \\ x \end{pmatrix}, \quad \xi(0) = 0$$

Note that $h(t) \in \mathbb{R}^m, a \in \mathbb{R}^m$; A is an $m \times 1$ matrix, Σ is an $m \times (m+1)$ matrix, $w(t) \in \mathbb{R}^{m+1}$, $b, B \in \mathbb{R}$ and Σ is a $1 \times (m+1)$ matrix.

Noting that $V(t) > 0$ for all t a.s., we see that the processes (17) and (19) are indistinguishable. From now on we use (19) to describe the dynamics of the wealth process. We restrict ourselves for simplicity to normal reflection in what follows.

Remark 4.1 *We use this special form of γ so as to be in the framework of Neuman boundary conditions. We can extend the results to general γ with nontangential assumption using the arguments given in Bensoussan and Lions* [12].

Let us begin with the finite horizon problem. The criterion function is

$$J_\theta^T(x, v, h(\cdot), w(\cdot)) = E^{h(\cdot),w(\cdot)}[\log V(T) + \frac{1}{\theta}\int_0^T |w(s)|^2 ds| \\ X(0) = x, V(0) = v] \tag{20}$$

As before, we introduce the HJI equation

$$\tfrac{\partial \psi}{\partial t} = \sup_h \cdot \inf_w \cdot [\nabla\psi \cdot \begin{pmatrix} h'(a + Ax + \Sigma w)v \\ b + Bx + \Lambda w \end{pmatrix} + \tfrac{1}{\theta}|w|^2$$

$$+ \tfrac{1}{2} Tr\left(\begin{pmatrix} h'\Sigma \\ \Leftarrow \end{pmatrix} H\psi \begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix}'\right] \tag{21}$$

$$= \sup_h \cdot \inf_w \cdot [\text{ the same expression as above}$$

with

$$\psi(T,x,v) = \log v$$
$$D\psi \cdot \gamma = 0 \text{ on } [0,T] \times \partial D \tag{22}$$

where H denotes the Hessian matrix.

Then we can prove the following:

Theorem 4.2 *The value function ψ of the finite horizon constrained game problem exists and is the unique solution of (22) in $C^{1+\alpha/2,2+\alpha}((0,T)\times D)\cap C([0,T]\times \bar{D})$, for some $\alpha > 0$.*

The arguments leading to the proof of Theorem 4.2, which are here omitted, lead to the following characterization of the optimal strategies:

Theorem 4.3 *There exists a saddle point $(h^*(\cdot), w^*(\cdot))$ to the finite horizon constrained game studied here.*

Finally, we can show that the value function ψ of this game problem is the value function of the original constrained dynamic asset allocation problem.

Theorem 4.4 *The value function ψ is the optimal value of the constrained asset allocation problem and $h^*(\cdot)$ of Theorem 4.2 is an optimal strategy.*

Let us now turn to the infinite horizon case. The associated cost functional is

$$\begin{aligned} J_\theta(x,v,h(\cdot),w(\cdot)) &= \lim_{T\to\infty} \frac{1}{T} E^{h(\cdot),w(\cdot)}[\log V(T) + \frac{1}{\theta}\int_0^T |w(t)|^2 dt| \\ V(0) = v, X(0) = x] \end{aligned} \tag{23}$$

The idea is to try to look at the asymptotic behavior of ψ as $T\to\infty$. This can only be done indirectly by reversing the time. Consider the following partial differential equation

$$\begin{aligned} \tfrac{\partial\phi}{\partial t} &= \sup_h \cdot \inf_w \cdot [\nabla\phi \begin{pmatrix} h'(a + Ax + \Sigma w)v \\ b + Bx \end{pmatrix} + \tfrac{1}{\theta}|w|^2 \\ &+ \tfrac{1}{2} Tr\left(\begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix} H\phi \begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix} \right)'] \\ &= \inf_w \cdot \sup_h \cdot [\text{ the same expression as above }] \end{aligned} \tag{24}$$

with

$$\phi(0,x,v) = \log v$$
$$\nabla\phi \cdot \gamma = 0 \text{ on } [0,T] \times \partial D \tag{25}$$

We now can prove the following theorem:

Theorem 4.5 *Let ϕ be a solution to (25). Then as $t \to \infty$, $\phi(t,x,v) - \phi(t,0,1) \to \tilde{\phi}(x,v)$ weakly in $W^{1,2}_{loc}$ and $\frac{\partial \phi}{\partial t} \to \rho_\theta$, where $(\rho_\theta, \tilde{\phi})$ satisfies*

$$\begin{aligned} \rho_\theta &= \sup_h \cdot \inf_w \cdot [\nabla \tilde{\phi} \begin{pmatrix} h'(a + Ax + \Sigma w) \\ b + Bx \end{pmatrix} + \tfrac{1}{\theta}|w|^2 \\ &+ \tfrac{1}{2} Tr \left(\begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix} H\tilde{\phi} \begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix}' \right)] \\ &= \inf_w \cdot \sup_h \cdot [\text{ the same expression as above }] \end{aligned} \tag{26}$$

$$\nabla \tilde{\phi} \cdot \gamma = 0 \quad \text{on } \partial D$$

Theorem 4.6 *ρ_θ is the value of the game and (h^*, w^*) is a saddle point if (h^*, w^*) satisfies*

$$\begin{aligned} &\sup_h \cdot \inf_w \cdot [\nabla \tilde{\phi} \begin{pmatrix} h'(a + Ax + \Sigma w) \\ b + Bx \end{pmatrix} + \tfrac{1}{\theta}|w|^2 \\ &+ \tfrac{1}{2} Tr \left(\begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix} H\tilde{\phi} \begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix}' \right)] \\ &= \inf_w \cdot [\nabla \tilde{\phi} \begin{pmatrix} h^*(a + Ax + \Sigma w) \\ b + Bx \end{pmatrix} + \tfrac{1}{\theta}|w|^2 \\ &+ \tfrac{1}{2} Tr \left(\begin{pmatrix} h^{*\prime}\Sigma v \\ \Lambda \end{pmatrix} H\tilde{\phi} \begin{pmatrix} h^{*\prime}\Sigma v \\ \Lambda \end{pmatrix} \right)] \end{aligned} \tag{27}$$

and

$$\begin{aligned} &\inf_w \cdot \sup_h \cdot [\nabla \tilde{\phi} \begin{pmatrix} h'(a + Ax + \Sigma w) \\ b + Bx \end{pmatrix} + \tfrac{1}{\theta}|w|^2 \\ &+ \tfrac{1}{2} Tr \left(\begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix} H\tilde{\phi} \begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix}' \right)] \\ &= \sup_h \cdot [\nabla \tilde{\phi} \begin{pmatrix} h'(a + Ax + \Sigma w^*) \\ b + Bx \end{pmatrix} + \tfrac{1}{\theta}|w^*|^2 \\ &+ \tfrac{1}{2} Tr \left(\begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix} H\tilde{\phi} \begin{pmatrix} h'\Sigma v \\ \Lambda \end{pmatrix}' \right)] \end{aligned} \tag{28}$$

where $(\rho_\theta, \tilde{\phi})$ *satisfies (26).*

Finally we have

Theorem 4.7 ρ_θ *is the value of the constrained dynamic asset allocation problem and* $h^*(\cdot)$ *of Theorem 4.5 is an optimal asset allocation strategy.*

5 Conclusion

We studied in this paper the risk sensitive dynamic asset allocation problem under the constraint that the "economic factors" in the model are nonnegative. Instead of attacking the problem directly, we first concentrated on the unconstrained case and converted the optimization problem into an equivalent differential game. The game problem has a simpler pay-off function and we established the equivalence of this with the unconstrained dynamic asset allocation problem. We then imposed constraint on the equivalent differential game. Using some recent results, we solve this problem completely in this paper. The solutions are in the form of a number of theorems. Due to space limitations, we are unable to give proofs of the theorems. For simplicity, we also restricted ourselves to the case of one economic factor. Detailed proofs of the theorem, along with extension to multiple economic factors, will be published elsewhere.

6 References

1. T.R. Bielecki and S.R. Pliska, "Risk-sensitive dynamic asset pricing", Applied Math. and Optimization, 39, pp. 337-360, 1999.
2. M.K. Ghosh and K. Suresh Kumar, "A stochastic differential game in the quadrant", Applied Math. and Optimization (to appear).
3. R.C. Merton, "Optimum consumption and portfolio rules in a continuous time model", Journal of Economic Theory, 3, pp. 373-413, 1971.
4. D.H. Jacobson, "Optimal stochastic linear systems with exponential criteria and their relation to deterministic differential games", IEEE Transactions on Automatic Control, AC-18, pp. 124-131, 1973.
5. P. Whittle, Optimization over Time: Dynamic Programming and Stochastic Control, Vols. 1 & II, Wiley, New York, 1982, 1983.
6. A. Bensoussan and J.H. van Schuppen, "Optimal control of partially observable stochastic systems with an exponential-of-integral performance index", SIAM Journal on Control and Optimization, 23, pp. 599-613, 1985.
7. P. Dai Pra, L. Meneghini and W.J. Runggaldier, "Connection between stochastic control and dynamic games", MCSS: Mathematics of Control, Signals & Systems, 9, pp. 303-326, 1996.

8. K. Fan, "Fixed point and minimax theorems in locally convex topological linear spaces", Proceedings of the National Academy of Science, U.S.A., 38, pp. 121-126, 1952.
9. V.E. Benes, "Existence of optimal strategies based on specified information for a class of stochastic decision problems", SIAM Journal on Control and Optimization, 8, pp. 179-188, 1970.
10. H. Nagai, "Bellman equations of risk-sensitive control", SIAM Journal on Control and Optimization, 34, pp. 74-101, 1996.
11. N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam, 1981.
12. A. Bensoussan and J.L. Lions, Applications of Variational Inequalities in Stochastic Control, North-Holland, Amsterdam, 1982.

Intensity-Based Valuation of Basket Credit Derivatives [a]

Tomasz R. Bielecki
Department of Mathematics, The Northeastern Illinois University
5500 North St. Louis Avenue, Chicago, IL 60625-4699 USA
e-mail: T-Bielecki@neiu.edu

Marek Rutkowski
Faculty of Mathematics and Information Science, Politechnika Warszawska
pl. Politechniki 1, 00-661 Warszawa, Poland
e-mail: markrut@mini.pw.edu.pl

Modeling of credit events and related credit derivatives in terms of hazard processes of associated random times has gained much attention in the literature in the recent years. Such an approach to the subject was termed the *intensity-based approach*. *Basket credit derivatives* are financial derivatives products deriving their value from changes in credit quality of several underlying credit entities (credit names). We present here some recent results aiming at intensity-based valuation of basket credit derivatives within the context of so-called *conditionally independent defaults.*

1 Notation and Set-up

We shall consider a finite collection of random times $\tau_1, \ldots, \tau_n$ defined on a common probability space $(\Omega, \mathcal{G}, \mathbf{Q}^*)$. The random times $\tau_1, \ldots, \tau_n$ are assumed to represent default times of underlying credit entities $i = 1, \ldots, n$. Since we are only interested in the valuation of derivative securities, we shall interpret $\mathbf{Q}^*$ as a martingale probability measure.

Unless explicitly stated otherwise, we assume that $\mathbf{Q}^*\{\tau_k = 0\} = 0$ and $\mathbf{Q}^*\{\tau_k > t\} > 0$ for every $t \in \mathbf{R}_+$ and $k = 1, \ldots, n$. The case of simultaneous defaults is not examined here; namely, we postulate that $\mathbf{Q}^*\{\tau_k = \tau_j\} = 0$ for arbitrary $k, j = 1, \ldots, n$ with $k \neq j$. We associate with the collection $\tau_1, \ldots, \tau_n$ of default times the ordered sequence $\tau_{(1)} \leq \tau_{(2)} \leq \ldots \leq \tau_{(n)}$ of random times. By definition, $\tau_{(1)} = \min(\tau_1, \tau_2, \ldots, \tau_n)$, and

$$\tau_{(i+1)} = \min\left(\tau_k : k = 1, \ldots, n,\ \tau_k > \tau_{(i)}\right)$$

for $i = 1, \ldots, n-1$. In particular, $\tau_{(n)} = \max(\tau_1, \tau_2, \ldots, \tau_n)$.

[a] The research of the authors was partially supported by State Committee for Scientific Research (Komitet Badań Naukowych) Grant PBZ-KBN-016/P03/1999. The research of the first author was additionally supported by NSF Grant DMS-9971307.

In addition to the family $\tau_1, \dots, \tau_n$ of random times, we postulate that we are also given a reference filtration, $\mathbf{F}$ say, on the probability space $(\Omega, \mathcal{G}, \mathbf{Q}^*)$. We introduce the enlarged filtration $\mathbf{G}$ by setting $\mathbf{G} = \mathbf{F} \vee \mathbf{H}^1 \vee \mathbf{H}^2 \vee \dots \vee \mathbf{H}^n$. It will be also convenient to denote $\mathbf{H} = \mathbf{H}^1 \vee \mathbf{H}^2 \vee \dots \vee \mathbf{H}^n$.

For an in-depth study of processes related to default times, such as: hazard processes, martingale hazard process and intensity of a hazard process, we refer to Jeanblanc and Rutkowski [3] or Bielecki and Rutkowski [1].

Finally, let us stress that we shall be exploiting below the arbitrage pricing theory for the purpose of valuation of basket credit derivatives. This theory demonstrates that the discounted value of an attainable contingent claim can be computed as a (conditional) expectation of the discounted claim; the expectation is taken with respect to a martingale probability which is consistent with the chosen discount factor process. In this note, the money market account is chosen as the discount process, and thus the probability $\mathbf{Q}^*$ is the spot martingale probability. We refer to the monograph by Musiela and Rutkowski [8] for an in-depth study of the arbitrage pricing theory.

2 Basket Credit Derivatives

Our goal is to derive valuation formulae for the i^{th}-to-default contingent claims. We shall only present here the case corresponding to mutually conditionally independent default times. For a study of more general cases we refer to Bielecki and Rutkowski [1].

We shall consider a general i^{th}-to-default claim, $CCT^{(i)}$ say, which matures at time T and is specified by the following covenants:

- if $\tau_{(i)} = \tau_k \leq T$ for some $k = 1, \dots, n$, then the claim pays at time $\tau_{(i)}$ the amount $Z^k_{\tau_{(i)}}$, where Z^k is a $\mathbf{G}$-predictable process, and it pays at time T a $\mathcal{G}_T$-measurable amount X_k,

- if $\tau_{(i)} > T$, the claimholder receives at time T a $\mathcal{G}_T$-measurable amount X.

According to the convention above, if the i^{th} default occurs in the time interval $[0, T]$ – that is, if $\tau_{(i)} = \tau_k \leq T$ for some k – an immediate recovery cash flow $Z^k_{\tau_{(i)}}$ is received at time $\tau_{(i)}$, and a delayed recovery cash flow X_k is passed to the claim-holder at the maturity date T. A more general convention (not examined here) for payoffs associated with the claim $CCT^{(i)}$ would also involve immediate recovery payoffs at each default time $\tau_j < \tau_{(i)}$.

Example 2.1 Duffie [2] considers an example of a first-to-default type claim $CCT^{(1)}$ that is defined by setting $X_k = 0$ for $k = 1, \dots, n$. The corresponding last-to-default contract is the claim $CCT^{(n)}$ with $X_k = 0$ for $k = 1, \dots, n$.

Example 2.2 Kijima and Muromachi [6] examine a special case of a first-to-default type claim $CCT^{(1)}$, which is termed the *default swap of type F*. It is defined by setting $Z^k \equiv 0$ for $k = 1, \ldots, n$. Another contingent claim considered by Kijima and Muromachi [6] – the so-called *default swap of type D* – may be seen as an example of the second-to-default contingent claim. Formally, they deal with the claim $CCT^{(2)}$ with the following specific features. First, they set $Z_k \equiv 0$ for $k = 1, \ldots, n$. Second, they postulate that, for each $k = 1, \ldots, n$, the recovery payoff on the set $\{\tau_{(2)} = \tau_k \leq T\}$ equals:

$$X_k = \sum_{l \neq k} (\tilde{X}_k + \tilde{X}_l) \mathbf{1}_{\{\tau_{(1)} = \tau_l\}},$$

where $\tilde{X}_j$ is a $\mathcal{G}_T$-measurable random variable for each $j = 1, \ldots, n$. Finally, the recovery payoff on the set $\{\tau_{(2)} > T\}$ equals:

$$X = \hat{X}_0 \mathbf{1}_{\{\tau_{(1)} > T\}} + \sum_{j=1}^{n} \hat{X}_j \mathbf{1}_{\{\tau_{(1)} = \tau_j \leq T\}},$$

where $\hat{X}_j$ is a $\mathcal{G}_T$-measurable random variable for each $j = 0, \ldots, n$. In this general formulation, a default swap of type D protects its holder against the first two defaults, provided that they both have occurred before or at the maturity date of the contract.

Example 2.3 Li [7] examines still another example of the i^{th}-to-default claim, specifically, he sets $Z^k \equiv 1$, $X_k = 0$ for $k = 1, \ldots, n$ and $X = 0$. Such a contract is known as the *digital default put of basket type.*

3 Conditionally Independent Default Times

In the next section we shall derive valuation formulae for selected basket credit derivatives. This will be done under an additional assumption of conditional independence of default times with respect to the underlying filtration. This assumption underpins a vast majority of works devoted to the intensity-based valuation of basket derivatives (see, for example, Kijima [5] and Kijima and Muromachi [6]). In this section we provided some discussion of the concept of conditional independence of default times.

Before we proceed with a formal definition of conditional independence of default times, we provide the intuitive meaning of this assumption. Observe that all reference credit names are subject to common risk factors that may trigger credit (default) events. In addition, each credit name is also subject to the so-called *idiosyncratic risk* that is specific for this particular name, and may trigger credit (default) events associated with this credit name as well. At the intuitive level, the assumption of conditional independence of default times means that once the common risk factors are fixed, the idiosyncratic risk factors become independent of each other.

Definition 3.1 The random times τ_i, $i = 1, \ldots, n$ are said to be *conditionally independent* with respect to the filtration $\mathbf{F}$ under $\mathbf{Q}^*$ if and only if the following condition is satisfied: for any $T > 0$ and arbitrary $t_1, \ldots, t_n \in [0, T]$ we have:

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T\} = \prod_{i=1}^{n} \mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T\}.$$

Remarks. (i) Note that in general the conditional independence of random times does not imply their independence; the converse implication does not hold either. We find it convenient to also introduce a slightly more general formulation of the conditional independence property (see Definition 3.4).
(ii) It should be emphasized that the property of conditional independence may not be invariant under an equivalent change of probability measure. Thus, if the random times τ_i, $i = 1, \ldots, n$ are conditionally independent with respect to $\mathbf{F}$ under $\mathbf{Q}^*$, this does not imply that these random times are also conditionally independent with respect to $\mathbf{F}$ under an equivalent probability measure $\mathbf{Q}$.

Let us stress that the following equality does not necessarily hold for every $t_1, \ldots, t_n \in [0, T]$ and $u \in [0, T[$

$$\mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T\} = \mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_u\}.$$

However, the family of random times constructed in Example 3.1 below enjoys the above property (this feature is reflected in equality (3)). Since this property will be frequently used in what follows, we now introduce the following assumption standing for the rest of this section.

Condition (C.1) For every $T > 0$, $u \in [0, T]$, and $i = 1, \ldots, n$ we have:

$$\mathbf{Q}^*\{\tau_i > u \,|\, \mathcal{F}_T\} = \mathbf{Q}^*\{\tau_i > u \,|\, \mathcal{F}_u\}.$$

For any $t \in \mathbf{R}_+$, we write $F_t = \mathbf{Q}^*\{\tau \le t \,|\, \mathcal{F}_t\}$, and we denote by G the $\mathbf{F}$-*survival process* of τ with respect to the filtration $\mathbf{F}$, given as:

$$G_t := 1 - F_t = \mathbf{Q}^*\{\tau > t \,|\, \mathcal{F}_t\}.$$

Notice that for any $0 \le t \le s$ we have $\{\tau \le t\} \subseteq \{\tau \le s\}$, and thus:

$$\mathbf{E}_{\mathbf{Q}^*}(F_s \,|\, \mathcal{F}_t) = \mathbf{Q}^*\{\tau \le s \,|\, \mathcal{F}_t\} \ge \mathbf{Q}^*\{\tau \le t \,|\, \mathcal{F}_t\} = F_t.$$

This shows that the process F (G, resp.) follows a bounded, non-negative $\mathbf{F}$-submartingale ($\mathbf{F}$-supermartingale, resp.) under $\mathbf{Q}^*$. We may thus deal with the right-continuous modification of F (of G) with finite left-hand limits.

Definition 3.2 Assume that $F_t < 1$ for $t \in \mathbf{R}_+$. The $\mathbf{F}$-*hazard process* of τ under $\mathbf{Q}^*$, denoted by Γ, is defined through the formula $1 - F_t = e^{-\Gamma_t}$. Equivalently, $\Gamma_t = -\ln G_t = -\ln(1 - F_t)$ for every $t \in \mathbf{R}_+$.

If $\Gamma_t = \int_0^t \gamma_u \, du$, then the $\mathbf{F}$-progressively measurable, non-negative process γ is called the $\mathbf{F}$-intensity of τ. The next definition introduces a related concept of the $\mathbf{F}$-martingale hazard process of a random time.

Definition 3.3 An $\mathbf{F}$-predictable, right-continuous, increasing process Λ (with $\Lambda_0 = 0$) is called a $(\mathbf{F}, \mathbf{G})$-*martingale hazard process* under $\mathbf{Q}^*$ of τ if and only if the process $\tilde{M}_t := H_t - \Lambda_{t \wedge \tau}$ follows a $\mathbf{G}$-martingale under $\mathbf{Q}^*$.

If, in addition, $\Lambda_t = \int_0^t \lambda_u \, du$, then the $\mathbf{F}$-progressively measurable, non-negative process λ is referred to as the $(\mathbf{F}, \mathbf{G})$-martingale intensity process of τ under $\mathbf{Q}^*$.

Example 3.1 *Canonical construction of conditionally independent default times.* We shall now provide an explicit construction of a conditionally independent family of random times with pre-specified $\mathbf{F}$-hazard processes.

Let Γ^i, $i = 1, \ldots, n$ be a given collection of $\mathbf{F}$-adapted increasing continuous stochastic processes, defined on a common filtered probability space $(\hat{\Omega}, \mathbf{F}, \mathbf{P}^*)$. We assume that $\Gamma^i_0 = 0$ and $\Gamma^i_\infty = \infty$, for $i = 1, \ldots, n$ (clearly $\Gamma^i_t < \infty$ for every $t \in \mathbf{R}_+$).

Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$ be an auxiliary probability space, endowed with a sequence ξ_i, $i = 1, \ldots, n$ of mutually independent random variables uniformly distributed on the interval $[0, 1]$. We consider the product space $(\Omega, \mathcal{G}, \mathbf{Q}^*) = (\hat{\Omega} \times \tilde{\Omega}, \mathcal{F}_\infty \otimes \tilde{\mathcal{F}}, \mathbf{P}^* \otimes \tilde{\mathbf{P}})$, and for any $i = 1, \ldots, n$ we set:

$$\tau_i = \inf \{ t \in \mathbf{R}_+ : \Gamma^i_t \geq -\ln \xi_i \}. \tag{1}$$

It might be useful to observe that each random variable $\eta_i := -\ln \xi_i$ is exponentially distributed under $\mathbf{Q}^*$, with unit parameter. Thus,

$$\tau_i = \inf \{ t \in \mathbf{R}_+ : \Gamma^i_t \geq \eta_i \},$$

where η_i, $i = 1, \ldots, n$ is a family of mutually independent random variables with unit exponential law. It is natural to endow the product space $(\Omega, \mathcal{G}, \mathbf{Q}^*)$ with the enlarged filtration $\mathbf{G} = \mathbf{F} \vee \mathbf{H}^1 \vee \ldots \vee \mathbf{H}^n$. For each t, the σ-field $\mathcal{G}_t$ represents all information available to an agent at time t, including the observations of all random times τ_i, $i = 1, \ldots, n$. Formally,

$$\mathcal{G}_t = \mathcal{F}_t \vee \sigma\big(\{\tau_1 < t_1\}, \ldots, \{\tau_n < t_n\} : t_1 \leq t, \ldots, t_n \leq t\big).$$

Let us finally observe that the sequence of random times constructed above satisfies the desired property that the equality $\mathbf{Q}^*\{\tau_i = \tau_j\} = 0$ holds for every $i, j = 1, \ldots, n$ such that $i \neq j$. The next lemma summarizes the properties of random times constructed in Example 3.1

Lemma 3.1 *For a given family $\Gamma^i, \ldots, \Gamma^n$ of $\mathbf{F}$-adapted increasing continuous processes, let the random times $\tau_1, \ldots, \tau_n$ be defined as in Example 3.1.*
(i) *The conditional joint probability of survival satisfies, for $t_1, \ldots, t_n \in \mathbf{R}_+$,*

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_\infty\} = \prod_{i=1}^n e^{-\Gamma^i_{t_i}} = e^{-\sum_{i=1}^n \Gamma^i_{t_i}}. \tag{2}$$

(ii) *For arbitrary $t_1, \ldots, t_n \in \mathbf{R}_+$ and any $T \geq \max(t_1, \ldots, t_n)$ we have:*

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T\} = \prod_{i=1}^n e^{-\Gamma^i_{t_i}} = e^{-\sum_{i=1}^n \Gamma^i_{t_i}}. \tag{3}$$

(iii) *Random times $\tau_1, \ldots, \tau_n$ are conditionally independent with respect to the filtration $\mathbf{F}$ under $\mathbf{Q}^*$.*
(iv) *For each $i = 1, \ldots, n$, the process Γ^i represents the $\mathbf{F}$-hazard process and the $(\mathbf{F}, \mathbf{G})$-martingale hazard process of the random time τ_i. In other words, the equality $\Gamma^i = \Lambda^i$ is valid for every $i = 1, \ldots, n$.*

Proof. First, observe that $\{\tau_i > t\} = \{\Gamma^i_t < -\ln \xi_i\} = \{e^{-\Gamma^i_t} > \xi_i\}$. Let us take arbitrary numbers $t_1, \ldots, t_n \in \mathbf{R}_+$. Each random variable $\Gamma^i_{t_i}$ is obviously $\mathcal{F}_\infty$-measurable, and thus

$$\begin{aligned}
&\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_\infty\} \\
&= \mathbf{Q}^*\{e^{-\Gamma^1_{t_1}} > \xi_1, \ldots, e^{-\Gamma^n_{t_n}} > \xi_n \,|\, \mathcal{F}_\infty\} \\
&= \mathbf{Q}^*\{e^{-x_1} > \xi_1, \ldots, e^{-x_n} > \xi_n \,|\, \mathcal{F}_\infty\}_{x_1 = \Gamma^1_{t_1}, \ldots, x_n = \Gamma^n_{t_n}} \\
&= \prod_{i=1}^n \mathbf{Q}^*\{e^{-x_i} > \xi_1\}_{x_i = \Gamma^i_{t_i}} = \prod_{i=1}^n \tilde{\mathbf{P}}\{e^{-x_i} > \xi_i\}_{x_i = \Gamma^i_{t_i}} = \prod_{i=1}^n e^{-\Gamma^i_{t_i}}.
\end{aligned}$$

This proves part (i). Equality (3) is a simple consequence of (2). Indeed, since for any $T \geq t_i$, the random variable $\Gamma^i_{t_i}$ is $\mathcal{F}_T$-measurable, it is apparent that the following chain of equalities holds:

$$\begin{aligned}
\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T\} &= \mathbf{E}_{\mathbf{Q}^*}\big(\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_\infty\}\big|\mathcal{F}_T\big) \\
&= \mathbf{E}_{\mathbf{Q}^*}\big(e^{-\sum_{i=1}^n \Gamma^i_{t_i}} \,\big|\, \mathcal{F}_T\big) = e^{-\sum_{i=1}^n \Gamma^i_{t_i}}.
\end{aligned}$$

In particular, for any i and every $t_i \leq T$ we have:

$$\mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T\} = \mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_\infty\} = e^{-\Gamma^i_{t_i}}.$$

To establish the conditional independence of random times τ_i, $i = 1, \ldots, n$ with respect to the filtration $\mathbf{F}$, it is enough to observe that, by virtue of part (ii), for any fixed $T > 0$ and arbitrary $t_1, \ldots, t_n \leq T$ we have:

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T\} = \prod_{i=1}^{n} e^{-\Gamma^i_{t_i}} = \prod_{i=1}^{n} \mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T\}.$$

For the last statement, notice that from Lemma 8.2.2 in Bielecki and Rutkowski [1] we know that Γ^i represents the $\mathbf{F}$-hazard process of τ_i and the $(\mathbf{F}, \mathbf{G}^i)$-martingale hazard process of τ_i, where $\mathbf{G}^i := \mathbf{F} \vee \mathbf{H}^i$. This means that the process $\tilde{M}^i_t = H^i_t - \Gamma^i_{t\wedge\tau_i}$ is a $\mathbf{G}^i$-martingale. We need to show that $\tilde{M}^i$ is also a $\mathbf{G}$-martingale. The process $\tilde{M}^i$ is manifestly $\mathbf{G}$-adapted. It suffices to check that for any $t \leq s$ we have:

$$\mathbf{E}_{\mathbf{Q}^*}(H^i_s - \Gamma^i_{s\wedge\tau_i} \,|\, \mathcal{G}_t) = \mathbf{E}_{\mathbf{Q}^*}(H^i_s - \Gamma^i_{s\wedge\tau_i} \,|\, \mathcal{G}^i_t).$$

Notice that the σ-fields $\mathcal{G}^i_s$ and $\tilde{H}_t := \mathcal{H}^1_t \vee \mathcal{H}^{i-1}_t \vee \ldots \vee \mathcal{H}^{i+1}_t \vee \mathcal{H}^n_t$ are conditionally independent given $\mathcal{G}^i_t$. Consequently, we obtain:

$$\mathbf{E}_{\mathbf{Q}^*}(H^i_s - \Gamma^i_{s\wedge\tau_i} \,|\, \mathcal{G}_t) = \mathbf{E}_{\mathbf{Q}^*}(H^i_s - \Gamma^i_{s\wedge\tau_i} \,|\, \mathcal{G}^1_t \vee \tilde{\mathcal{H}}_t) = \mathbf{E}_{\mathbf{Q}^*}(H^i_s - \Gamma^i_{s\wedge\tau_i} \,|\, \mathcal{G}^i_t).$$

We conclude that Γ^i is the $(\mathbf{F}, \mathbf{G})$-martingale hazard process of τ_i. $\diamond$

Let us introduce a property, which is apparently stronger than the conditional independence of random times with respect to a given filtration $\mathbf{F}$.

Definition 3.4 The random times $\tau_1, \ldots, \tau_n$ are *dynamically conditionally independent* with respect to $\mathbf{F}$ under $\mathbf{Q}^*$ if and only if for any $0 \leq t < T$ and arbitrary $t_1, \ldots, t_n \in [t, T]$ we have:

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T \vee \mathcal{H}_t\} = \prod_{i=1}^{n} \mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T \vee \mathcal{H}_t\}.$$

or, equivalently (it is clear that $\mathcal{F}_T \vee \mathcal{H}_t = \mathcal{F}_T \vee \mathcal{G}_t$ for any $t \leq T$),

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T \vee \mathcal{G}_t\} = \prod_{i=1}^{n} \mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T \vee \mathcal{G}_t\}.$$

The following Lemma 3.2 is a slight generalization of Lemma 5.1.4 in Bielecki and Rutkowski [1]. Its proof is omitted.

Lemma 3.2 *Let $\tau_1, \ldots, \tau_n$ be defined on the probability space $(\Omega, \mathcal{G}, \mathbf{Q}^*)$. For any sub-$\sigma$-field $\mathcal{F}$ of $\mathcal{G}$, we denote $J = \mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F} \vee \mathcal{H}_t\}$, where $\mathcal{H}_t = \mathcal{H}_t^1 \vee \ldots \vee \mathcal{H}_t^n$. If $t_i \geq t$ for $i = 1, \ldots, n$, then we have:*

$$J = \mathbf{1}_{\{\tau_1 > t, \ldots, \tau_n > t\}} \frac{\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}\}}{\mathbf{Q}^*\{\tau_1 > t, \ldots, \tau_n > t \,|\, \mathcal{F}\}}.$$

Proposition 3.1 *The random times $\tau_1, \ldots, \tau_n$ are conditionally independent with respect to the filtration $\mathbf{F}$ under $\mathbf{Q}^*$ if and only if they are dynamically conditionally independent with respect to the filtration $\mathbf{F}$ under $\mathbf{Q}^*$.*

Proof. It is enough to show that the conditional independence implies the dynamical conditional independence. The conditional independence of $\tau_1, \ldots, \tau_n$ with respect to $\mathbf{F}$ is equivalent to the following property: for each $T > 0$ and for arbitrary Borel subsets $A_1, \ldots, A_n$ of the interval $[0, T]$ we have:

$$\mathbf{Q}^*\{\tau_1 \in A_1, \ldots, \tau_n \in A_n \,|\, \mathcal{F}_T\} = \prod_{i=1}^{n} \mathbf{Q}^*\{\tau_i \in A_i \,|\, \mathcal{F}_T\}.$$

It is clear that this implies that for any $t \leq T$, the σ-fields $\mathcal{H}_t^1, \ldots, \mathcal{H}_t^n$ are mutually conditionally independent given $\mathcal{F}_T$. For $t \leq t_i \leq T$ we have: $\mathcal{H}_t^i \subseteq \mathcal{H}_{t_i}^i$ and the σ-fields $\mathcal{H}_{t_i}^i$ and $\tilde{\mathcal{H}}_{t_i} := \mathcal{H}_{t_i}^1 \vee \ldots \vee \mathcal{H}_{t_i}^{i-1} \vee \mathcal{H}_{t_i}^{i+1} \vee \ldots \vee \mathcal{H}_{t_i}^n$ are conditionally independent given $\mathcal{F}_T$. This yields:

$$\mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T \vee \mathcal{H}_t\} = \mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T \vee \mathcal{H}_t^i\}.$$

Then, by virtue of Lemma 3.2 we obtain:

$$\mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T \vee \mathcal{H}_t^i\} = \mathbf{1}_{\{\tau_i > t\}} \frac{\mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T\}}{\mathbf{Q}^*\{\tau_i > t \,|\, \mathcal{F}_T\}}.$$

Let us denote $J = \mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T \vee \mathcal{H}_t\}$. Applying Lemma 3.2 with $\mathcal{F} = \mathcal{F}_T$, we find that:

$$J = \mathbf{1}_{\{\tau_1 > t, \ldots, \tau_n > t\}} \frac{\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T\}}{\mathbf{Q}^*\{\tau_1 > t, \ldots, \tau_n > t \,|\, \mathcal{F}_T\}}.$$

Consequently, using again the conditional independence of $\tau_1, \ldots, \tau_n$, we get:

$$J = \prod_{i=1}^{n} \mathbf{1}_{\{\tau_i > t\}} \frac{\mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T\}}{\mathbf{Q}^*\{\tau_i > t \,|\, \mathcal{F}_T\}} = \prod_{i=1}^{n} \mathbf{Q}^*\{\tau_i > t_i \,|\, \mathcal{F}_T \vee \mathcal{H}_t\}.$$

This ends the proof. $\diamondsuit$

We have the following, rather obvious, modification of Lemma 3.1.

Lemma 3.3 *For a family $\Gamma^i, \ldots, \Gamma^n$ of* $\mathbf{F}$*-adapted increasing continuous processes, define the random times $\tau_1, \ldots, \tau_n$ as in Example 3.1. Then:*
(i) *the random times τ_i, $i = 1, \ldots, n$ are dynamically conditionally independent with respect to the filtration* $\mathbf{F}$ *under* $\mathbf{Q}^*$,
(ii) *for every $t \geq 0$ and every $t_1, \ldots, t_n \in [t, \infty)$, the joint conditional probability of survival satisfies:*

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_\infty \vee \mathcal{G}_t\} = \mathbf{1}_{\{\tau_1 > t_1, \ldots, \tau_n > t_n\}}\, e^{\sum_{i=1}^n (\Gamma_t^i - \Gamma_{t_i}^i)},$$

(iii) *for arbitrary $T > t \geq 0$ and any $t_1, \ldots, t_n \in [t, T]$, we have:*

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T \vee \mathcal{G}_t\} = \mathbf{1}_{\{\tau_1 > t_1, \ldots, \tau_n > t_n\}}\, e^{\sum_{i=1}^n (\Gamma_t^i - \Gamma_{t_i}^i)}.$$

Proof. The proof is left to the reader. $\diamondsuit$

Let us now focus on the minimum of default times $\tau_1, \ldots, \tau_n$. If each hazard process Γ^i admits the $\mathbf{F}$-intensity γ^i, equality (3) becomes:

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T\} = \prod_{i=1}^n \exp\Big(-\int_0^{t_i} \gamma_u^i \, du\Big). \tag{4}$$

Let us now focus on the random time $\tau_{(1)} = \tau_1 \wedge \cdots \wedge \tau_n$. The $\mathbf{F}$-hazard process $\Gamma^{(1)}$ of this random time satisfies $\Gamma^{(1)} = \sum_{i=1}^n \Gamma^i$, since (3) implies that:

$$e^{-\Gamma_t^{(1)}} = \mathbf{Q}^*\{\tau_{(1)} > t \,|\, \mathcal{F}_t\} = \mathbf{Q}^*\{\tau_1 > t, \ldots, \tau_n > t \,|\, \mathcal{F}_t\} = e^{-\sum_{i=1}^n \Gamma_t^i}.$$

Thus, in view of result (5.11) in Bielecki and Rutkowski[1], for any $\mathcal{F}_s$-measurable random variable Y and any $t \leq s$, the following equality holds:

$$\mathbf{E}_{\mathbf{Q}^*}(\mathbf{1}_{\{\tau_{(1)} > s\}}\, Y \,|\, \mathcal{G}_t) = \mathbf{1}_{\{\tau_{(1)} > t\}}\, \mathbf{E}_{\mathbf{Q}^*}\big(Y e^{\Gamma_t^{(1)} - \Gamma_s^{(1)}} \,|\, \mathcal{F}_t\big). \tag{5}$$

Notice that for any $\mathcal{G}_s$-measurable random variable Y and any $t \leq s$ we have:

$$\mathbf{E}_{\mathbf{Q}^*}(\mathbf{1}_{\{\tau_{(1)} > s\}}\, Y \,|\, \mathcal{G}_t) = \mathbf{E}_{\mathbf{Q}^*}(\mathbf{1}_{\{\tau_{(1)} > s\}}\, Y \,|\, \tilde{\mathcal{G}}_t),$$

where $\tilde{\mathbf{G}}$ stands for the filtration associated with $\tau_{(1)}$; that is, $\tilde{\mathbf{G}} = \mathbf{F} \vee \mathbf{H}^{(1)}$, where the filtration $\mathbf{H}^{(1)}$ is generated by the process $H_t^{(1)} = \mathbf{1}_{\{\tau_{(1)} \leq t\}}$.

3.1 Case of signed intensities

Some authors (e.g., Kijima and Muromachi [6]) examine credit risk models in which the negative values of intensities of random times involved are not precluded. They rightly indicate that negative values of the intensity process clearly contradict the interpretation of the intensity as the conditional probability of survival over an infinitesimal time interval. Nevertheless, the construction of conditionally independent random times also goes through in this case. Let us analyze this issue in some detail.

Assume that we are given a collection Γ^i, $i = 1, \ldots, n$ of $\mathbf{F}$-adapted continuous stochastic processes, with $\Gamma_0^i = 0$, defined on a filtered probability space $(\hat{\Omega}, \mathbf{F}, \hat{\mathbf{P}})$. We introduce a finite family τ_i, $i = 1, \ldots, n$, of random times on the enlarged probability space $(\Omega, \mathcal{G}, \mathbf{Q}^*)$, through formula (1), that is:

$$\tau_i = \inf \{ t \in \mathbf{R}_+ : \Gamma_t^i \geq -\ln \xi_i \}.$$

The random times $\tau_1, \ldots, \tau_n$ possess most of the required properties, but in general the hazard processes of these random times do not coincide with processes Γ^i as the following result shows.

Lemma 3.4 *The random times τ_i, $i = 1, \ldots, n$ are conditionally independent with respect to $\mathbf{F}$ under $\mathbf{Q}^*$. In particular, for every $t_1, \ldots, t_n \leq T$ we have:*

$$\mathbf{Q}^*\{\tau_1 > t_1, \ldots, \tau_n > t_n \,|\, \mathcal{F}_T\} = \prod_{i=1}^{n} e^{-\tilde{\Gamma}^i_{t_i}} = e^{-\sum_{i=1}^n \tilde{\Gamma}^i_{t_i}}, \tag{6}$$

where $\tilde{\Gamma}$ is the increasing process associated with Γ^i, i.e., $\tilde{\Gamma}_t^i := \sup_{u \leq t} \Gamma_u^i$.

Proof. The proof goes along similar lines as the proof of Lemma 3.1. It is enough to observe that we have:

$$\{\tau_i > t\} = \{\tilde{\Gamma}_t^i < -\ln \xi_i\} = \{e^{-\tilde{\Gamma}_t^i} > \xi_i\}.$$

Notice that the inclusion $\{\tilde{\Gamma}_t^i < -\ln \xi_i\} \subseteq \{\Gamma_t^i < -\ln \xi_i\}$ is always true, but in general $\{\tilde{\Gamma}_t^i < -\ln \xi_i\} \neq \{\Gamma_t^i < -\ln \xi_i\}$. $\diamondsuit$

In view of the last result, if default times are obtained through the construction described in this section, the intensity of each default time τ_i becomes automatically zero on the (random) set $\{\gamma^i < 0\}$. To conclude, the 'true' intensity of τ_i equals $\tilde{\gamma}^i = \max(\gamma^i, 0)$.

4 Valuation of the i^{th}-to-Default Contract

Our goal in this section is to compute the initial price $S_0^{(i)}$ for the i^{th}-to-default claim $CCT^{(i)}$ under the assumption of conditional independence of default times. We assume throughout that processes Z^k, $k = 1, \ldots, n$ are $\mathbf{F}$-predictable, and random payoffs X_k, $k = 1, \ldots, n$ and X are $\mathcal{F}_T$-measurable. These assumptions make the subsequent results less universal than some results derived in the next chapter. For instance, the recovery payoffs that explicitly depend on the timing of previous defaults are formally excluded. However, if these restrictions were not imposed, we would not be able to profit from the postulated conditional independence of default times with respect to the reference filtration $\mathbf{F}$.

To derive a representation for the value process of a general i^{th}-to-default claim $CCT^{(i)}$, we need to introduce some auxiliary notation. Let $i, j \in \{1, \ldots, n\}$ be fixed. By $\Pi^{(i,j)}$ we denote the collection of specific partitions of the set $\{1, \ldots, n\}$. Namely, if $\pi \in \Pi^{(i,j)}$ then $\pi = \{\pi_-, \{j\}, \pi_+\}$, where $\pi_- = \{k_1, k_2, \ldots, k_{i-1}\}$, $\pi_+ = \{l_1, l_2, \ldots, l_{n-i}\}$, and

$$j \notin \pi_-, \ j \notin \pi_+, \ \pi_- \cap \pi_+ = \emptyset, \ \pi_- \cup \pi_+ \cup \{j\} = \{1, \ldots, n\}.$$

For a fixed $i \in \{1, \ldots, n\}$ and any $j \in \{1, \ldots, n\}$, the partition $\pi = \{\pi_-, \{j\}, \pi_+\}$ should be interpreted as follows: the index j is the index of the i^{th} defaulting entity. The set π_- contains indices of all the names that default prior to the default of the j^{th} entity. Finally, the set π_+ includes all indices corresponding to the entities whose defaults occur after the default of the j^{th} entity.

Example 4.1 In this example, we consider $n = 2$ credit entities. For $i = 1$ (i.e., in the case of the first-to-default claim) and $j = 1, 2$ we have:

$$\Pi^{(1,1)} = \Big\{\{\emptyset, \{1\}, \{2\}\}\Big\}, \quad \Pi^{(1,2)} = \Big\{\{\emptyset, \{2\}, \{1\}\}\Big\}.$$

Similarly, in the case of the second-to-default claim, we have:

$$\Pi^{(2,1)} = \Big\{\{\{2\}, \{1\}, \emptyset\}\Big\}, \quad \Pi^{(2,2)} = \Big\{\{\{1\}, \{2\}, \emptyset\}\Big\}.$$

In this example, each set $\Pi^{(i,j)}$ contains only one partition; for example, the only element of $\Pi^{(1,1)}$ is the partition $\pi = \{\emptyset, \{1\}, \{2\}\}$.

Example 4.2 Let us now consider the case of $n = 4$. Let us take, for instance, $j = 3$. Then $\Pi^{(1,3)} = \Big\{\{\emptyset, \{3\}, \{1, 2, 4\}\}\Big\}$,

$$\Pi^{(2,3)} = \Big\{\{\{1\}, \{3\}, \{2, 4\}\}, \{\{2\}, \{3\}, \{1, 4\}\}, \{\{4\}, \{3\}, \{1, 2\}\}\Big\},$$

$$\Pi^{(3,3)} = \Big\{\{\{1, 2\}, \{3\}, \{4\}\}, \{\{1, 4\}, \{3\}, \{2\}\}, \{\{2, 4\}, \{3\}, \{1\}\}\Big\},$$

and finally, $\Pi^{(4,3)} = \Big\{\{\{1, 2, 4\}, \{3\}, \emptyset\}\Big\}$.

For any numbers $i, j \in \{1, \ldots, n\}$ and arbitrary $\pi \in \Pi^{(i,j)}$, we write $\tau(\pi_-) = \max\{\tau_k : k \in \pi_-\}$ and $\tau(\pi_+) = \min\{\tau_l : l \in \pi_+\}$, where we set by convention: $\max\emptyset = -\infty$ and $\min\emptyset = \infty$. It is clear that $\tau(\pi_-)$ ($\tau(\pi_+)$, resp.) is the default time that immediately precedes (follows, resp.) the time of the i^{th} default.

We shall first examine the general case, and subsequently the special case of non-random recovery payoffs and hazard processes. Let $B_t = \exp(\int_0^t r_u\,du)$ be the process modeling the savings account. It is not difficult to check that the initial price of the i^{th}-to-default payoff satisfies (cf. Proposition 10.2.3 in Bielecki and Rutkowski[1]):

$$\begin{aligned}
S_0^{(i)} &= \mathbf{E}_{\mathbf{Q}^*}\Big(\sum_{j=1}^n B_{\tau_j}^{-1} Z_{\tau_j}^j \sum_{\pi\in\Pi^{(i,j)}} \mathbf{1}_{\{\tau(\pi_-)<\tau_j<\tau(\pi_+),\,\tau_j\le T\}}\Big)\\
&+ \mathbf{E}_{\mathbf{Q}^*}\Big(B_T^{-1}\sum_{j=1}^n X_j \sum_{\pi\in\Pi^{(i,j)}} \mathbf{1}_{\{\tau(\pi_-)<\tau_j<\tau(\pi_+),\,\tau_j\le T\}}\Big)\\
&+ \mathbf{E}_{\mathbf{Q}^*}\Big(B_T^{-1}X\sum_{j=1}^n \sum_{\pi\in\Pi^{(i,j)}} \mathbf{1}_{\{\tau(\pi_-)<\tau_j<\tau(\pi_+),\,\tau_j> T\}}\Big)\\
&=: J_1 + J_2 + J_3.
\end{aligned}$$

Since $B_0 = 1$, we shall frequently omit B_0 from the formulae. In view of the assumed conditional independence of random times $\tau_1,\dots,\tau_n$, for the first term we obtain:

$$\begin{aligned}
J_1 &= \mathbf{E}_{\mathbf{Q}^*}\Big\{\mathbf{E}_{\mathbf{Q}^*}\Big(\sum_{j=1}^n Z_{\tau_j}^j B_{\tau_j}^{-1} \sum_{\pi\in\Pi^{(i,j)}} \mathbf{1}_{\{\tau(\pi_-)<\tau_j<\tau(\pi_+),\,\tau_j\le T\}}\,\Big|\,\mathcal{F}_T\Big)\Big\}\\
&= \mathbf{E}_{\mathbf{Q}^*}\Big\{\sum_{j=1}^n\int_0^T Z_u^j B_u^{-1}\Big(\sum_{\pi\in\Pi^{(i,j)}}\Big[\prod_{k\in\pi_-}\mathbf{Q}^*\{\tau_k<u\,|\,\mathcal{F}_T\}\Big]\\
&\qquad\times\Big[\prod_{l\in\pi_+}\mathbf{Q}^*\{\tau_l>u\,|\,\mathcal{F}_T\}\Big]\Big)\,d\mathbf{Q}^*\{\tau_j\le u\,|\,\mathcal{F}_T\}\Big\}\\
&= \mathbf{E}_{\mathbf{Q}^*}\Big\{\sum_{j=1}^n\int_0^T Z_u^j e^{-\int_0^u r_s\,ds}\Big(\sum_{\pi\in\Pi^{(i,j)}}\Big[\prod_{k\in\pi_-}\big(1-e^{-\Gamma_u^k}\big)\Big]\\
&\qquad\times\Big[e^{-\sum_{l\in\pi_+}\Gamma_u^l}\Big]\Big)\gamma_u^j e^{-\Gamma_u^j}\,du\Big\}.
\end{aligned} \tag{7}$$

For J_2, we have:

$$J_2 = \mathbf{E}_{\mathbf{Q}^*}\Big\{\mathbf{E}_{\mathbf{Q}^*}\Big(B_T^{-1}\sum_{j=1}^n X_j \sum_{\pi\in\Pi^{(i,j)}} \mathbf{1}_{\{\tau(\pi_-)<\tau_j<\tau(\pi_+),\,\tau_j\le T\}}\,\Big|\,\mathcal{F}_T\Big)\Big\}$$

$$= \mathbf{E}_{\mathbf{Q}^*}\Big\{B_T^{-1}\sum_{j=1}^{n} X_j \int_0^T \Big(\sum_{\pi\in\Pi^{(i,j)}} \Big[\prod_{k\in\pi_-} \mathbf{Q}^*\{\tau_k < u \,|\, \mathcal{F}_T\}\Big]$$

$$\times\Big[\prod_{l\in\pi_+} \mathbf{Q}^*\{\tau_l > u \,|\, \mathcal{F}_T\}\Big]\Big)\, d\mathbf{Q}^*\{\tau_j \le u \,|\, \mathcal{F}_T\}\Big\}$$

$$= \mathbf{E}_{\mathbf{Q}^*}\Big\{e^{-\int_0^T r_s ds}\sum_{j=1}^{n} X_j \int_0^T \Big(\sum_{\pi\in\Pi^{(i,j)}} \Big[\prod_{k\in\pi_-} (1-e^{-\Gamma_u^k})\Big]$$

$$\times\Big[e^{-\sum_{l\in\pi_+}\Gamma_u^l}\Big]\Big)\,\gamma_u^j e^{-\Gamma_u^j}\, du\Big\},$$

and the last term satisfies:

$$J_3 = \mathbf{E}_{\mathbf{Q}^*}\Big\{\mathbf{E}_{\mathbf{Q}^*}\Big(XB_T^{-1}\sum_{j=1}^{n}\sum_{\pi\in\Pi^{(i,j)}} \mathbf{1}_{\{\tau(\pi_-)<\tau_j<\tau(\pi_+),\,\tau_j>T\}}\,\Big|\,\mathcal{F}_\infty\Big)\Big\}$$

$$= \mathbf{E}_{\mathbf{Q}^*}\Big\{XB_T^{-1}\sum_{j=1}^{n} \int_T^\infty \Big(\sum_{\pi\in\Pi^{(i,j)}} \Big[\prod_{k\in\pi_-} \mathbf{Q}^*\{\tau_k < u \,|\, \mathcal{F}_\infty\}\Big]$$

$$\times\Big[\prod_{l\in\pi_+} \mathbf{Q}^*\{\tau_l > u \,|\, \mathcal{F}_\infty\}\Big]\Big)\, d\mathbf{Q}^*\{\tau_j \le u \,|\, \mathcal{F}_\infty\}\Big\}$$

$$= \mathbf{E}_{\mathbf{Q}^*}\Big\{Xe^{-\int_0^T r_s ds}\sum_{j=1}^{n} \int_T^\infty \Big(\sum_{\pi\in\Pi^{(i,j)}} \Big[\prod_{k\in\pi_-} (1-e^{-\Gamma_u^k})\Big]$$

$$\times\Big[e^{-\sum_{l\in\pi_+}\Gamma_u^l}\Big]\Big)\,\gamma_u^j e^{-\Gamma_u^j}\, du\Big\}.$$

In case of $i = 1$, the above result agrees with the more abstract expression established in Proposition 10.2.4 in Bielecki and Rutkowski [1].

Let us assume that $Z_t^j = z^j(t)$, where $z^j : [0,T] \to \mathbf{R}$, $i = j,\dots,n$ are deterministic (integrable) functions of time. In addition, we assume that the terminal payoffs $X_j = x_j$ and $X = x$, where x_j, $j = 1,\dots,n$ and x are constants. Finally, the intensities of default times $\gamma_t^j = \gamma^j(t)$, $j = 1,\dots,n$ are assumed to be non-random. Though these assumptions are rather stringent, they are nevertheless widely common in literature, since they lead to a simple result for the values of various kinds of i^{th}-to-default claims. Let us stress that the interest rates are not assumed to be non-random here.

Proposition 4.1 *Let $B(0,T)$ stand for the price of a default-free zero-coupon bond maturing at time T. Assume that the default times τ_j, $j = 1,\dots,n$ are conditionally independent with respect to the filtration $\mathbf{F}$, with deterministic intensities γ^j. The price of the i^{th}-to-default claim with deterministic recovery* *payoffs at time* $t = 0$ *equals:*

$$S_0^{(i)} = \sum_{j=1}^{n} \int_0^T B(0,u) z^j(u) g_{ij}(u) \gamma^j(u) e^{-\int_0^u \gamma^j(s)ds} \, du$$

$$+ B(0,T) \sum_{j=1}^{n} x_j \int_0^T g_{ij}(u) \gamma^j(u) e^{-\int_0^u \gamma^j(s)ds} \, du$$

$$+ B(0,T) x \sum_{j=1}^{n} \int_T^\infty g_{ij}(u) \gamma^j(u) e^{-\int_0^u \gamma^j(s)ds} \, du,$$

where, for every $u \in \mathbf{R}_+$, *we set:*

$$g_{ij}(u) := \sum_{\pi \in \Pi^{(i,j)}} e^{-\sum_{l \in \pi_+} \int_0^u \gamma^l(s)ds} \prod_{k \in \pi_-} \Big(1 - e^{-\int_0^u \gamma^k(s)ds} \Big).$$

Proof. It is enough to recall that $B(0,t) = \mathbf{E}_{\mathbf{Q}^*}(B_t^{-1})$ for $t \in [0,T]$. $\diamondsuit$

We shall now examine some particular cases of i^{th}-to-default claims. We maintain the assumption that the default times are $\mathbf{F}$-conditionally independent, but we do not postulate that their $\mathbf{F}$-intensities are deterministic.

Let us find the initial price, $S_0^{(F)}$ say, of a default swap of type F, which is an example of the first-to-default contract. To this end, we first consider a general first-to-default claim $CCT^{(1)}$. Using previously established formulae (or by direct calculations), we find that $S_0^{(1)} = J_1 + J_2 + J_3$, where the terms J_1, J_2, J_3 can be evaluated as follows. First, the term J_1 – associated with the recovery payoff $Z_{\tau_j}^j$ at default time $\tau_j \le T$ when the j^{th} reference entity is the first-to-default – is given by the formula:

$$J_1 = B_0 \, \mathbf{E}_{\mathbf{Q}^*} \Big(\sum_{j=1}^{n} B_{\tau_j}^{-1} Z_{\tau_j}^j \mathbf{1}_{\{\tau_j = \tau_{(1)}, \, \tau_j \le T\}} \,\Big|\, \mathcal{G}_0 \Big)$$

$$= \mathbf{E}_{\mathbf{Q}^*} \Big(\sum_{j=1}^{n} \int_0^T Z_u^j e^{-\int_0^u r_s ds} e^{-\sum_{l \ne j} \Gamma_u^l} \gamma_u^j e^{-\Gamma_u^j} \, du \Big).$$

Then, the term J_2 – corresponding to the random payoff X_j at maturity T when the j^{th} reference entity is the first-to-default – satisfies:

$$J_2 = B_0 \, \mathbf{E}_{\mathbf{Q}^*} \Big(B_T^{-1} \sum_{j=1}^{n} X_j \mathbf{1}_{\{\tau_j = \tau_{(1)}, \, \tau_j \le T\}} \,\Big|\, \mathcal{G}_0 \Big)$$

$$= \mathbf{E}_{\mathbf{Q}^*} \Big(e^{-\int_0^T r_s ds} \sum_{j=1}^{n} \int_0^T X_j e^{-\sum_{l \ne j} \Gamma_u^l} \gamma_u^j e^{-\Gamma_u^j} \, du \Big).$$

Finally, the last term – associated with the payoff X at maturity T in case there was no default prior to T – is given by the formula:

$$J_3 = B_0\,\mathbf{E}_{\mathbf{Q}^*}\big(XB_T^{-1}\mathbb{1}_{\{\tau_{(1)}>T\}}\,\big|\,\mathcal{G}_0\big) = \mathbf{E}_{\mathbf{Q}^*}\Big(XB_T^{-1}\sum_{j=1}^{n}\mathbb{1}_{\{\tau_j=\tau_{(1)},\,\tau_j>T\}}\Big)$$
$$= \mathbf{E}_{\mathbf{Q}^*}\Big(Xe^{-\int_0^T r_s\,ds}\sum_{j=1}^{n}\int_T^{\infty} e^{-\sum_{l\neq j}\Gamma^l_u}\gamma^j_u e^{-\Gamma^j_u}\,du\Big).$$

Under an additional assumption of constant recovery payoffs we have:

$$J_2 = \hat{x}\,\mathbf{E}_{\mathbf{Q}^*}\big(B_T^{-1}\,\mathbb{1}_{\{\tau_{(1)}\le T\}}\big) = \hat{x}B(0,T) - \hat{x}B_0\,\mathbf{E}_{\mathbf{Q}^*}\big(B_T^{-1}\mathbb{1}_{\{\tau_{(1)}>T\}}\big),$$

and

$$J_3 = xB_0\,\mathbf{E}_{\mathbf{Q}^*}\big(B_T^{-1}\mathbb{1}_{\{\tau_{(1)}>T\}}\big),$$

provided that $X_j = \hat{x}$ for $j = 1,\dots,n$, and $X = x$, where $\hat{x}$ and x are real numbers. In the case of the default swap of type F, we clearly have $Z^j \equiv 0$ for every j, and thus the following result – originally due to Kijima and Muromachi [6] – is valid.

Proposition 4.2 *The value at time $t = 0$ of a default swap of type F, with the constant payoffs $\hat{x}$ and x, equals:*

$$S_0^{(F)} = \hat{x}B(0,T) + (x-\hat{x})B_0\,\mathbf{E}_{\mathbf{Q}^*}\big(B_T^{-1}e^{-\sum_{j=1}^{n}\Gamma^j_T}\big). \tag{8}$$

Proof. It is enough to observe that:

$$S_0^{(F)} = J_2 + J_3 = \hat{x}B(0,T) + (x-\hat{x})B_0\,\mathbf{E}_{\mathbf{Q}^*}\big(B_T^{-1}\mathbb{1}_{\{\tau_{(1)}>T\}}\big),$$

and to apply the conditioning with respect to the σ-field $\mathcal{F}_T$. $\diamond$

Let us notice that a similar representation can be derived in the case of a default swap of type D (see Bielecki and Rutkowski [1]).

1. T.R. Bielecki and M. Rutkowski, *Credit Risk: Modeling, Valuation and Hedging* (Springer-Verlag, Berlin, 2002).
2. D. Duffie, First-to-default valuation, preprint, Stanford University, 1998.
3. M. Jeanblanc and M. Rutkowski, Modelling of default risk: mathematical tools, preprint, Université d'Évry and Warsaw University of Technology, 2000.
4. M. Jeanblanc and M. Rutkowski, Default risk and hazard process, in: *Mathematical Finance – Bachelier Congress 2000,* eds. H. Geman, D. Madan, S.R. Pliska and T. Vorst (Springer-Verlag, Berlin, 2001).

5. M. Kijima, Valuation of a credit swap of the basket type, *Review of Derivatives Research* **4**, 81-97 (2000).
6. M. Kijima and Y. Muromachi, Credit events and the valuation credit derivatives of basket type, *Review of Derivatives Research* 4, 55-79 (2000).
7. D.X. Li, The valuation of the ith-to-default basket credit derivatives, preprint, RiskMetrics Group, 1999.
8. M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modelling* (Springer-Verlag, Berlin, 1997).

Comonotonicity of Backward Stochastic Differential Equations

Zengjing Chen and Xiangrong Wang
Department of Mathematics
Shandong University
Jinan, 250100, P.R. of China

Abstract

Pardoux and Peng introduced a class of nonlinear backward stochastic differential equations (shortly BSDEs) in 1990, according to Pardoux and Peng's theorem, the solution of this kind of BSDEs consists of a pair of adapted processes, say (y, z). Since then, many researchers have been exploring the properties of this pair solution (y, z), especially the properties of part y. In this paper, we shall explore the properties of z. We give a comonotonic theorem for part z.

Keywords: Backward stochastic differential equation (BSDE), Comonotonicity, Partial differential equation(PDE).

1 Introduction

Fixed time horizon $T > 0$, under some suitable assumptions on ξ and g, Pardoux and Peng (1990) showed that there exists a pair adapted process, say (y, z), satisfying backward stochastic differential equation (in short BSDE):

$$y_t = \xi + \int_t^T g(y_s, z_s, s)ds - \int_t^T < z_s, dW_s >, \ 0 \leq t \leq T. \qquad (*)$$

Here and next $< \cdot, \cdot >$ is the inner product in R^d.

Since then, many researchers have been working on this subject and related properties of the solutions of BSDEs, due to the connection of this subject with mathematical finance, stochastic control, partial differential equation, stochastic game and stochastic geometry and mathematical economics. Among these results is the comparison theorem of BSDEs with respect to y (first introduced by Peng (1991)). Such a comparison theorem, as said by El Karoui in [K, p15],"plays the same role that the maximum principle in the theory of partial differential equation".

An interesting question is how to compare part z of the solution (y, z) of BSDE(*) ? In fact, because z in BSDE(*) is a volatility, it is not easy to compare z in the same way as to compare y.

In this paper, we try to explore the comonotonicity of z. That is, let (y^ξ, z^ξ) and (y^η, z^η) be the solutions of BSDE (*) corresponding to terminal value ξ and η, respectively, we shall give a sufficient condition on ξ and η under which

$$z_t^\xi \odot z_t^\eta \geq 0, \quad \text{a.e.} \quad t \in [0, T].$$

Here and next, for any $z, x \in R^d$, $z \odot x$ is denoted by $z \odot x := (z_1x_1, z_2x_2, \cdots, z_dx_d)$ and $z \odot x \geq 0$ means $z_ix_i \geq 0$ for each i, z_i and x_i are the i-th components of z and x, $i = 1, 2, \cdots, d$.

2 BSDE and Related Properties

In this section, we shall present some notations and recall briefly some basic lemmas we use in this paper.

Fix $T \in [0, \infty)$, let $(W_t)_{0 \leq t \leq T}$ be a d-dimensional standard Brownian motion defined on a completed probability space $(\Omega, \mathcal{F}, P)$ and $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ be the natural filtration generated by Brownian motion $(W_t)_{0 \leq t \leq T}$, i.e.

$$\mathcal{F}_t = \sigma\{W_s;\ s \leq t\}.$$

We assume $\mathcal{F} = \mathcal{F}_T$.

$L^2(0,T) := \left\{X :\ X_t \text{ is } \mathcal{F}_t\text{-adapted process with } \|X\|_L^2 := E\int_0^T |X_s|^2 ds < \infty\right\}$;
$L^2(\Omega, \mathcal{F}, P) := \{\xi :\ \xi \text{ is } \mathcal{F}\text{-measurable random variable such that } E|\xi|^2 < \infty\}$.

We say function $g :\ R \times R^d \times [0, T] \to R$ satisfy (H1) and $(H2)$ if the following conditions hold:

(H1) For any$(y, z) \in R \times R^d$, $\int_0^T |g(y, z, s)|^2 ds < \infty$;
(H2) g satisfies uniform Lipschitz condition, i.e. there exists $\mu > 0$ such that,

$$|g(y_1, z_1, t) - g(y_2, z_2, t)| \leq \mu(|y_1 - y_2| + |z_1 - z_2|),$$
$$(y_i, z_i) \in R^{1+d}, t \geq 0, i = 1, 2.$$

The next lemma is a special case of Pardoux and Peng's theorem:

Lemma 1 **(Pardoux & Peng,1990)** *Suppose that g satisfies (H1) and (H2), $\xi \in L^2(\Omega, \mathcal{F}, P)$, then there exists a unique pair of adapted processes $(y, z) \in L^2(0,T) \times L^2(0,T)$ satisfying BSDE:*

$$y_t = \xi + \int_t^T g(y_s, z_s, s) ds - \int_t^T < z_s, dW_s > . \tag{1}$$

The following Lemma can be found in [P1, P2].

Lemma 2 *Suppose that g_1 and g_2 satisfy (H1) and (H2), let $\xi_1, \xi_2 \in L^2(\Omega, \mathcal{F}, P)$ and (y^i, z^i) $(i = 1, 2)$ be the solutions of BSDE (1) corresponding to $\xi = \xi_1$, $g = g_1$ and $\xi = \xi_2$, $g = g_2$, respectively. Then there exists a constant $c > 0$ such that*

$$E[\sup_{0\le s\le T} |y_s^1 - y_s^2|^2 + \int_0^T |z_s^1 - z_s^2|^2 ds] \le cE[|\xi_1 - \xi_2|^2 + (\int_0^T |\overline{g}_s| ds)^2],$$

where $\overline{g}_s := g_1(y_s^1, z_s^1, s) - g_2(y_s^1, z_s^1, s)$.

Remark 1 *The lemma implies that if ξ_2 converges to ξ_1 in $L^2(\Omega, \mathcal{F}, P)$ and $g_1(y^1, z^1, \cdot)$ converges to $g_2(y^1, z^1, \cdot)$ in $L^2(0, T)$, then (y^2, z^2) converges to (y^1, z^1)in $L^2(0, T) \times L^2(0, T)$.*

Assumption A. Let $b(t, x) : [0, T] \times R \to R, \sigma(t, x) : [0, T] \times R \to R^{1\times d}$ be continuous in (t, x) and uniformly Lipschitz continuous in $x \in R$.

By the existence theorem of stochastic differential equation (shortly SDE), there exists a unique strong solution $\{X_s^{t,x}\}$ satisfying SDE:

$$\begin{cases} dX_s = b(s, X_s)ds + \sigma(s, X_s)dW_s, \\ X_t = x, \ s \in [t, T]. \end{cases} \tag{2}$$

Let $\Phi(x)$ be a continuous function defined on R such that $\Phi(X_T^{t,x}) \in L^2(\Omega, \mathcal{F}, P)$ and $(y^{t,x}, z^{t,x})$ be the solution of BSDE:

$$y_s = \Phi(X_T^{t,x}) + \int_s^T g(y_r, z_r, r)dr - \int_s^T < z_r, dW_r >, \quad s \in [0, T]. \tag{3}$$

The following Lemma can be found in [P3, MPY, YZ]:

Lemma 3 *All the functions b, σ, Φ, g are smooth (which are assumed to be C^3) with bounded derivatives, let $(y_s^{t,x}, z_s^{t,x})$ be the solution of BSDE (3), then*

(i) $u(t, x) := y_t^{t,x} \in C^{1,2}([0, T] \times R, R)$ is the unique solution of the following partial differential equation (shortly PDE):

$$\begin{cases} \partial_t u(t, x) + \mathcal{L}u(t, x) + g(t, u(t, x), \sigma^*(t, x)\partial_x u(t, x)) = 0, \\ u(T, x) = \Phi(x). \end{cases} \tag{4}$$

where $\mathcal{L}u(t, x) := \frac{1}{2}\sigma(t, x)\sigma^*(t, x)\partial_x^2 u(t, x) + b(t, x)\partial_x u(t, x)$.

(ii) $z_s^{t,x} = \sigma^*(s, X_s^{t,x})\partial_x u(s, X_s^{t,x})$, a.e. $s \in [t, T]$, where $\partial_x u$ is the partial derivative of $u(t, x)$ with respect to x.

Remark 2 *In particular, in (2) and (3), if let $t = 0$, then $(y_s^{0,x}, z_s^{0.x})$, the solution of BSDE (3) with $t = 0$, is the solution of BSDE (3) with terminal value $\Phi(X_T^{0,x})$, we will apply this property in next section.*

3 Comonotonic theorem

In this section, we shall give a comonotonic theorem of BSDEs.

For any $\xi, \eta \in L^2(\Omega, \mathcal{F}, P)$, let (y^ξ, z^ξ) and (y^η, z^η) be the solutions of BSDEs:

$$y_t^\xi = \xi + \int_t^T g_1(y_s^\xi, z_s^\xi, s)ds - \int_t^T < z_s^\xi, dW_s > \tag{5}$$

and

$$y_t^\eta = \eta + \int_t^T g_2(y_s^\eta, z_s^\eta, s)ds - \int_t^T < z_s^\eta, dW_s >, \tag{6}$$

where g_1 and g_2 satisfies assumption (H1) and (H2).

For the purpose of this paper, let us consider random variables ξ and η which satisfy that there exist two functions Φ and Ψ such that ξ and η are of the forms

$$\xi = \Phi(X_T), \quad \eta = \Psi(Y_T),$$

where X_T and Y_T are the values of $\{X_t\}$ and $\{Y_t\}$, the solutions of the following SDEs, at time T, respectively:

$$\begin{cases} dX_s = b_1(s, X_s)ds + \sigma_1(s, X_s)dW_s, 0 \le s \le T, \\ X_0 = x, \ x \in R, \end{cases} \tag{7}$$

and

$$\begin{cases} dY_s = b_2(s, Y_s)ds + \sigma_2(s, Y_s)dW_s, 0 \le s \le T, \\ Y_0 = y, y \in R, \end{cases} \tag{8}$$

where b_i, σ_i $(i = 1, 2)$ satisfy Assumption A.

We now begin to find the conditions on Φ, Ψ, σ_1 and σ_2 under which z^ξ and z^η satisfy

$$z_t^\xi \odot z_t^\eta \ge 0, \text{ a.e. } t \in [0, T]. \tag{9}$$

Let us introduce the following definition.

Definition 1 *The functions Φ and Ψ are called comonotonic, if both Φ and Ψ are of the same monotonicity. That is, if Φ is an increasing (decreasing) function, so is Ψ. Furthermore, if Φ and Ψ are strictly monotonic, Φ and Ψ are called strictly comonotonic.*

Let us first observe an example which shows that if Φ and Ψ are not comonotonic, then inequality (9) is not true.

Example 1 *Suppose that $\{W_t\}$ is 1-dimensional Brownian motion, let $\xi = (W_T)^2$ and $\eta = W_T$, let us consider the BSDEs:*

$$y_t = (W_T)^2 - \int_t^T ds - \int_t^T < z_s, dW_s >$$

and

$$\overline{y}_t = W_T + \int_t^T \overline{z}_s ds - \int_t^T < \overline{z}_s, dW_s > .$$

Solving the above BSDEs, it follows $(y_t, z_t) = (W_t^2, 2W_t)$ *and*

$$(\overline{y}_t, \overline{z}_t) = (W_t + T - t, 1),$$

thus

$$z_t \odot \overline{z}_t = z_t \overline{z}_t = 2W_t,$$

which does not satisfy (9), the main reason is that x^2 *and* x *are not comonotonic.*

However, we have the following theorem:

Theorem 1 *Suppose that* (y^ξ, z^ξ) *and* (y^η, z^η) *are the solutions of BSDE(5) and (6) with terminal values* $\xi = \Phi(X_T)$ *and* $\eta = \Psi(Y_T)$ *such that* $\xi, \eta \in L^2(\Omega, \mathcal{F}, P)$, *where* (X_t) *and* (Y_t) *are the solutions of SDE(7) and (8) and* Φ *and* Ψ *are two continuous functions.*

(i) If Φ *and* Ψ *are comonotonic and* $\sigma_1(t, X_t) \odot \sigma_2(t, Y_t) \geq 0$, *then*

$$z_t^\xi \odot z_t^\eta \geq 0, \qquad a.e.\ t \in [0, T];$$

(ii) Furthermore, if Φ *and* Ψ *are strictly comonotonic, then if and only if*

$$\sigma_1(t, X_t) \odot \sigma_2(t, Y_t) > 0, \qquad a.e.\ t \in [0, T]$$

that we have

$$z_t^\xi \odot z_t^\eta > 0, \qquad a.e.\ t \in [0, T].$$

Proof. We prove this theorem into two steps:

Step 1: We assume that $\Phi, \Psi, b_i, \sigma_i, g_i (i = 1, 2)$ are smooth to be C^3.
Let $\{X_s^{t,x}\}$ be the solution of SDE:

$$\begin{cases} dX_s^{t,x} = b_1(s, X_s^{t,x})dt + \sigma_1(s, X_s^{t,x})dW_s, \\ X_t = x, \quad s \in [t, T]. \end{cases}$$

and $\{Y_s^{t,y}\}$ be the solution of SDE:

$$\begin{cases} dY_s^{t,y} = b_2(s, Y_s^{t,y})dt + \sigma_2(s, Y_s^{t,y})dW_s, \\ Y_t = y, \quad s \in [t, T]. \end{cases}$$

Furthermore, let $(y_s^{t,x,\Phi}, z_s^{t,x,\Phi})$ and $(y_s^{t,y,\Psi}, z_s^{t,y,\Psi})$ be the solutions of the BSDE:

$$y_t = h + \int_t^T g(y_s, z_s, s)ds - \int_t^T < z_s, dW_s >$$

corresponding to terminal values $h = \Phi(X_T^{t,x}), g = g_1$ and $h = \Psi(Y_T^{t,y}), g = g_2$, respectively, then, by Remark 2, $X_T = X_T^{0,x}$, $Y_T = Y_T^{0,y}$ and $z_s^\xi = z_s^{0,x,\Phi}$, $z_s^\eta = z_s^{0,y,\Psi}$.

But by Lemma 3(ii),

$$\begin{cases} z_s^{t,x,\Phi} = \sigma_1^*(s, X_s^{t,x})\partial_x u(s, X_s^{t,x}), \text{a.e. } s \in [0,T]; \\ z_s^{t,y,\Psi} = \sigma_2^*(s, Y_s^{t,y})\partial_y v(s, Y_s^{t,y}), \text{a.e. } s \in [0,T]. \end{cases}, \tag{10}$$

where $u(t,x) := y_t^{t,x,\Phi}$ and $v(t,y) := y_t^{t,y,\Psi}$.

Thus, let $t = 0$, it follows

$$\begin{aligned} z_s^\xi \odot z_s^\eta &= z_s^{0,x,\Phi} \odot z_s^{0,y,\Psi} \\ &= \sigma_1^*(s, X_s)\partial_x u(s, X_s) \odot \sigma_2^*(s, Y_s)\partial_y v(s, Y_s) \\ &= \sigma_1(s, X_s) \odot \sigma_2(s, Y_s)\partial_x u(s, X_s)\partial_y v(s, Y_s), \text{ a.e. } s \in [0,T]. \end{aligned} \tag{11}$$

Let us now prove (i):

Since Φ and Ψ are comonotonic, by Comparison Theorem for stochastic differential equation (SDE in short) ' , $X_T^{t,x}$ and $Y_T^{t,y}$ are increasing in x, y, thus for fixed T and t, $\Phi(X_T^{t,\cdot})$ and $\Psi(Y_T^{t,\cdot})$ are almost sure comonotonic, by Comparison Theorem of BSDE in [P1], we can conclude that $y_t^{t,x,\Phi}$ and $y_t^{t,x,\Psi}$ are comonotonic with x,

that is, $u(t,\cdot)$ and $v(t,\cdot)$ are comonotonic, which implies

$$\partial_x u(t,x)\partial_y v(t,y) \geq 0, t \in [0,T].$$

From (11), we obtain

$$z_t^\xi \odot z_t^\eta \geq 0$$

because of $\sigma_1(t, X_t) \odot \sigma_2(t, Y_t) \geq 0$.

We now prove (ii): If Φ and Ψ are strictly comonotonic, by strictly Comparison Theorem of BSDE, $u(t,x)$ and $v(t,y)$ are strictly comonotonic, which implies

$$\partial_x u(t,x)\partial_y v(t,y) > 0, t \in [0,T].$$

From (11), the proof of (ii) is complete.

Step 2: If b_i, σ_i, g_i $(i = 1,2), \Phi$ and Ψ are not belong to C^3, we can choose smooth functions $b_i^\epsilon, \sigma_i^\epsilon, g_i^\epsilon (i = 1,2), \Phi_\epsilon$ and Ψ_ϵ such that $b_i^\epsilon, \sigma_i^\epsilon, g_i^\epsilon, \Phi_\epsilon, \Psi_\epsilon$ converge to $b_i, \sigma_i, g_i, \Phi, \Psi$ uniformly over compact sets respectively.

Let $(y_s^{\xi,\epsilon}, z_s^{\xi,\epsilon})$ and $(y_s^{\eta,\epsilon}, z_s^{\eta,\epsilon})$ be the solutions of BSDEs, respectively :

$$y_t = \Phi_\epsilon(X_T^\epsilon) + \int_t^T g_1^\epsilon(y_s, z_s, s)ds - \int_t^T < z_s, dW_s >;$$

and

$$y_t = \Psi_\epsilon(Y_T^\epsilon) + \int_t^T g_2^\epsilon(y_s, z_s, s)ds - \int_t^T < z_s, dW_s >;$$

and $\{X_t^\epsilon\}$, $\{Y_t^\epsilon\}$ be the solutions of SDEs respectively:

$$\begin{cases} dX_s^\epsilon = b_1^\epsilon(s, X_s^\epsilon)ds + \sigma_1^\epsilon(s, X_s^\epsilon)dW_s, 0 \leq s \leq T, \\ X_0 = x, \ x \in R, \end{cases}$$

and

$$\begin{cases} dY_s^\epsilon = b_2^\epsilon(s, Y_s^\epsilon)ds + \sigma_2^\epsilon(s, Y_s^\epsilon)dW_s, 0 \leq s \leq T, \\ Y_0 = y, y \in R. \end{cases}$$

It is easy to check that $X^\epsilon \to X$ and $Y^\epsilon \to Y$ in $L^2(0,T)$.

Applying Lemma 2,

$$z^{\xi,\epsilon} \to z^\xi, \quad z^{\eta,\epsilon} \to z^\eta, \text{as} \quad \epsilon \to 0$$

in $L^2(0,T)$. By Step 1,

$$z_s^{\xi,\epsilon} \odot z_s^{\eta,\epsilon} \geq 0, \text{a.e.} s \in [0,T],$$

hence

$$z_s^\xi \odot z_s^\eta \geq 0, \text{a.e.} s \in [0,T].$$

Remark 3 *In Theorem 1, σ_1 and σ_2 are the volatilities of SDE(7) and (8), but z^ξ and z^η are the volatilities of BSDE(5) and (6), thus Theorem 1 shows a relation of volatilities between forward SDEs and Backward SDEs.*

In (10), let $t = 0$, we can obtain immediately:

Theorem 2 *Let $\{X_t\}$ be the solution of SDE(7) and there exists $T > 0$ such that $\Phi(X_T) \in L^2(\Omega, \mathcal{F}, P)$, let (y_t^Φ, z_t^Φ) be the solution of BSDE:*

$$y_t = \Phi(X_T) + \int_t^T g(y_s, z_s, s)ds - \int_t^T < z_s, dW_s >,$$

(i) If Φ is a continuous increasing function, then

$$z_t^\Phi \odot \sigma_1(t, X_t) \geq 0, a.e.\ t \in [0,T].$$

(ii) If Φ is a continuous decreasing function, then

$$z_t^\Phi \odot \sigma_1(t, X_t) \leq 0, a.e.\ t \in [0,T].$$

Proof. Let us observe (10). In (10), let $t = 0$, then $z_s^{\Phi} = z_s^{0,x,\Phi}$, not that $u(t,x) = y_t^{t,x\Phi}$ is the value of $\{y_s^{t,x,\Phi}\}_{0\leq s\leq T}$, the solution of the following BSDE, at time $s = t$:

$$y_s^{t,x,\Phi} = \Phi(X_T^{t,x}) + \int_s^T g_1(y_r^{t,x,\Phi}, z_r, r)dr - \int_s^T < z_r, dW_r >, s \in [0,T],$$

since Φ is increasing, by comparison theorem of BSDE, $y_t^{t,x\Phi}$ is increasing in x, hence

$$\partial_x u(s,x) \geq 0,$$

from (10),

$$z_s^{\Phi} = \sigma_1(s, X_s) \odot \sigma_1(s, X_s)\partial_x u(s, X_s) \geq 0.$$

The proof of (i) is complete. The proof of (ii) is similar to the proof of (i).
□

The assumption that Φ and Ψ are continuous in Theorem 1 and Theorem 2 is not necessary, let us now consider the case where Φ and Ψ are indicator functions:

Suppose that $\{X_t\}$ is the solution of SDE (7) such that $X_T \in L^2(\Omega, \mathcal{F}, P)$, for any $a < b < c$, let

$$A := \{X_T < a\}, \quad B := \{X_T \geq b\}, \quad C := \{X_T \geq c\}.$$

Let (y^A, z^A), (y^B, z^B) and (y^C, z^C) be the solutions of the following BSDE corresponding to $\xi = I_A, I_B$ and I_C, respectively:

$$y_t = \xi + \int_t^T g(y_s, z_s, s)ds - \int_t^T < z_s, dW_s > .$$

Here and next I_A is an indicator function.

Obviously, $C \subset B$ and $A \cap B = \emptyset$, let us observe $z^C \odot z^B$ and $z^A \odot z^B$.

Lemma 4 *Using the above notations, suppose that b_1, σ_1, g satisfy the assumption of Theorem 1, then*

(i) $z_t^C \odot z_t^B \geq 0$;

(ii) $z_t^A \odot z_t^B \leq 0$, *a.e.* $t \in [0,T]$.

Proof. First, let us construct C^3 functions which converge to indicator functions.

Indeed, for any $n = 1, 2, \cdots$, we denote by $\Phi_n(x,a)$, $\Phi_n(x,b)$ and $\Phi_n(x,c)$:

$$\Phi_n(x,a) := e^{-nd^{(a)}(x)}, \quad \Phi_n(x,b) := e^{-nd^{(b)}(x)}, \quad \Phi_n(x,c) := e^{-nd^{(c)}(x)},$$

where

$$d^{(a)}(x) = \begin{cases} (x-a)^3, & x > a, \\ 0, & x \leq a \end{cases};$$

$$d^{(b)}(x) = \begin{cases} (b-x)^3, & x < b, \\ 0, & x \geq b; \end{cases};$$

$$d^{(c)}(x) = \begin{cases} (c-x)^3, & x < c, \\ 0, & x \geq c. \end{cases}$$

It is easy to check that for each $n \geq 1$, $\Phi_n(\cdot, a), \Phi_n(\cdot, b), \Phi_n(\cdot, c) \in C^3$ and

$$\Phi_n(x,a) \to I_{(x \leq a)}, \quad \Phi_n(x,b) \to I_{(x \geq b)}, \quad \Phi_n(x,c) \to I_{(x \geq c)} \text{ as } n \to \infty.$$

Let us now prove (i):

Suppose that $(y^{n,a}, z^{n,a})$, $(y^{n,b}, z^{n,b})$ and $(y^{n,c}, z^{n,c})$ are the solutions of following BSDE corresponding to $\xi = \Phi_n(X_T, a)$, $\xi = \Phi_n(X_T, b)$ and $\xi = \Phi_n(X_T, c)$, respectively,

$$y_t = \xi + \int_t^T g(y_s, z_s, z)ds - \int_t^T < z_s, dW_s > .$$

Since $\Phi_n(\cdot, b)$ and $\Phi_n(\cdot, c)$ are comonotonic, applying Theorem 1,

$$z_s^{n,b} \odot z_s^{n,c} \geq 0, \quad \text{a.e. } s \in [0,T].$$

Note that $\Phi_n(X_T, b) \to I_B$ and $\Phi_n(X_T, c) \to I_C$ as $n \to \infty$ in $L^2(\Omega, \mathcal{F}, P)$, by Lemma 2

$$z^{n,b} \to z^B, \quad z^{n,c} \to z^C, \text{as } n \to \infty$$

in $L^2(0,T)$. The proof of (i) is complete.

Now let us prove (ii), indeed, it is easy to check that $(-y_t^A + 1, -z_t^A)$ is the solution of BSDE:

$$y_t = I_{(X_T > a)} + \int_t^T \overline{g}(y_s, z_s, s)ds - \int_t^T < z_s, dW_s >$$

where $\overline{g}(y,z,t) := -g(-y+1, -z, t)$.

We now prove

$$(-z_s^A) \odot z_s^B \geq 0, \text{a.e. } s \in [0,T].$$

Indeed, similar to the proof of (i), we can construct a C^3 function $\overline{\Phi}_n(x,a)$ such that $\overline{\Phi}_n(X_T, a) \to I_{(X_T > a)}$ as $n \to \infty$ and $\overline{\Phi}_n(x,a)$, $\Phi_n(x,b)$ are comonotonic.

Let $(y^{n,a,\overline{g}}, z^{n,a,\overline{g}})$ be the solution of BSDE:

$$y_t = \overline{\Phi}_n(X_T, a) + \int_t^T \overline{g}(y_s, z_s, s)ds - \int_t^T z_s dW_s$$

then by Lemma 2, $z^{n,a,\overline{g}} \to -z^A$ as $n \to \infty$ in $L^2(0,T)$.

Since $\overline{\Phi}_n(x,a)$ and $\Phi_n(x,b)$ are comonotonic , by Theorem 1,

$$z_s^{n,a,\overline{g}} \odot z_s^{n,b} \geq 0, \text{a.e. } s \in [0,T].$$

Let $n \to \infty$, note that $z^{n,b} \to z^B$ in $L^2(0,T)$, thus

$$(-z_s^A) \odot z_s^B \geq 0, \text{a.e. } s \in [0,T],$$

the proof of (ii) is complete.

Remark 4 *In Lemma 4, if we replace X_T by $\Phi(X_T)$, where Φ is an continuous increasing function, then the corresponding result in Lemma 4 is still true.*

Acknowledgements

The work was partly done while Chen was visiting IMPA Brazil, whose hospitality he deeply appreciated. Research supported in part by Foundation for University Key Teachers by Ministry of Education, P.R. China.

References

[CE] Z.Chen, L. Epstein *Epstein, Ambiguity, risk and asset return in continuous time,* to appear in Econometrica.

[H] Y.Hu, *Probabilistic interpretation for a systems of quasi elliptic PDE with Niemann boundary conditions,* Stoch. Prob. and Appli., 48, 107-121, 1993.

[K] N.El Karoui,*Backward stochasti differential equations–a general introduction* , Ptiman Research Notes in Mathematics Series, No.364, P 7-25,1997.

[PP] E. Pardoux and S. Peng, *Adapted Solution of a Backward Stochastic Differential Equation.* Systems and Control Letters, 14, 55-61, 1990.

[P1] S.Peng, *BSDE and related g–expectation,* Ptiman Research Notes in Mathematics Series, No. 364, PP 141-159, 1997.

[P2] S. Peng , *Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doom-Meyer type.* Proba. Theory & Related Fields, 113, 473-499, 1999.

[P3] S.Peng, *Probabilistic interpretation for systems of quasisimilar Parabolic PDE and application,* Stoch. and Stoch. Reports, 37, 61-74, 1991.

[MPY] J.Ma,J.Protter, J.Yong, *Solving Forward-backward stochastic differential equations—a four step scheme.* Proba. Theory and Related fields, 98, 339-359, 1994.

[YZ] J.Yong, X. Zhou, *Stochastic Controle: Hamiltonian Systems and HJB Equations,* Springer,New York, 1999.

Some Lookback Option Pricing Problems

Xin Guo
IBM T. J. Watson Research Center, P.O. Box 218
Yorktown Heights, NY 10598, USA

Abstract

We review some path-dependent option pricing problems in the financial market (Black-Scholes or incomplete) in the context of optimal stopping problems. Our focus is on the effectiveness and the limitations of the well-known technique of the "principle of smooth fit". We demonstrate concrete examples where this principle is sufficient in deriving closed-form solutions. We also provide cases where the smooth fit is necessary but not sufficient. We finally discuss an optimal stopping problem with regime switching where we extend the technique of smooth fit to allow instantaneous and discontinuous jumps to obtain closed-form solutions for pricing exotic options.

1 Introduction

Option pricing is one of the central problems in the study of financial markets. An option is a financial instrument that gives its holder the right but not the obligation to sell or buy the underlying asset (for example, a share of stock) on specific terms at a fixed instant T or an arbitrary time $t \leq T$ during a certain period of time $[0, T]$ in the future. Different types of options assign different payoff functions. For instance, the payoff of certain lookback options may depend on the minimum or maximum stock price achieved during the life of the option.

The valuation of options depends on the stochastic process for the underlying asset and on the proper choice of the probability measure. In 1900, L. Bachelier [1] proposed the following model for X_t, the stock price at time t,

$$dX_t = \mu dt + \sigma dW_t,$$

where μ, σ are constants and W_t is the standard Wiener process. In 1973, F. Black and M. Scholes abandoned Bachelier's Brownian motion model in favor of a geometric Brownian motion law

$$dX_t = \mu X_t dt + \sigma X_t dW_t.$$

The Black-Scholes model is "adduced to avoid the anomalies of Bachelier's unlimited liability" and its "log-normal asymptotes leads to rational pricing functions for warrants and options which satisfy complicated boundary conditions" (P. Samuelson's remark [25]). R. Merton [20] enriched and renovated the derivation of the Black-Scholes option pricing formula by introducing the idea of hedging in a continuously trading market. Upon applying Ito's formalisms, he deduced the Black-Scholes-Merton warrant-pricing functions, which depend only on certain interest rates and the common stock's relative variance. In the early 1980s, M. Harrison and D. Kreps [16] and S. Pliska [17] used martingales to provide a general mathematical framework for a frictionless security market. One of their major results states that the absence of arbitrage is effectively the same as the existence of an embedded unique martingale measure under which the price of derivatives is the expected discounted value of its future cash flow.

Their seminal work inspired mathematicians and led them to the center-stage of quantitative finance. Since then, standard hedge options together with various kinds of security derivatives, especially exotic types of options, have been introduced into this arena.

Values of path-dependent options are closely related to optimal stopping problems. This connection was first demonstrated in 1965 by H. P. McKean[23], who derived a closed-form solution for an optimal stopping problem that arises in pricing perpetual American put options. Since then, there has been a surge of renewed interest in optimal stopping problems and their applications in financial market. The relevant mathematical literature is voluminous and a small fraction of the references are: [18,19,26,28].

Along with the development of optimal stopping and option pricing problems comes the revival of traditional mathematical techniques. Among them, the "principle of smooth fit" has played an increasingly important role and is one of the most notable ones. Widely considered as the basis for the fundamental connection between optimal stopping and variational inequalities, this principle often enables one to explicitly solve optimal control/stopping problems in the continuous case whose closed-form solutions in discrete versions are not available.

We will review the effectiveness and the limitations of this well-known technique [2,6,3,9]. We demonstrate concrete examples where this principle is sufficient in deriving closed-form solutions. We also provide cases where the smooth fit is necessary but not sufficient. We finally discuss an optimal stopping problem with regime switching in which we extend the technique of smooth fit to allow instantaneous and discontinuous jumps to obtain closed-form solutions for pricing perpetual lookback options[a].

[a] Although there are many types of path-dependent options [10,18,19,26,28] such as American options, Asian options, and lookback options, our focus in this paper is on lookback options and the related optimal stopping problems.

2 Some lookback options in the Black-Scholes model

L. Shepp and A. Shiryayev [26] consider the following option. The owner of the option can choose any exercise date, represented by the stopping time τ ($0 \leq \tau \leq \infty$) and gets a payoff of either s (a fixed constant) or the maximum stock price achieved up to the exercise date, whichever is larger, discounted by $e^{-r\tau}$, where r is a fixed number. For this perpetual lookback type option, they coin the name "Russian option" out of respect for the great Russian mathematician A. Kolmogorov who enunciated the smooth fit technique [6,3].

Valuation of Russian options is based on solving the following optimal stopping time problem. Let $X = \{X_t, t \geq 0\}$ be the price process for a stock with $X_0 = x > 0$, and $S_t = \max\{s, \sup_{0 \leq u \leq t} X_u\}$, where $s > x$ is a given constant. How to compute the value of V,

$$V(x, s) = \sup_{\tau} E_{x,s} e^{-r\tau} S_\tau,$$

where τ is a stopping time with respect to the filtration $\mathcal{F}_{X_t} = \{X(s), s \leq t\}$, meaning no clairvoyance is allowed.

The stopping time τ^* for which the above maximum is achieved is called the *optimal stopping time.* τ^* is the first hitting time of a *stopping region*, that is, the region when $V(x, s) = s$. When this region is reached, the best policy is simply to *pull out of the game*, namely, stop and cash the reward s. In its complimentary region called the *continuation region*, one has $V(x, s) \geq s$, which means it pays to continue holding the option although one has the option to stop. The key to solving optimal stopping problems is to find the "free boundary" $X_t = g(S_t)$ between the continuation region and the stopping region

Assuming stock price fluctuations X_t to follow the Black-Scholes geometric Brownian motion model, L. Shepp and A. Shiryayev obtain an explicit solution to the above problem.

Their main result is that the value function $V(x, s)$ is finite if and only if $r > \mu$. And when $r > \mu$, the optimal stopping time τ^* is

$$\tau^* = \inf\{t > 0 \mid X_t \leq \alpha S_t \mid X_0 = x, S_0 = s\},$$

where

$$\alpha = \left(\frac{1 - 1/\gamma_1}{1 - 1/\gamma_0}\right)^{1/(\gamma_0 - \gamma_1)}, \tag{2.1}$$

and $\gamma_0 > 1 > 0 > \gamma_1$ are the solutions of

$$r = \mu\gamma + 1/2\sigma^2\gamma(\gamma - 1).$$

Compared to McKean's American put option problem with the payoff function $(X_t - K)^+$ (where K is a constant and the so called "strike price"), the payoff function for Russian options is more complex and involves the running maximum of the Brownian motion. Value function for Russian options is uniquely determined via the principle of smooth fit. It is also closely related to the fact that X_t/S_t is Markovian for a geometric Brownian motion X_t (as was first pointed out by P. Levy). In fact, this observation led to a simpler derivation of the above optimal stopping problem[27]: finding the free boundary $X_t = g(S_t)$ was reduced to finding a threshold of the Markov process (X_t/S_t) via a first passage time technique.

Building on Shepp and Shiryayev's analysis, D. Duffie and M. Harrison derive a unique arbitrage-free price for the Russian option[8] by assuming the existence of a dividend payoff for the underlying asset: the value is finite if and only if the dividend payout rate $\delta = r - \mu$ is strictly positive.

In[14], we combine McKean' problem with that of Shepp and Shiryayev and consider a more general type of lookback options, whose payoff function $c(s)$ is a function of S_t and K: the amount by which the maximum stock price achieved during the life ($T \leq \infty$) of the option exceeds a fixed (strike) price (say K). It has features of American options in that it gives the holder the choice of an arbitrary exercise time. However, unlike standard American call options whose payoff depends on the difference between the spot price at the execution time and the strike price K, the payoff of this option involves the running maximum of the stock price up to its execution time. Therefore, we call it the *perpetual lookback American option.*

The corresponding optimal stopping time problem is to find the value of

$$V^*(x,s) = \sup_{0 \leq \tau \leq \infty} E_{x,s}[e^{-r\tau} c(S_\tau)],$$

where $S_t = \max_{0 \leq u \leq t} X_u$ is the running maximum of X_t, $X_0 = x$, τ is a stopping time meaning no clairvoyance is allowed, and $c(s)$ is the utility function.

In[14], based on the Black-Scholes geometric Brownian motion model for X_t, we prove the existence of the optimal stopping time τ^*, provided that $c(s)$ satisfies certain growth conditions. Consequently, the value functions rules are explicitly characterized and derived.

Our main result is the following. If $c(s)$ satisfies the following properties:
(1) $c(s) = 0$, when $s < K$ for some constant K;
(2) $\lim_{s \to \infty} c(s)/s$ exists and equals some positive constant;
(3) For $s > K$, $c(s)$ has a derivative strictly bounded away from zero;
(4) $\tilde{l}(s) = c(s)/(sc'(s))$ is increasing to a finite limit which is bigger than $F(0)$, then

THEOREM 2.1 *(i) The value function $V^*(x,s)$ is finite iff $r > \max(0,\mu)$.*
(ii) When $r > \max(0,\mu)$,

$$V^*(x,s) = \begin{cases} c(s), & x \leq g(s), s > K, \\ \frac{c(s)}{\gamma_0-\gamma_1}\left(\gamma_0(\frac{x}{g(s)})^{\gamma_1} - \gamma_1(\frac{x}{g(s)})^{\gamma_0}\right), & s > x \geq g(s), s > K, \\ A\frac{\gamma_1}{\gamma_1-\gamma_0}x^{\gamma_0}, & s < K, x < s, \end{cases}$$

where $0 \leq g(s) < \alpha s$ for all $s > K$ and $g(s)$ is the (unique) solution to the differential equation

$$c'(s)\left[\gamma_0\left(\frac{s}{g(s)}\right)^{\gamma_1} - \gamma_1\left(\frac{s}{g(s)}\right)^{\gamma_0}\right] = -\gamma_0\gamma_1 c(s)\frac{g'(s)}{g(s)}\left[\left(\frac{s}{g(s)}\right)^{\gamma_0} - \left(\frac{s}{g(s)}\right)^{\gamma_1}\right],$$

such that when $s \to \infty$, $g(s)/s$ is asymptotic to α, i.e., $\lim_{s\to\infty}\frac{g(s)}{s} = \alpha$ and $0 \neq A = \lim_{s\to K+}\frac{c(s)}{g(s)^{\gamma_0}}$. Here $\gamma_0 > 1 > 0 > \gamma_1$ are the solutions of

$$r = \mu\gamma + 1/2\sigma^2\gamma(\gamma-1),$$

and α satisfies

$$F(\alpha) = \lim_{s\to\infty}\frac{c(s)}{sc'(s)}.$$

where

$$F(x) = \frac{1}{-\gamma_0\gamma_1}\frac{\gamma_0 x^{\gamma_0-\gamma_1} - \gamma_1}{1 - x^{\gamma_0-\gamma_1}}.$$

The optimal stopping time τ^* is

$$\tau^* = \inf\{t \geq 0 \mid X_t = g(S_t), S_t > K \mid X_0 = x, S_0 = s, x > g(s)\}.$$

In particular, let $c(s) = (s-K)^+$. Then, we have

COROLLARY 2.2 *(i) The value function $V^*(x,s)$ is finite iff $r > \max(0,\mu)$.*
(ii) When $r > \max(0,\mu)$, the stopping time τ^* is given by

$$\tau^* = \inf\{t \geq 0 \mid X_t = g(S_t), S_t > K\},$$

starting from $X_0 = x > 0, S_0 = s, x > g(s)$, where $0 \leq g(s) < \alpha s$ for all $s > K$ and $g(s)$ is the (unique) solution to the differential equation

$$\gamma_0\left(\frac{s}{g(s)}\right)^{\gamma_1} - \gamma_1\left(\frac{s}{g(s)}\right)^{\gamma_0} = -\gamma_0\gamma_1(s-K)\frac{g'(s)}{g(s)}\left[\left(\frac{s}{g(s)}\right)^{\gamma_0} - \left(\frac{s}{g(s)}\right)^{\gamma_1}\right],$$

such that when $s \to \infty$, $g(s)/s$ *is asymptotic to* α, *i.e.*, $\lim_{s\to\infty} \frac{g(s)}{s} = \alpha$ *and* $0 \neq A = \lim_{s\to K+} \frac{s-K}{g(s)^{\gamma_0}}$.

Notice that here

$$\alpha = \left(\frac{1 - 1/\gamma_1}{1 - 1/\gamma_0}\right)^{1/(\gamma_0 - \gamma_1)},$$

and is the same as that in [26].

In spite of this connection, the free boundary in this lookback American option problem seems totally irrelevant to the fact that X_t/S_t is Markovian. This is in contrast to the result in [26] where the Markov structure of X_t/S_t is extensively exploited. It further confirms our understanding that the structure solution in [26] is intrinsically related to the linearity of the payoff function when $K = 0$.

Another interesting point is the following. Although it is often the case in explicit solutions of optimal stopping problems that smooth fit uniquely identifies the boundary, and vice versa (cf. [13,18,23,28] and the references therein), the intriguing feature of the solution to perpetual lookback American options is the fact that smooth fit itself is not enough. Rather, one needs an additional condition to ensure proper growth of the value function in order to carry out the martingale argument. This is why the condition on the asymptotic behavior of $g(s)$ is required.

3 Option pricing in a market model with regime switching

One of the beauties of the Black-Scholes model lies in its simplicity which enables us to obtain closed-form solutions for various types of options, standard or exotic, as we saw in the previous section. Therefore, despite mounting empirical evidence that show the limitations of the Black-Scholes model, it remains irreplaceable.

In our search for a less simplistic yet simple model, we propose [15] a market model by incorporating the information structure among investors' community, represented by a stochastic process $\epsilon(t)$.

We theorize that the market activity is accompanied by changes in the information structure. In other words, instead of focusing only on the process X_t, we turn our attention to the joint process $(X_t, \epsilon(t))$.

Consequently, the fluctuations of the stock price X_t is assumed to follow an equation of the form

$$dX_t = X_t \mu_{\epsilon(t)} dt + X(t) \sigma_{\epsilon(t)} dW_t,$$

where the stochastic process $\epsilon(t)$ is independent of W_t and represents the state of information in the investors' community. For each state i, there is a *known* drift parameter μ_i and a known volatility parameter σ_i. $(\mu_{\epsilon(t)}, \sigma_{\epsilon(t)})$ take different values when $\epsilon(t)$ is in different states. We assume that $\epsilon = \epsilon(t)$ is a Markov process which moves among M ($= 2$ or more) states.

In particular, when $M = 2$, for which $\epsilon(t)$ alternates between 0 and 1 and where $\sigma_0 \neq \sigma_1$, we assume that

$$P(\tau_i > t) = e^{-\lambda_i t}, \qquad i = 0, 1.$$

It is easy to see that in this model, the joint process (X_t, S_t) is Markovian, although X_t alone is no longer a Markov process. This reflects the idea that the stock price change is not independent of information distribution of the investors' community.

Furthermore, this model is not "complete" [16,17] because of the additional process $\epsilon(t)$. In other words, $\epsilon(t)$ is a bounded adapted process with respect to the σ-algebra $\mathcal{F}_t$ generated by X_t (denoted as $\mathcal{F}^X$), but is not adapted to the σ-algebra generated by W_t (written as $\mathcal{F}^W$).

To remedy this, D. Duffie proposed one way to complete the market by issuing a security named COS ("Change Of States"): at each time t, there is a market for this security that pays one unit of account (say, a dollar) at the next time $\tau(t) = \inf\{u > t \mid \epsilon(u) \neq \epsilon(t)\}$ that the Markov chain $\epsilon(t)$ changes state. That contract then becomes worthless (i.e., has no future dividends), and a new contract is issued that pays at the next change of state, and so on.

Under natural pricing, this COS will complete the market and provide unique arbitrage-free prices to hedge options on the underlying risk asset. Moreover, assuming the existence of this COS, we can find a martingale measure Q under which X_t satisfies

$$dX(t) = (r - d_{\epsilon(t)})X(t)\,dt + X_t \sigma_{\epsilon(t)}\,dB^Q,$$

where B^Q is a standard Brownian motion under Q.

Based on this framework, we investigate option pricing problems for standard hedge options such as European options and perpetual lookback options. Although pricing European options [15] requires no more than detailed and delicate analysis of the Laplace transform, pricing path-dependent options in this model with regime switching demands more than direct applications of existing techniques.

The most interesting one is the optimal stopping problem for pricing perpetual lookback options [13]. Its closed-form solution is obtained by extending the technique of the "principle of smooth fit" to allow discontinuous jumps.

We show that when $M = 2$ and the hidden Markov process $\epsilon(t)$ switches from one state to another, there is a discontinuous jump over the free boundary. The optimal stopping rule is given by

$$\tau^* = \inf\{t > 0, X_t < a_i S_t \mid X_0 = x, S_0 = s, \epsilon(0) = i\}$$

where a_i are functions of the parameters $\mu_i, \sigma_i, \lambda_i$.

We also show that for cases where $M > 2$, there are no explicit closed-form solutions for lookback options, since one has to solve an algebraic equation of order $2M$. The proof of the result is via martingale theory.

Not surprisingly, the success in obtaining an explicit closed-form solution for the model with regime switching relies heavily on the Markov structure of $(X_t/S_t, \epsilon(t))$. To validate this in comparison to P. Levy's observation, we need only replace the classical Wiener space in [27] generated by $(\Omega, \mathcal{F}, F^W = (\mathcal{F}_t^W)_{t\geq 0}, P)$ with $(\Omega, \mathcal{F}, F^X = (\mathcal{F}_t^X)_{t\geq 0}, P)$.

Despite the fact that the solution structure is relevant to the Markovianness of $(X_t/S_t, \epsilon(t))$, it is much more difficult to solve the optimal stopping problem via the first passage time technique. The difficulty comes from the "instantaneous jump" because of which one needs to solve an integral equation system; this appears hard.

Although the technique can be applied directly to other path-dependent options such as American put options, it lends little help in case of $M > 2$ where the alternative is numerical calculations. In order to facilitate numerical simulations, we present one way of discretizing the continuous market model inspired by that of Cox, Ross, and Rubinstein [7]. The idea of the proof of convergence from the discrete model to the continuous one relies on the well-known Skorohod imbedding technique [15].

In [12], we explore some aspects of first passage time problems for the model with regime switching. This result provides different perspectives of the understanding of this model which would be useful for investigation of various problems such as hedging and pricing.

Our current project is the statistical testing of this Gaussian mixture model and comparison of this model with various market models in option pricing [5].

1. L. Bachelier (1900): Theorie de la speculation. *Annales Scientifiques de L'École Normale Supérieure* 3d ser., 17, 21–88.
2. J. Bather and H. Chernoff (1967): Sequential decisions in the control of a spaceship. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* 3, 181–207.
3. A. Bensoussan and J. L. Lions (1978): Applications des inéquations variationelles en controle stochastique. *Applications of Variational Inequalities in Stochastic Control*, North-Holland.

4. F. Black and M. Scholes (1973): The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.
5. D. X. Chan, X. Guo, and E. Pednault (2001): Determining the number of states in a hidden Markov process for Gaussian mixture models. *Manuscript.*
6. H. Chernoff (1961): Sequential tests for the mean of a normal distribution. *Proc. 4th. Berk. Symp. Stat. Prob.* 1: 79–91.
7. J. Cox, S. Ross, and M. Rubinstein (1979): Option pricing: A simplified approach. *Journal of Financial Economics* 7, 229–263.
8. D. Duffie and M. Harrison (1993): Arbitrage pricing of perpetual lookback options. *Annals of Applied Probability* 3, 641–651.
9. A. Friedman (1976): *Stochastic Differential Equations and Applications II*, Academic Press.
10. H. Geman and M. Yor (1993): Bessel processes, Asian options, and perpetuities. *Mathematical Finance* 3, 349–375.
11. X. Guo (1999): Inside Information and Stock Fluctuations. *Ph.D dissertation, Mathematics Department.* Rutgers University.
12. X. Guo (2001): When the 'bull' meets the 'bear'— a first passage time problem for a hidden Markov process. *Methodology and Computation in Applied Probability*, Kluwer.
13. X. Guo (2001): An explicit solution to an optimal stopping problem with regime switching. *Journal of Applied Probability* 38(2), 464–481.
14. X. Guo and L. Shepp (2001): Some optimal stopping problems with non-trivial boundaries for pricing exotic options. *Journal of Applied Probability* 38(3), 1–12.
15. X. Guo (2001): Information and option pricings. *Quantitative Finance* 1, 38–44.
16. M. Harrison and D. Kreps (1979): Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* 20, 381–408.
17. M. Harrison and S. Pliska (1981): Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications* 11, 215–260.
18. S. D. Jacka (1991): Optimal stopping and the american put option. *Mathematical Finance* 1(2), 1–14.
19. I. Karatzas (1988): On the pricing of American options. *Applied Mathematics and Optimization* 17, 37–60.
20. R. Merton (1973): The theory of rational option pricing. *Bell Journal of Economics and Management Science* 4, 141–183.
21. D. Revuz and M. Yor (1991): *Continuous Martingale and Brownian Motion.* Springer-Verlag.
22. H. Robbins, D. Sigmund, and Y. Chow (1971): *Great Expectations: The Theory of Optimal Stopping.* Houghton-Nifflin.
23. H. P. McKean (1965): Appendix: A free boundary problem for the heat equation arising from a problem in mathematical economics. *Industrial Management Review* 6, 32–39.

24. P. Samuelson (1965): Rational theory of warrant pricing. *Industrial Management Review* 6, 13–31.
25. P. Samuelson (1973): Mathematics of speculative price (with an appendix on continuous-time speculative processes by Robert C. Merton). *SIAM Review* 15(1), 1–42.
26. L. Shepp and A. N. Shiryaev (1993): The Russian option: Reduced regret. *Annals of Applied Probability* 3, 631–640.
27. L. A. Shepp and A. N. Shiryaev (1994): A new look at pricing of the "Russian options". *Theory Probability and its Application* 39(1), 103–119.
28. P. L. J. Van Moerbeke (1976): On optimal stopping and free boundary problems. *Archive for Rational Mechanics and Analysis* 60(2), 101–148.

Option Pricing in a Market Where the Volatility Is Driven by Fractional Brownian Motions

Yaozhong Hu
Department of Mathematics, University of Kansas
405 Snow Hall, Lawrence, KS 66045-2142
Email: hu@math.ukans.edu

Abstract

In this paper the stochastic volatility model of Stein and Stein is extended to treat the long memory character of the volatility. It is proposed to model the volatility by a mean reverting Langevin equation driven by fractional Brownian motions. The risk-minimizing hedging price for European call options is obtained and its computation is discussed.

1 Introduction

Ever since the work of Black and Scholes[1] and Merton[16], option pricing theory has received a great deal of attention from researchers of various disciplines.

The original work of Black and Scholes assumes that the stock prices follows the geometric Brownian motions. For simplicity, let us assume that the financial market is given by two securities, described by

$$\begin{aligned} d\beta_t &= r\beta_t dt\,, \qquad \beta_0 = 1 \\ dX_t &= \mu X_t dt + \sigma X_t dW_t\,, \quad X_0 = x \quad \text{is given}\,. \end{aligned} \tag{1.1}$$

where r, μ, and σ are constants and $(W_t, t \geq 0)$ is a standard Brownian motion on some probability space $(\Omega, \mathcal{F}, P)$. Usually β_t denotes the price of a bond and X_t denotes the price of a stock at current time t.

An (European call) option is the right, but not the obligation, to to buy a share of the stock at a specific expiration time T with the striking price K. Thus the profit of the holder of the option at time T is $(X_T - K)^+$, where a^+ denotes the positive part of a, *i.e.* $a^+ = a$ if $a > 0$ and $a^+ = 0$ if $a \leq 0$.

With an arbitrage argument, it is known that the fair price to hold a European call option at the initial time 0 is given by the famous Black and

---

AMS 1991 subject classifications. 60H10, 60H30, 60J65, 91B28.
Key words and phrases: Fractional Brownian Motions, Black Scholes formula, Option price, European option, determinant, trace, Lie algebra
This work is supported in part by the National Science Foundation under Grant No. EPS-9874732, matching support from the State of Kansas and General Research Fund of the University of Kansas.

Scholes formula:

$$C_{BS}(t,x,\sigma) = x\Phi(d_1) - Ke^{-rT}\Phi(d_2)\,, \tag{1.2}$$

where

$$d_1 = \frac{\log(x/K) + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}\,,$$

and

$$\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-y^2/2}dy\,.$$

There have been great amount of extensions of the classical model (1.1). One class of extensions is to substitute the *volatility* σ by a random process driven by another (maybe correlated) Brownian motion. More precisely, σ in (1.1) is replaced by $\sigma_t = f(Y_t)$, where f is a given function and Y_t satisfies a stochastic differential equation. Three particular stochastic differential equations have been attracted more attention. We refer to [9] and the references therein for more details.

In a few earlier work, there have been efforts to apply fractional Brownian motions to the modeling of the stock return in order to capture the long range dependence character of some financial markets. However, it is proved [18] that if one replaces W_t in (1.1) by a fractional Brownian motion B_t^H, there will be arbitrage opportunities in the market (see also for instance, [14], [15], [19]). In [8] it is introduced a new type of stochastic integral. This new stochastic calculus has been applied to the above arbitrage problem ([14]).

In this paper, we shall consider the financial markets whose volatilities are described by stochastic differential equations driven by fractional Brownian motions. More precisely, we shall consider the markets of two securities, described by

$$\begin{aligned} d\beta_t &= r\beta_t dt\,, \qquad \beta_0 = 1 \\ dX_t &= \mu X_t dt + \sigma_t X_t dW_t\,, \quad X_0 = x \quad \text{is given}\,. \end{aligned} \tag{1.3}$$

where r, μ, and $(W_t, t \geq 0)$ are as above and $\sigma_t = f(Y_t)$ for $f = |x|$ and

$$dY_t = dY_t = \alpha(m - Y_t)dt + \beta dB_t^H\,, \qquad Y_0 = y \quad \text{is given}\,. \tag{1.4}$$

Here B_t^H is a fractional Brownian motion with Hurst parameter H, of the type introduced in [17]. The modeling in the financial market using long memory process has appeared in some economic literature, see for instance, [2-6]. In particular, in [4], it is proposed that the volatility be modeled by $\sigma_t = e^{Y_t}$ and $dY_t = k(\theta - Y_t)dt + \gamma d\tilde{B}_t^H$, where $\tilde{B}_t^H$ is a fractional Brownian motion of type different to that introduced by [17]. Some properties of this model have also been studied in [4]. In some sense, the model of [4] is the extension of a model of Scott to the long memory situation.

Comparing to the earlier work, our model is the extension of the model of Stein and Stein (see for example, [9] and [20]) to long memory setting. We will focus on the computational aspects of the fair price.

In Section 2, we shall sketch the obtention of the risk-minimizing hedging price for the European option and in Section 3, we will present two ways to compute the fair price obtained in Section 2.

2 Risk Minimizing Hedging Price

First let us derive some properties of the market.

From (1.3), it follows that

$$\gamma_n = \operatorname{cov}(X_1, X_{n+1} - X_n) = 0\,.$$

Thus the stock return process is not a long memory process! I fact this process is a semimartingale. However, the volatility process σ_t is a long memory process, this means that $\sum_{n=1}^{\infty} \gamma_n = \infty$.

Since we introduce another independent random factors, we can expect that the market is now no longer complete.

Lemma 2.1 *The market defined by (1.3)-(1.4) is incomplete.*

Proof Denote $z_t = X_t/\beta_t\,, t \geq 0$. Then

$$dz_t = (\mu - r)Z_t dt + \sigma_t Z_t dW_t\,.$$

Let

$$\eta_t = \exp\left\{\int_0^t \gamma_u dW_u - \frac{1}{2}\int_0^t |\gamma_u|^2 du\right\}\,, \quad t \geq 0\,, \tag{2.1}$$

where $\gamma_u = (r - \mu)/\sigma_u$ and introduce a new measure

$$\frac{d\mathbb{P}'}{d\mathbb{P}} = \eta_T\,, \quad \mathbb{P} - a.s. \tag{2.2}$$

It is easy to see that η is a square integrable martingale. By Girsanov theorem, $\tilde{W}_t = W_t - \int_0^t \gamma_u du$ is a $\mathbb{P}'$ Brownian motion and then

$$z_t = z_0 + \int_0^t \sigma_u z_u d\tilde{W}_u$$

is a martingale. Since B_t^H is a fractional Brownian motion, independent of W_t, there is a standard Brownian motion B_t such that it is independent of W_t such that

$$B_t^H = \int_0^t K(t,s)dB_s$$

(see [7]) and B_t^H and B_t generates the same filtration. It is easy to see that

$$\eta_t' = \exp\left(B_t - t/2\right)\,, \quad t \geq 0$$

is another square integrable martingale. Define $\tilde{\mathbb{P}}$ by $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \eta_T\eta_T'$. Then $\tilde{\mathbb{P}}$ is another measure such that $z_t, 0 \leq t \leq T$ is a martingale. Thus we found two distinct equivalent martingale measures. This implies that the market defined by (1.3)-(1.4) is incomplete. □

Lemma 2.2 *$\mathbb{P}'$ defined by (2.2) is the minimal martingale measure associated with $\mathbb{P}$.*

Proof This theorem can be proved in similar way as in [15], p.228-229, Lemma 13.3, combined with the proof the previous theorem. We shall not give more detail. □

Theorem 2.3 *Let $(X_T - K)^+$ be a European call option settled at time T. Then the risk-minimizing hedging price is*

$$v = \mathbb{E}^{\mathbb{P}'}\left[(X_T - K)^+\beta_T^{-1}\right]\,. \tag{2.3}$$

Proof This theorem can be proved in similar way as in [15], p.229. □

To compute the risk-minimizing hedging price given by (2.3), first let us follow the idea presented in [9]. The explicit solution of (1.3) is given by

$$X_T = x_0 \exp\left\{\int_0^T \sigma_s dW_s + \int_0^T \left(\mu - \frac{1}{2}\sigma_s^2\right) ds\right\}$$

Given $\mathcal{G}_T = \sigma\left(B_t^H, 0 \leq t \leq T\right)$, $\int_0^T \sigma_s dW_s$ is a Gaussian random variable with mean 0 and variance $\sigma^2 = \int_0^T \sigma_s^2 ds$. Thus it follows from the Black and Scholes formula that the risk-minimizing hedging price is

$$\begin{aligned} v &= \mathbb{E}^{\mathbb{P}'}\left\{\mathbb{E}^{\mathbb{P}'}\left[(X_T - K)^+\beta_T^{-1}|\mathcal{G}_T\right]\right\} \\ &= \mathbb{E}^{\mathbb{P}'}\left\{C_{BS}(x,\sigma)\right\}\,, \end{aligned}$$

where C_{BS} is defined by (1.2) and

$$\sigma = \sqrt{\int_0^T \sigma_s^2 ds}\,. \tag{2.4}$$

Now we summarize the above results as follows. The market consists of two securities, one is bond and another is stock. The stock price follows a "generalized" geometric Brownian motion with volatility modeled by the absolute value of a mean reverting Gaussian process, driven by fractional Brownian motion. Namely, the market is described by

$$\begin{cases} d\beta_t = r\beta_t dt\,, \quad \beta_0 = 1 \\ dX_t = \mu X_t dt + |Y_t| X_t dW_t\,, \quad X_0 = x \quad \text{is given} \\ dY_t = \alpha(m - Y_t)dt + \beta dB_t^H\,, \quad Y_0 = y \quad \text{is given}\,, \end{cases} \tag{2.5}$$

where $r > 0$, $\mu > 0$, $x > 0$, y, α and β are given constants, $(W_t\,, t \geq 0)$ is a Brownian motion and $(B_t^H\,, t \geq 0)$ is a fractional Brownian motion, independent of W. We also assume that $B_t^H = \int_0^T \kappa(t,s)dB_t$ for a Brownian motion $(B_t\,, t \geq 0)$, independent of W. All the process are considered in the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where the expectation is denoted by $\mathbb{E}$.

The risk minimizing hedging price for the European option $(X_T - K)^+$ is

$$C(t,x,y) = \mathbb{E}\,\left[C_{BS}(t,x,K,T,\sqrt{\bar{\sigma}^2})\right]\,, \tag{2.6}$$

where $C_{BS}(t,x,K,T,\sqrt{\bar{\sigma}^2})$ is defined by 1.2 and

$$\bar{\sigma}^2 = \frac{1}{T}\int_0^T Y_s^2 ds\,. \tag{2.7}$$

Since the linearity of the equation for Y_t, we can solve the third equation of (2.5) to obtain an explicit expression for Y_t:

$$Y_t = \alpha(t) + \int_0^t \beta(t,s)dB_s^H\,, \tag{2.8}$$

where

$$\alpha(t) = e^{-\alpha t}y + m - me^{-\alpha t}$$

and

$$\beta(t,s) = \beta e^{-\alpha(t-s)}\,.$$

Thus Y_t is a Gaussian process.

3 The Expectation of Quadratic Forms of Gaussian Processes

To find the risk-minimizing hedging price, one needs to compute the expectation (2.6). We consider $C_{BS}(t,x,K,T,\sqrt{z/T})$ as a function of $z \geq 0$ (and consider x, K, and T as constants), and denoted it by $C_{BS}(z)$. It is clear that $C_{BS}(z)$ is a bounded function of z. Its Fourier transform is then in $L^1(\mathbb{R})$. Using inverse Fourier transform formula, we can find a function $g(\xi)$, which may depend on x, K, and T, such that

$$C_{BS}(t, x, K, T, \sqrt{z/T}) = C_{BS}(z) = \int_{\mathbb{R}} g(\xi) e^{i\xi z} g(\xi) d\xi$$

and $\int_{\mathbb{R}} |g(\xi)| d\xi < \infty$, where $i = \sqrt{-1}$. Note that g does not depend on the volatility.

To compute the expectation of (2.6), we need to compute

$$\mathbb{E}\left\{ e^{iZ + i\xi \int_0^T Z_t^2 dt} \right\}, \tag{3.1}$$

where Z is a Gaussian random variable, $(Z_t, t \geq 0)$ is a Gaussian process of mean 0, Z and $(Z_t, t \geq 0)$ is jointly Gaussian.

When Z is a Brownian motion, the computation of (3.1) can be carried out in many ways, one can use the so-called Cameron formula, and one may also use partial differential equations. However, in the fractional Brownian motion, or in more general Gaussian process setting, the explicit computation of (3.1), to the best of my knowledge, is still unknown.

First we extend a well-known method to our situation. This method appeared to be due to K. Itô (see also [12] for an application to the computation of some Feynman integrals). Let $(Z_t, 0 \leq t \leq T)$ be a Gaussian process with mean 0 and covariance

$$K(t, s) = \mathbb{E}\ (Z_t Z_s)\,, \quad 0 \leq s, t \leq T\,. \tag{3.2}$$

Assume that $\displaystyle\int_0^T \int_0^T |K(t, s|^2 ds dt < \infty$. Denote $L^2([0, T])$ the Hilbert space of square integrable functions on $[0, T]$. Define

$$Tf(t) = \int_0^T K(t, s) f(s) ds\,, \quad \forall\, t \in [0, T]\,. \tag{3.3}$$

Then T is a non-negative definite, Hilbert-Schmidt operator on $L^2([0, T])$. Thus there is an orthonormal basis $\{e_1, e_2, \cdots\}$ of $L^2([0, T])$ and $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots$ such that

$$\sum_{n=1}^{\infty} \lambda_n^2 < \infty$$

and

$$Te_n = \lambda_n e_n\,, \quad n = 1, 2, \cdots$$

Therefore

$$K(t,s) = \lim_{m\to\infty} K_m(t,s)\,,$$

where

$$K_m(t,s) = \sum_{n=1}^{m} \lambda_n e_n(t) e_n(s)$$

and the convergence is in $L^2([0,T]^2)$ (in the operator's Hilbert-Schmidt norm).

Introduce

$$\Theta_m(t) = \sum_{n=1}^{m} \sqrt{\lambda_n} e_n(t) G_n\,,$$

where $\{G_1, G_2, \cdots\}$ is a sequence of independent standard normal variables. Θ_m converges in L^2 and then almost surely (a subsequence) converges to

$$\Theta_t = \sum_{n=1}^{\infty} \sqrt{\lambda_n} e_n(t) G_n\,, \tag{3.4}$$

It is obvious that the Gaussian process $(Z_t\,, 0 \le t \le T)$ is identical in law to the Gaussian process $(\Theta_t,\,, 0 \le t \le T)$.

When $Z = 0$ the computation of (3.1) can be reduced to the computation of

$$\begin{aligned}
\mathbb{E}\ \exp\left\{ i\xi \int_0^T Z_t^2 dt \right\} &= \lim_{m\to\infty} \mathbb{E}\ e^{i\xi \int_0^T \left(\sum_{n=1}^m \sqrt{\lambda_n} e_n(t) G_n\right)^2} \\
&= \lim_{m\to\infty} \mathbb{E}\ e^{i\xi\left(\sum_{n=1}^m \lambda_n G_n^2\right)} \\
&= \lim_{m\to\infty} \prod_{n=1}^{m} \mathbb{E}\ e^{i\xi\lambda_n G_n^2} \\
&= \lim_{m\to\infty} \prod_{n=1}^{m} (1 - 2i\xi\lambda_n)^{-1/2} \\
&= \prod_{n=1}^{\infty} (1 - 2i\xi\lambda_n)^{-1/2}
\end{aligned}$$

This limit exists when $\sum_{n=1}^{\infty} \lambda_n < \infty$.

Using the notation of determinant of an operator, we obtain

Theorem 3.1 *Let T defined by (3.3) be a trace class operator. Then*

$$\mathbb{E}\ \exp\left\{i\xi\int_0^T Z_s^2 ds\right\} = \det\left(I - 2i\xi T\right)^{-1/2}. \tag{3.5}$$

In a similar argument, one can obtain that if Z is a Gaussian random variable such that Z and $(Z_t\,, 0 \le t \le T)$ is jointly Gaussian and

$$R_t = \mathbb{E}\ (ZZ_t)\ , \quad 0 \le t \le T \tag{3.6}$$

is in $L^2([0,T])$. Then we have

Theorem 3.2 *Let T defined by (3.3) be a trace class operator and R defined by (3.6) is in $L^2([0,T])$. Then*

$$\mathbb{E}\ \exp\left\{iZ + i\xi\int_0^T Z_s^2 ds\right\} = \det\left(I - 2i\xi T\right)^{-1/2}\exp\{-\frac{1}{2}\langle R, \left(I - 2i\xi T\right)^{-1} R\rangle\,, \tag{3.7}$$

where $\langle f, g\rangle = \int_0^T f(t)g(t)dt$ denotes the scalar product of $L^2([0,T])$.

We may need to compute $\mathbb{E}\ \exp\left\{iZ + i\xi\int_0^T Z_s^2 ds\right\}$ for many different expiration times T's and in most situations, it is very difficult to find the eigenvalues and the eigenfunction of T. Now we propose an another approach to compute it. For simplicity, we assume $Z = 0$. We begin by writing the determinant of $I - 2i\xi T$ as

$$\det(I - 2i\xi T)^{-1/2} = \exp\left\{-\frac{1}{2}\ \mathrm{Tr}\left[\ln(I - 2i\xi T)\right]\right\},$$

where $\mathrm{Tr}A$ is the trace of an operator A: If A is associated with kernel $a(t,s)\,, 0 \le s, t \le T$, then

$$\mathrm{Tr}A = \int_0^T a(t,t)dt\,,$$

See for example, [13] and in particular the references therein for more discussion of trace.

The trace is much easy to compute than the determinant. However, we need to compute the function of an operator T: $U = \ln(I - 2i\xi T)$.

Now we are going to consider the algebraic aspect of the computation of $U_t = \ln(I - 2i\xi T_t), 0 \le t \le T$.

We write $T_t = \exp(U_t) - I$. Differentiating it with respect to t, we obtain using the Lie algebra argument

$$\begin{aligned}\frac{dT_t}{dt} &= \exp(U_t)\int_0^1 \exp(-vU_t)\frac{dU_t}{dt}\exp(vU_t)dv \\ &= \exp(U_t)\int_0^1 \exp(-v\mathrm{ad}_{U_t})\dot{U}_t dv \\ &= \exp(U_t)\frac{I-\exp(-\mathrm{ad}_{U_t})}{ad_{U_t}}\dot{U}_t\end{aligned}$$

where $\mathrm{ad}_U V = UV - VU$ for two operators U and V and $\mathrm{ad}_U^n V$ is defined recursively:

$$\mathrm{ad}_U^n V = \mathrm{ad}_U(\mathrm{ad}_U^{n-1}V)\,.$$

We refer to[10] for more detail about the above computation.

From the above we have

$$\begin{aligned}\dot{U}_t &= \frac{ad_{U_t}}{I-\exp(-\mathrm{ad}_{U_t})}\left(\exp(-U_t)\dot{T}_t\right) \\ &= \frac{ad_{U_t}}{I-\exp(-\mathrm{ad}_{U_t})}\left([I+T_t]^{-1}\dot{T}_t\right)\end{aligned}$$

Therefore we obtain

Theorem 3.3

$$\det(I-T_t)^{-1/2} = \exp\left\{-\frac{1}{2}\ \mathit{Tr}(U_t)\right\}\,, \tag{3.8}$$

where U_t is the solution of the following operator equation:

$$\dot{U}_t = \frac{\mathrm{ad}_{U_t}}{I-\exp(-\mathrm{ad}_{U_t})}\,[I+T_t]^{-1}\,\dot{T}_t \tag{3.9}$$

It is clear that the solution of (3.9) is $U_t = \ln(I-T_t)$. However, we expect that (3.9) may provide some new way to compute the logarithm of an operator. For example, one may use (3.9) to construct time discretization schemes to approximate U_t.

1. Black F. and Scholes, M. The pricing of options and corporate liabilities, Journal of Political Economy 81 (1973), 637-654.
2. Baillie, R. T. Long memory processes and fractional integration in econometrics. J. Econometrics 73 (1996), no. 1, 5–59.

3. Comte, F.; Renault, E. Long memory continuous time models. J. Econometrics 73 (1996), no. 1, 101–149.
4. Comte, F.; Renault, E. Long memory in continuous-time stochastic volatility models. Math. Finance 8 (1998), no. 4, 291–323.
5. Ding, Z.; Granger, C. W. J. and Engle, R.F. A long memory property of stock market returns and a new model, Journal of Empirical Finance, 1 (1993), 83-108.
6. Ding, Z.; Granger, C. W. J. Modeling volatility persistence of speculative returns: a new approach. J. Econometrics 73 (1996), no. 1, 185–215.
7. Decreusefond, L.; Üstünel, A. S. Stochastic analysis of the fractional Brownian motion. Potential Anal. 10 (1999), no. 2, 177–214.
8. Duncan, T. E.; Hu, Y.Z.; Pasik-Duncan, B. Stochastic calculus for fractional Brownian motion. I. Theory. SIAM J. Control Optim. 38 (2000), 582-612
9. Fouque, J.-P.; Papanicolaou, G. and Sircar, K. R. Derivatives in financial markets with stochastic volatility. Cambridge University Press, Cambridge, 2000.
10. Hu, Y. Z. Calculs formels sur les EDS de Stratonovitch. Séminaire de Probabilités, XXIV, 1988/89, 453-460, Lecture Notes in Math., 1426, Springer, Berlin, 1990.
11. Hu, Y.Z. On the differentiability of functions of an operator. Séminaire de Probabilités, XXIX, 218-219, Lecture Notes in Math., 1613, Springer, Berlin, 1995.
12. Hu, Y. Z.; Meyer, P.-A. Chaos de Wiener et intégrale de Feynman. Séminaire de Probabilités, XXII, 51-71, Lecture Notes in Math., 1321, Springer, Berlin, 1988.
13. Hu, Y. Z.; Meyer, P. A. On the approximation of multiple Stratonovich integrals. Stochastic processes, 141-147, Springer, New York, 1993.
14. Hu, Y.Z. and Øksendal, B. Fractional white noise calculus and applications to finance. Preprint of University of Oslo, 1999.
15. Kallianpur, G. and Karandikar, R. L. Introduction to option pricing theory. Birkhäuser Boston, Inc., Boston, MA, 2000.
16. Merton, R. C. Theory of rational option pricing. Bell J. Econom. and Management Sci. 4 (1973), 141-183.
17. Mandelbrot B.B. and Van Ness J.W. Fractional Brownian motions, fractional noises and applications. SIAM Rev. 10 (1968) 422-437.
18. Rogers, L.C.G. Arbitrage with fractional Brownian motion. Math. Finance 7 (1997), 95–105.
19. Shiryaev, A. On arbitrage and replication for fractal models: In A. Shiryaev and A. Sulem (editors): Workshop on Mathematical Finance,

INRIA, Paris 1998.
20. Stein, E. and Stein, J. Stock price distributions with stochastic volatility: An analytic approach. Review of Financial studies 4 (1991), 727-752.

Optimal Investment and Consumption with Fixed and Proportional Transaction Costs

Hong Liu[a]
The Olin School of Business, Washington University
St. Louis, MO 63130
liuh@olin.wustl.edu

Abstract
We consider the optimal investment and consumption policy for a constant absolute risk averse investor who faces fixed and/or proportional transaction costs when trading a stock and maximizes his expected utility from intertemporal consumption. We show that the Hamilton-Jacobi-Bellman PDE with free boundaries can be reduced to an ODE, which greatly simplifies the problem. Using the stochastic impulse and singular control techniques, we then derive the optimal investment and consumption policy. In particular, when there are both fixed and proportional costs, it is shown that the optimal stock investment policy is to keep the dollar amount invested in the stock between two constant levels and upon reaching these two thresholds, the investor jumps to the corresponding optimal target level.

Keyword: transaction cost, investment, consumption.

1 Introduction

This paper studies the optimal investment and consumption policy for a constant absolute risk averse investor. The investor faces fixed and/or proportional transaction costs in trading a stock and maximizes his expected utility from intertemporal consumption.

When there are only proportional costs, the problem we study amounts to a stochastic singular control, as in Davis and Norman (1990) and Cuoco and Liu (2000). On the other hand, when there are fixed costs, the problem amounts to a stochastic impulse control problem, as in Eastham and Hastings (1988), Morton and Pliska (1995) and Korn (1998). We show that the Hamilton-Jacobi-Bellman PDE with free boundaries associated with this problem can be reduced to an ODE with free boundaries. Because of this reduction from PDE to ODE, this paper, in contrast to the existing literature, can explicitly specify the form of the optimal investment and consumption policy up to some constants.

Also related are papers which assume quasi-fixed transaction costs. Duffie and Sun (1990) and Morton and Pliska (1995) assume the trading costs is proportional to the total wealth at the time of the trade. As in the presence of fixed costs, the presence of quasi-fixed costs also leads to an optimal impulse control problem. While the assumption of quasi-fixed cost simplifies analysis, it is at best an approximation of the fixed cost faced by most investors.

2 The Model

2.1 The Asset Market

Throughout this paper we assume a probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\{\mathcal{F}_t\}$. Uncertainty in the model is generated by a standard one dimensional Brownian motion w. We assume that w_t is adapted.

There are two assets our investor can trade. The first asset ("the bond") is a money market account growing at a continuously compounded, constant rate $r > 0$. The second asset ("the stock") is risky. The investor can buy the stock at the ask price S_t and sell the stock at the bid price $(1-\alpha)S_t$, where $0 \leq \alpha < 1$ represents the proportional transaction cost rate. In addition, the investor has to pay a fixed brokerage fee $F \geq 0$ for each transaction (for both purchases and sales) in the stock. The stock pays dividend at a rate of $\delta \geq 0$. The ex-dividend stock price S_t follows a geometric Brownian motion process:

$$dS_t = \mu S_t dt + \sigma S_t dw_t, \tag{1}$$

where all parameters are assumed to be constants, the expected return of the stock is greater than the interest $(\mu + \delta > r)$ and $\sigma > 0$.

The investor derives utility from the intertemporal consumption c. All the consumption purchases and transaction cost payments are made from the bank account.

When $\alpha + F > 0$, the above model gives rise to equations governing the evolution of the amount invested in the bond x_t, and the amount invested in the stock y_t:

$$dx_t = rx_t dt + \delta y_t dt - c_t dt - dI_t + (1-\alpha)dD_t - F(1_{\{dI_t>0\}} + 1_{\{dD_t>0\}}), \tag{2}$$

$$dy_t = \mu y_t dt + \sigma y_t dw_t + dI_t - dD_t, \tag{3}$$

where the processes D and I represent the cumulative dollar amount of sales and purchases of the stock respectively. These processes are nondecreasing, right continuous adapted processes with $D(0) = I(0) = 0$. $x_0 > 0$ and $y_0 \geq 0$ are the given initial positions in the bond and the stock respectively.

2.2 The Investor's Problem

There is a single perishable consumption good (the numeraire). We assume that the investor derives his utility from intertemporal consumption of this good. In addition, as in Merton (1969), Wang (1993) and Lo, Mamaysky and Wang (2000), we assume that the investor has a constant absolute risk aversion preference, i.e., $u(c) = -e^{-\beta c}$ for some $\beta > 0$. We let $\Theta(x_0, y_0)$ denote the set of admissible trading strategies (I, D, c) such that (2) and (3) are satisfied and in addition

$$E[\int_0^\infty e^{-\rho t - r\beta x_t} dt < \infty, \tag{4}$$

where the last condition is imposed, as in Lo, Mamaysky and Wang (2000), to rule out any arbitrage opportunity such as the doubling strategy or Ponzi scheme.

The investor's problem is to choose admissible trading strategies D, I and c to maximize $E[\int_0^\infty e^{-\rho t} u(c_t) dt]$ subject to (2), (3) and (4). We define the value function at time t as

$$v(x, y) = \sup_{(I,D,c)\in\Theta(x,y)} E[\int_t^\infty e^{-\rho(s-t)}(-e^{-\beta c_s}) ds | \mathcal{F}_t, x_t = x, y_t = y]. \tag{5}$$

3 The Proportional Transaction Cost Case

In the case where $\alpha > 0$ and $F = 0$, stock trading will be infrequent. In this section we provide a heuristic derivation of the optimal policy. All the proofs are omitted to save space and available upon request. Similar to Davis and Norman (1990) and Liu and Loewenstein (2001), the trading region splits into three regions: Buy, Sell and No Transaction (NT). In the Buy region, the investor buys until reaching the closest NT boundary. In the Sell region, the investor sells until reaching the closest NT boundary. Inside NT, the investor does not trade. Therefore in NT, the value function must satisfy the HJB equation

$$\max_c \left(\frac{1}{2}\sigma^2 y^2 v_{yy} + \mu y v_y + r x v_x + \delta y v_x - c v_x - \rho v - e^{-\beta c}\right) = 0 \tag{6}$$

The optimal consumption is thus

$$c^* = -\frac{1}{\beta} \log\left(\frac{v_x}{\beta}\right), \tag{7}$$

which implies that (6) becomes

$$\frac{1}{2}\sigma^2 y^2 v_{yy} + \mu y v_y + r x v_x + \delta y v_x + \frac{v_x}{\beta} \log\left(\frac{v_x}{\beta}\right) - \rho v - \frac{v_x}{\beta} = 0. \tag{8}$$

For the convenience of exposition, let $z = r\beta y$ be the scaled amount in the stock. From now on, we will refer z as the (scaled) amount in stock. We conjecture that

$$v(x,y) = -\frac{1}{r}e^{-r\beta x - \varphi(r\beta y)}, \tag{9}$$

for a function $\varphi : \mathbb{R} \to \mathbb{R}$. Let another function ψ be the restriction of φ in the no transaction region, i.e., $\varphi(z) \equiv \psi(z)$ for any z in NT. (8) then becomes a nonlinear ODE:

$$\frac{1}{2}\sigma^2 z^2 \psi_{zz} - \frac{1}{2}\sigma^2 z^2 \psi_z^2 + \mu z \psi_z - r\psi + \delta z + (\rho - r) = 0. \tag{10}$$

We note that the above ODE is independent of the amount in the bond x. This suggests that if the boundary conditions can also be reduced to conditions in terms of only z, then the optimal stock transaction policy will depend only on the amount in the stock, but not the holding in the bond. We will show that this is indeed the case. We thus conjecture that the optimal stock transaction policy is characterized by two critical numbers $\underline{z}$ and $\bar{z}$. When the (scaled) amount in the stock z is below $\underline{z}$, the investor buys enough to reach $\underline{z}$. When it is above $\bar{z}$, the investor sells enough to reach $\bar{z}$.

Now let us consider the conditions in the transaction region. Because in the Buy region, the proposed transaction policy is to transact immediately to the closest NT boundary, the marginal (indirect) utility from the bond must be always equal to the marginal utility from the stock holding. Therefore, the differential equation in the Buy region is

$$v_y = v_x \tag{11}$$

and similarly, in the Sell region the differential equation is

$$v_y = (1-\alpha)v_x. \tag{12}$$

In addition, by the optimality of the boundaries (see Duma (1991)), v is a C^2 function in all regions.

The proposed transaction policy and the C^2 property (see Dumas (1990)) of the value function then implies the following six boundary conditions in terms of ψ:

$$\psi(\underline{z}) = C_1 + \underline{z}, \tag{13}$$

$$\psi_z(\underline{z}) = 1, \tag{14}$$

$$\psi_{zz}(\underline{z}) = 0, \tag{15}$$

$$\psi(\bar{z}) = C_2 + (1-\alpha)\bar{z}, \tag{16}$$

$$\psi_z(\bar{z}) = 1 - \alpha, \tag{17}$$

$$\psi_{zz}(\bar{z}) = 0. \tag{18}$$

Therefore indeed all the boundary conditions (13)-(18) are all independent of the holdings in the bond. Thus the above conjecture on the form of the optimal policy is justified.

The above free boundary problem can be easily solved by the following scheme: first, fix $\underline{z}$ (the no transaction cost optimal amount is a natural start), solve the ODE (10) subject to (14) and (15); then solve (17) for $\bar{z}$; finally, check if (18) is satisfied. If it is then the solution is found, otherwise, repeat the above procedure until (18) is satisfied. It turns out that this one dimensional search can be easily done.

Let

$$\varphi(z) = \begin{cases} C_2 + (1-\alpha)z & \text{if } z \geq \bar{z} \\ \psi(z) & \text{if } \underline{z} < z < \bar{z} \\ C_1 + z & \text{if } z \leq \underline{z}. \end{cases} \tag{19}$$

We then have the following result on the value function and the optimal trading strategy.

Theorem 1 *Suppose $F = 0$ and $\alpha > 0$. There exist constants C_1, C_2, $\underline{z}$ and $\bar{z}$ so that $\varphi(z)$ as defined in (19) is a C^2 function on $(-\infty, \infty)$ and $v(x, y) = -\frac{1}{r}e^{-r\beta x - \varphi(r\beta y)}$ is the value function. Moreover, the optimal transaction policy is to transact the minimal amount in order to maintain z between $\underline{z}$ and $\bar{z}$.*

We make the following assumptions about parameters for all the subsequent numerical illustrations throughout the rest of the paper unless otherwise stated. From Ibbotson and Sinquefeld (1982), we set the mean $\mu - r$ and the volatility σ of the annual excess return on the market portfolio to be 5.9% and 22% respectively. In addition, following Grossman and Laroque (1990), we set the real risk free rate r to be 1% and the time preference parameter ρ to be 0.01. For simplicity we focus on the case when the stock does not pay any dividend, i.e., $\delta = 0$. In Lo, Mamaysky and Wang (2000), they calibrate β to be between 0.0001 and 5.000. In all the subsequent numerical illustrations, we set it to the low end 0.001 to emphasize the effect of transaction costs. Of course, this is by no means a calibration for our model.

Figure 1 displays the optimal no transaction boundaries $\underline{z}$ and $\bar{z}$ as a function of the proportional transaction cost (The middle thin line is the Merton line). Without transaction cost ($\alpha = 0$), the investor will always keep \$121,900 in the stock (note that this is the actual amount which is equal to the scaled amount in the figure divided by $r\beta$), which is represented by the middle thin line. In the presence of the transaction cost, it is no longer optimal to always maintain a fixed amount in the stock. Instead, the investor will let the amount in the stock to fluctuate within a range. When $\alpha = 0.01$, for example, the investor will not adjust the amount he invests in the stock until it reaches \$99,400 or \$144,700. Thus the presence of transaction cost has a significant impact. Also note that as the transaction cost increases the Buy boundary comes down and the Sell boundary goes up.

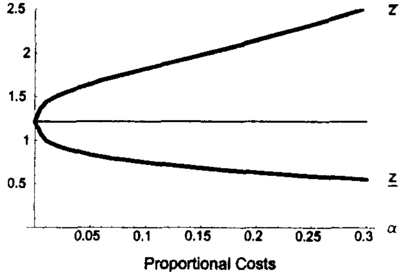


Figure 1: Boundaries as functions of the proportional transaction cost.

4 The Fixed Transaction Cost Case

In the presence of only the fixed cost ($F > 0$ but $\alpha = 0$), we conjecture that the optimal policy is characterized by three (instead of two as in the previous section) critical numbers: $\underline{y}, y^*$ and $\bar{y}$. When the amount in the stock reached the lower bound $\underline{y}$ or the upper bound $\bar{y}$, it is optimal to transact to y^*.

In the no transaction region, the HJB equation (6) in the previous section still holds. In addition, the forms for the value function in the transaction regions, i.e., where $y \leq \underline{y}$ or $y \geq \bar{y}$, are the same as in the previous section. What are different are the boundary conditions. Because of the presence of the fixed cost, the value function is generally not C^2 across the boundaries. However, it is still C^1.

In terms of ψ and the scaled amount $\underline{z} = r\beta\underline{y}$, $z^* = r\beta y^*$ and $\bar{z} = r\beta\bar{y}$, we have the following seven boundary conditions:

$$\psi(\underline{z}) = C_1 + \underline{z}, \tag{20}$$

$$\psi_z(\underline{z}) = 1, \tag{21}$$

$$\psi_z(z^*) = 1, \tag{22}$$

$$\psi(\bar{z}) = C_2 + \bar{z}, \tag{23}$$

$$\psi_z(\bar{z}) = 1, \tag{24}$$

$$\psi(z^*) = C_1 + r\beta F + z^*, \tag{25}$$

and

$$\psi(z^*) = C_2 + r\beta F + z^*. \tag{26}$$

The following theorem records results on the value function and the optimal trading strategy.

Theorem 2 *Assume $F > 0$ and $\alpha = 0$. Suppose there exist constants C_1, C_2, $\underline{z}$, z^* and $\bar{z}$ such that $\varphi(z)$ as defined in (19) with ψ being a solution of (10) subject to (20)-(26) is C^2 in $\mathbb{R} - \{\underline{z}, \bar{z}\}$ and in addition,*

$$\underline{z} < \frac{(\mu - r + \delta) - \sqrt{(\mu - r + \delta)^2 - 2\sigma^2(r(\psi(z^*) - z^* - r\beta F) - \rho + r)}}{\sigma^2}, \tag{27}$$

$$\bar{z} > \frac{(\mu - r + \delta) + \sqrt{(\mu - r + \delta)^2 - 2\sigma^2(r(\psi(z^*) - z^* - r\beta F) - \rho + r)}}{\sigma^2}, \tag{28}$$

$$\psi_z(z) \le 1, \quad \forall z \ge z^*, \tag{29}$$

$$\psi_z(z) \ge 1, \quad \forall z \le z^*. \tag{30}$$

Then

$$v(x, y) = -\frac{1}{r} e^{-r\beta x - \varphi(r\beta y)}$$

is the value function. Moreover, the optimal transaction policy is to transact the minimal amount in order to reach z^* only when $z < \underline{z}$ or $z > \bar{z}$.

Figure 2 displays the optimal no transaction boundaries $\underline{z}$ and $\bar{z}$ and the optimal target z^* as a function of the fixed cost. In the presence of the fixed transaction cost, it is no longer optimal to transact an infinitesimal amount to keep the amount in the stock in a range. When $F = \$5$, for example, the investor will allow the actual amount in the stock to fluctuate between \$102,000 and \$139,800. If the actual amount reaches \$105,200, the investor will buy \$16,600 worth of stock. On the other hand, if the actual amount reaches \$139,800, the investor will sell \$18,000 worth of stock. Thus the presence of fixed transaction cost also has a significant impact. In addition, note that as in the previous case, as transaction cost increases the Buy boundary comes down and the Sell boundary goes up. However, the sensitivity of the optimal target y^* to the change in the transaction costs is very small. It only decreases from \$121,900 to \$121,500 as the fixed cost increases from 0 to \$30, making z^* indistinguishable from the Merton line in the figure.

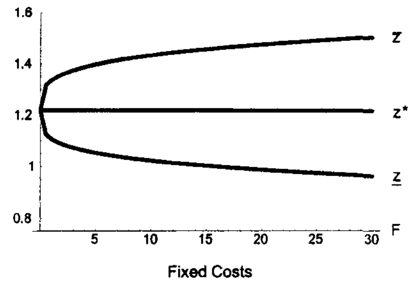


Figure 2: Boundaries as functions of the fixed transaction costs.

5 The Fixed and Proportional Transaction Cost Case

When both F and α are positive, i.e., there exists fixed and proportional costs for each transaction, the problem is even more complicated. We conjecture that in the presence of fixed and proportional transaction costs, there exist four (instead of three in the previous section) critical numbers $\underline{y}, \underline{y}^*, \bar{y}^*$ and $\bar{y}$ with $\underline{y} < \underline{y}^* < \bar{y}^* < \bar{y}$ to characterize the optimal trading strategy. To be specific, the optimal policy is to transact immediately to the Buy target boundary $\underline{y}^*$ if $y \leq \underline{y}$ and jump to the Sell target boundary $\bar{y}^*$ if $y \geq \bar{y}$.

Let $\underline{z} = r\beta\underline{y}, \underline{z}^* = r\beta\underline{y}^*, \bar{z}^* = r\beta\bar{y}^*, \bar{z} = r\beta\bar{y}$ denote the scaled boundaries. The transaction policy and the C^1 property of the value function then implies the following eight boundary conditions in terms of ψ:

$$\psi(\underline{z}) = C_1 + \underline{z}, \tag{31}$$

$$\psi_z(\underline{z}) = 1, \tag{32}$$

$$\psi_z(\underline{z}^*) = 1, \tag{33}$$

$$\psi_z(\bar{z}) = 1 - \alpha, \tag{34}$$

$$\psi_z(\bar{z}^*) = 1 - \alpha, \tag{35}$$

$$\psi(\bar{z}) = C_2 + (1 - \alpha)\bar{z}, \tag{36}$$

$$\psi(\underline{z}^*) = C_1 + r\beta F + \underline{z}^*, \tag{37}$$

$$\psi(\bar{z}^*) = C_2 + r\beta F + (1 - \alpha)\bar{z}^*. \tag{38}$$

Therefore we need to find $\underline{z}, \underline{z}^*, \bar{z}^*, \bar{z}, C_1$ and C_2 such that the ODE (10) is satisfied and these eight boundary conditions are satisfied.

Let φ be as defined in (19). We then have the following result on the value function and the optimal trading strategy.

Theorem 3 *Assume $F > 0$ and $\alpha > 0$. Suppose there exist constants C_1, C_2, $\underline{z}$, $\underline{z}^*$, $\bar{z}^*$ and $\bar{z}$ such that $\varphi(z)$ as defined in (19) with ψ being a solution of (10) subject to (31)-(38) is C^2 in $\mathbb{R} - \{\underline{z}, \bar{z}\}$ and in addition,*

$$\underline{z} < \frac{(\mu - r + \delta) - \sqrt{(\mu - r + \delta)^2 - 2\sigma^2(r(\psi(\underline{z}^*) - \underline{z}^* - r\beta F) - \rho + r)}}{\sigma^2} \tag{39}$$

$$\bar{z} > \frac{(\mu - r + \frac{\delta}{1-\alpha}) - \sqrt{(\mu - r + \frac{\delta}{1-\alpha})^2 - 2\sigma^2(r(\psi(\bar{z}^*) - (1-\alpha)\bar{z}^* - r\beta F) - \rho + r)}}{(1-\alpha)\sigma^2}, \tag{40}$$

$$1 - \alpha < \psi_z(z) < 1, \quad \forall \underline{z}^* < z < \bar{z}^* \tag{41}$$

$$\psi_z(z) \leq 1 - \alpha, \quad \forall z \geq \bar{z}^* \tag{42}$$

$$\psi_z(z) \geq 1, \quad \forall z \leq \underline{z} \tag{43}$$

Then

$$v(x, y) = -\frac{1}{r}e^{-r\beta x - \varphi(r\beta y)}$$

is the value function. Moreover, the optimal transaction policy is to transact the minimal amount in order to reach $\underline{z}^*$ if $z \leq \underline{z}$ and to reach $\bar{z}^*$ if $z \geq \bar{z}$.

Figure 3 shows the optimal boundaries $\underline{z}$, $\underline{z}^*$, $\bar{z}^*$ and $\bar{z}$ as functions of the fixed cost, setting $\alpha = 0.01$. In the presence of the fixed and proportional transaction cost, it is no longer optimal to jump to the same boundary as in the previous section when transacting. If $F = \$5$, for example, the investor will buy \$10,800 worth of stock when the actual amount reaches \$93,500. If the market goes up and the actual amount increases to \$152,600, the investor will sell \$14,300 worth of stock. In addition, note that as the fixed cost decreases to zero, $\underline{z}$ and $\underline{z}^*$ approaches the $\underline{z}$ in the case with no fixed cost and similarly for $\bar{z}$ and $\bar{z}^*$.

Figure 4 shows the optimal boundaries $\underline{z}$, $\underline{z}^*$, $\bar{z}^*$ and $\bar{z}$ as functions of the proportional cost, setting $F = \$5$. If $\alpha = 0.05$, for example, the investor will buy \$8,200 worth of stock when the actual amount reaches \$79,600. If the market goes up and the actual amount increases to \$171,900, the investor will sell \$13,500 worth of stock. As the proportional transaction cost increases, both the size of a purchase after reaching the Buy boundary $\underline{z}$ and the size of a sale after reaching the Sell boundary $\bar{y}$ decrease. In addition, note that as the proportional cost approaches zero, $\underline{z}^*$ and $\bar{z}^*$ approaches the z^* in the case with no proportional cost.

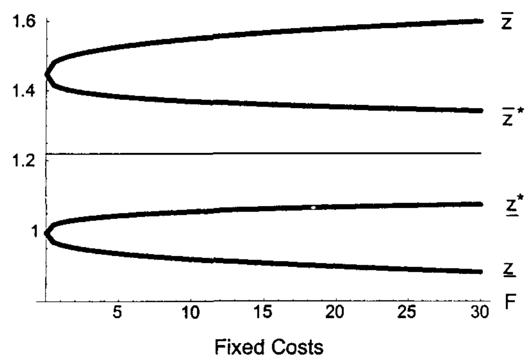


Figure 3: Boundaries as functions of the fixed transaction costs.

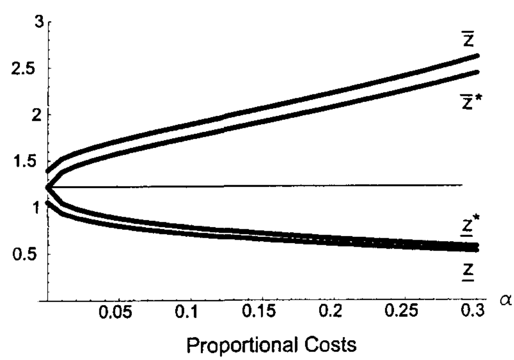


Figure 4: Boundaries as functions of the proportional transaction costs.

6 Concluding Remarks

The approach of this paper may be extended to some other models with both fixed and proportional transaction costs. First, if a CARA investor derives utility from intertemporal consumption and terminal wealth and has an exponentially distributed horizon, all main results of this paper seem still valid. Second, if a CARA investor with an exponentially distributed horizon derives utility only from terminal wealth and the interest rate is zero, a closed form solution (up to a Gamma function and several constants to be determined by solving nonlinear algebraic equations.) can be obtained following the same derivation specified in this paper. Finally, using the randomization methodology proposed by Liu and Loewenstein (2001), we can approximate the solution to a finite horizon optimal consumption and investment problem with fixed and proportional transaction costs.

References

CUOCO, D., AND H. LIU. 2000. Optimal consumption of a divisible durable good. *Journal of Economic Dynamics and Control* 24, 561-613.

DAVIS, M. H. A., AND A.R. NORMAN. 1990. Portfolio selection with transaction costs. *Mathematics of Operations Research* 15, 676-713.

DUFFIE, D., AND T. SUN. 1990. Transaction costs and portfolio choice in a discrete-continuous time setting. *Journal of Economic Dynamics and Control* 14, 35-51.

DUMAS, B. 1991. Super contact and related optimality conditions. *Journal of Economic Dynamics and Control* 4, 675-685.

EASTHAM, J. F., AND K. J. HASTINGS. 1988. Optimal impulse control of portfolios. *Mathematics of Operation Research* 13, 588-605.

GROSSMAN, S. J., AND G. LAROQUE. 1990. Asset pricing and optimal portfolio choice in the presence of illiquid durable consumption goods. *Econometrica* 58, 23-51.

IBBOTSON, R. AND R. SINQUEFELD. 1982. Stocks, bonds, bills and inflation: the past and the future, 1982 edition, Charlottesville: Financial Research Analyst's Foundation.

KORN, R. 1998. Portfolio optimization with strictly positive transaction costs and impulse controls. *Finance and Stochastics.* 2, 85-114.

LIU, H., AND M. LOEWENSTEIN. 2001. Optimal portfolio selection with transaction costs and finite horizons. Forthcoming in *Review of Financial Studies.*

LO, A., H. MAMAYSKY AND J. WANG. 2000. Asset prices and trading volume under fixed transaction costs. Working paper, MIT.

MERTON, R. C. 1969. Lifetime portfolio selection under uncertainty: the continuous-time case. *Review of Economics and Statistics* 51, 247-257.

MORTON, A. J., AND S. R. PLISKA 1995. Optimal portfolio management with fixed transaction costs. *Mathematical Finance* 5, 337-356.

WANG, J 1993. A model of intertemporal asset prices under asymmetric information. *Review of Economic Studies* 60, 249-282.

Sharp Estimates of Ruin Probabilities for Insurance Models Involving Investments

Jin Ma [a]
Department of Mathematics, Purdue University,
West Lafayette, IN 47907-1395, USA
E-mail: majin@math.purdue.edu

Xiaodong Sun
Department of Mathematics, Purdue University,
West Lafayette, IN 47907-1395, USA
E-mail: xsun@math.purdue.edu

In this note we introduce a general method for estimating ruin probabilities for insurance models that allows the insurance company to invest in a financial market. Our method is based on a new type of exponential martingale parametrized by a rate function. We show by examples that many existing Lundberg-type bounds can be reduced to finding an appropriate rate function. To study the asymptotics of the Lundberg bounds in such a general setting, we establish the relation between the ruin probability and a special type of *storage process* characterized by a generalized reflected SDE with discontinuous paths. Based on such a relation we use large deviation techniques to derive, in some special but non trivial cases, the limiting behavior of ruin probabilities, as well as the adjustment coefficients.

1 Introduction

We consider the classical ruin problems with general risk reserve process of the following form:

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \langle h_s, dW_s \rangle - \int_0^{t+} \int_{\mathbb{R}+} f(s, x, \cdot) N_p(dsdx), \quad (1.1)$$

where $b(s,x) = r_s x + \rho(s,x) + a_s$ is a (random) function, W is a d-dimensional Brownian motion, and p is a stationary Poisson point process of class (QL) and N_p is its corresponding counting measure. We assume that the processes r, a and h are adapted to the filtration $\{\mathcal{F}_t^W \vee \mathcal{F}_t^p\}_{t\geq 0}$.

The risk model (1.1) contains many existing ones as special case, as we shall see in the §2. But the most important feature of (1.1) is that it contains the case when the insurance company invests its reserve in a financial market with contains both risky and riskless assets. In that case, r is the interest rate of the money market, $\boldsymbol{a_s} = \langle \pi_s, \theta_s \rangle$, where $\boldsymbol{\pi}$ **is** the (stock) portfolio process, and θ is the "risk premium", and $h_s = \sigma_s^T \pi_s$, where σ is the volatility matrix of the market (for full derivation, see Ma and Sun[9]).

The main results of this report contains two parts. First, we introduce a general *exponential martingale*, parametrized by a "rate function". We show by examples that various known Cramér-Lundberg-type bounds studied separately before with different methods can all be derived by this method via carefully choosing the rate functions. For example, if the rate function is of the form $I_\delta(t,x) = \delta x \exp\{-\int_0^t r_s ds\}$, with δ being a parameter, then we can derive the Lundberg bounds for classical models, discounted reserve models, and perturbed reserve models; Setting $I(t,x) = I(x) = \int_0^t \gamma(z)dz$, then we can recover the Lundberg bound of Asmussen and Nielsen[1] and determine the "local adjustment coefficients". We shall prove that, by solving a first or second order integro-differential equaiton/inequality for the rate function I, we can derive some even sharper bounds, including those discovered recently by Nyrhinen[12] and Kalashnikov-Norberg[7]. Therefore, this method provides a unified framework for estimating the Lundberg bounds of ruin probability.

In the second part of the paper we establish a relation between the ruin probability and a *storage-type process*, in the spirit of that in Asmussen and Petersen[2]. We show that, in light of the general theory of storage processes (see, e.g., Prabhu[14]), the storage process corresponding to our general ruin problem takes the form of a *stochastic differential equation with discountinuoius paths and reflecting boundary conditions* (SDEDR, for short), similar to the one studied by Ma[8]. We then use this relation to derive the adjustment coefficients in some special, but non-trivial ruin problems for general insurance models, by investigating the asymptotics of ruin probabilities using large deviation techniques. The idea of such technique was used by, for example, Djehiche[5] and Martin-Löf[11], but the storage process of this kind and the result in such generality seem to be new.

For mathematical clarity throughout this paper we assume all the uncertainties or randomness come from a common, complete probability space $(\Omega, \mathcal{F}, P)$, on which is defined a d-dimensional Brownian motion W (source of uncertainty in the financial market), and a Poisson point process p (source of uncertainty of the claim process), independent of W. We define $\mathbf{F}^p \triangleq \{\mathcal{F}_t^p\}_{t\geq 0}$, $\mathbf{F}^W \triangleq \{\mathcal{F}_t^W\}_{t\geq 0}$, and $\mathbf{F} \triangleq \{\mathcal{F}_t^p \vee \mathcal{F}_t^W\}_{t\geq 0}$, where $\mathbf{F}^p$ and $\mathbf{F}^W$ are the (augmented) natural filtrations generated by p and W, respectively. We denote the *compensator* of N_p by $\hat{N}_p(dtdx) = E(N_p(dtdx)) = \nu(dx)dt$, where $\nu(dx)$ is the characteristic measure of p, and the *compensated* random measure $\tilde{N}_p \triangleq N_p - \hat{N}_p$. Finally, we denote F_p to be the space of all random fields $f(t,x,\omega); [0,T] \times \mathrm{I\!R} \times \Omega \longmapsto \mathrm{I\!R}_+$ such that for fixed x, $f(\cdot, x, \cdot)$ is $\mathbf{F}^p$-

predictable, and that $\int_0^t \int_{\mathbb{R}_+} |f(s,x,\cdot)|\nu(dx)ds < \infty$, a.s.

We shall make use of the following *standing assumptions* throughout.

(A1) The random function $b : [0,T] \times \mathbb{R} \mapsto \mathbb{R}$ is continuous, adapted to $\mathbf{F}^W$ for each fixed $x \in \mathbb{R}$, and is uniformly Lipschitz in the variable x, uniformly in (t,ω).

(A2) The random field $f(\cdot,\cdot,\cdot) \in F_p$, such that $f(t,x,\cdot) \geq 0$, $\forall(t,x)$, a.s.; and that there exists a $\delta_0 > 0$,

$$\int_0^t \int_{\mathbb{R}_+} \exp\{\delta_0 f(s,x,\cdot)\}\nu(dx)ds < \infty, \quad \forall t \geq 0, \quad P\text{-a.s.}$$

2 Exponential (Local) Martingales

Recall that $b(t,x) = r_t x + \rho(t,x) + a_s$. Besides (A1), in this section we assume that r is positive and uniformly bounded, and ρ is deterministic. Furthermore, we assume that $\int_0^t [|a_s|^2 + |h_s|^2]ds < \infty$, $\forall t \in [0,T]$, a.s.

To construct the exponential martingale we proceed as follows. First, for any function $I \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ we define

$$Z_t^I \triangleq \int_0^t \int_{\mathbb{R}^+} [\exp\{I(s,X_s) - I(s,X_s - f(s,x,\cdot))\} - 1]v(dx)ds \quad (2.1)$$

$$V_t^I \triangleq \int_0^t \{\partial_x I(s,X_s)[r_s X_s + \rho(s,X_s) + a_s] + \partial_t I(s,X_s)\}ds \quad (2.2)$$

$$Y_t^I \triangleq \int_0^t \{(\partial_x I(s,X_s))^2 - \partial_{xx}^2 I(s,X_s)\}|h_s|^2 ds. \quad (2.3)$$

$$(2.4)$$

Denote

$$K_t^I \triangleq -V_t^I + \frac{1}{2}Y_t^I + Z_t^I, \quad \text{and} \quad L_t^I \triangleq \exp\{-I(t,X_t) - K_t^I\}, \quad t \geq 0. \quad (2.5)$$

Since the process X is càdlàg , one shows that $V_t^I < \infty$ and $Y_t^I < \infty$, $\forall t$, P-a.s. For Z^I, K^I and L^I, however, we need some more assumptions.

Definition 2.1 *A function $I \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ is called a rate function satisfying Hypothesis A, if the process Z^I defined by (2.1) satisfies $Z_t^I < \infty$ (hence $K_t^I < \infty$, $L_t^I < \infty$), $\forall t \geq 0$, P-a.s.*

Let I be a rate function satisfying Hypothesis A, and let $F^I(t,v,x,y,z) \triangleq \exp(v - I(t,x) - \frac{1}{2}y - z)$. Then clearly $L_t^I = F^I(t,V_t^I,X_t,Y_t^I,Z_t^I)$, $t \geq 0$.

Applying Itô's formula and noting the definitions of V^I, Y^I, and Z^I we have

$$L_t^I = L_0^I - \int_0^t F_x^I \langle h_s, dW_s\rangle + \int_0^{t^+}\int_{\mathbb{R}_+} \{e^{I(s,X_{s-})-I(s,X_{s-}-f(s,x,\cdot))} - 1\} F^I \tilde{N}_p(dxds).$$

The following theorem is therefore obvious.

Theorem 2.2 *Suppose that I is a rate function satisfying Hypothesis A. Then the process $\{L_t^I : t \geq 0\}$ is an* $\mathbf{F}$*-local martingale.* ■

An important example of the rate function I satisfying Hypothesis A is

$$I(t,x) = I_\delta(t,x) = \delta x e^{-\int_0^t r_s ds}, \tag{2.6}$$

where $\delta \in \mathbb{R}$ is some constant. Let us denote $\beta_t = -\int_0^t r_s ds$ to be the discount factor, and $\widetilde{X}_t = e^{\beta_t} X_t$ to be the discounted risk reserve. Then $I_\delta(t, X_t) = \delta \widetilde{X}_t$, and an easy application of Itô's formula shows that $\widetilde{X}$ satisfies the SDE:

$$\tilde{X}_t = x + \int_0^t e^{\beta_s}(\widetilde{\rho}(s,\widetilde{X}_s) + a_s)ds + \int_0^t e^{\beta_s}\langle h_s, \sigma_s dW_s\rangle - \int_0^{t^+} e^{\beta_s} dS_s, \tag{2.7}$$

where $\widetilde{\rho}(t,\widetilde{X}_t) = \rho(t, e^{-\beta_t}\widetilde{X}_t))$, and $dS_t = \int_{R+} f(t,x,\cdot)N_p(dtdx)$. Now for any $\delta > 0$, $\gamma \geq 0$, we define the following processes: for $t \geq 0$,

$$m_t^f(\gamma) \stackrel{\triangle}{=} \int_{\mathbb{R}_+} [e^{\gamma f(t,x,\omega)} - 1]\nu(dx); \; Z_t^\delta \stackrel{\triangle}{=} \int_0^t m_s^f(\delta e^{\beta_s})ds; \; Z_t^{\delta,0} \stackrel{\triangle}{=} \int_0^t m_s^f(\delta)ds, \tag{2.8}$$

and define the following two subsets of $\mathbb{R}_+$:

$$\mathcal{D} = \{\delta \geq 0 : Z_t^\delta < \infty, \forall t, \text{ a.s.}\}; \quad \mathcal{D}_0 = \{\delta \geq 0 : Z_t^{\delta,0} < \infty, \forall t, \text{ a.s.}\}. \tag{2.9}$$

Then $\mathcal{D}_0 \subseteq \mathcal{D}$, and for $\delta \in \mathcal{D}$ we can rewrite V^I, Y^I as

$$V_t^\delta = \delta \int_0^t e^{\beta_s}[\widetilde{\rho}(s,\widetilde{X}_s) + a_s]ds; \quad Y_t^\delta = \delta^2 \int_0^t e^{2\beta_s}|h_s|^2 ds. \tag{2.10}$$

Moreover, K^I becomes $K_t^\delta = -V_t^\delta + \frac{1}{2}Y_t^\delta + Z_t^\delta$, and L^I of (2.5) becomes $L_t^\delta \stackrel{\triangle}{=} \exp\{-\delta\widetilde{X}_t - K_t^\delta\}$, $t \geq 0$. In this case we have (see Ma and Sun[9]):

Theorem 2.3 *Suppose that the assumptions (A1)–(A2) hold. Then the process $\{L_t^\delta : t \geq 0\}$ enjoys the following properties:*

(i) For every $\delta \in \mathcal{D}$, $\{L_t^\delta : t \geq 0\}$ is an $\mathbf{F}$*-local martingale.*

(ii) If the processes a and h are bounded, and that f is deterministic, then for every $\delta \in \mathcal{D}_0$, $\{L_t^\delta : t \geq 0\}$ is an $\mathbf{F}$*-martingale.*

(iii) If in addition to (ii), r is also deterministic, then (ii) holds for all $\delta \in \mathcal{D}$. ■

3 Lundberg-Type Bounds

We now use the exponential martingales L^I and L^δ to derive various Lundberg-type bounds. First note that as a positive local martigale, L_t^I is a supermartingale. Thus applying the Optional Sampling Theorem with the stopping time $\tau \triangleq \inf\{t, X_t < 0\}$, and using the assumption on I as well as Jensen's inequality, one derives the following theorem (see Ma and Sun[9]).

Theorem 3.1 (Lundberg Bounds) *Assume that the rate function I satisfies Hypothesis A, such that $I(t,x) \le 0$, for all t and $x \le 0$. Then, it holds that*

$$\psi(x,T) \le e^{-I(0,x)} E \sup_{0\le t\le T} \exp(K_t^I), \tag{3.1}$$

$$\psi(x) \le e^{-I(0,x)} E \sup_{t\ge 0} \exp(K_t^I), \tag{3.2}$$

where K^I is defined by (2.5). ∎

The following modification of Theorem 3.1 is sometimes more convenient.

Corollary 3.2 *Assume all assumptions of Theorem 3.1 are in force. Then the following Lundberg bounds hold:*

$$\psi(x,T) \le e^{-I(0,x)} E \sup_{0\le t\le T} \exp(K_t^I(X^+)), \tag{3.3}$$

$$\psi(x) \le e^{-I(0,x)} E \sup_{t\ge 0} \exp(K_t^I(X^+)), \tag{3.4}$$

where $K^I(X^+)$ is the same as K^I by replacing all X by $X^+ \triangleq X \vee 0$. ∎

In the case when the rate function $I(t,x) = I_\delta(t,x)$ we have

Theorem 3.3 *Assume (A1) and (A2). Then, for every $\delta \in \mathcal{D}$, the ruin probabilities $\psi(x,T)$ and $\psi(x)$ have the following upper bounds:*

$$\psi(x,T) \le e^{-\delta x} E \sup_{0\le t\le T} \exp(K_t^\delta), \quad \psi(x) \le e^{-\delta x} E \sup_{t\ge 0} \exp(K_t^\delta). \tag{3.5}$$

Furthermore, if $\widetilde{\rho}(t,x) \ge c_\rho > 0$, then the process K^δ can be replaced by

$$\widetilde{K}_t^\delta \triangleq -\widetilde{V}_t^\delta + \frac{1}{2} Y_t^\delta + Z_t^\delta, \tag{3.6}$$

where $\widetilde{V}_t^\delta \triangleq \delta \int_0^t e^{\beta_s}(c_\rho + a_s)ds$, $t \ge 0$. Finally, if we define $\widetilde{\delta} = \sup\{\delta \in \mathcal{D} : E\{\sup_{t\ge 0} \exp(\widetilde{K}_t^\delta)\} < \infty\}$, then for any $\varepsilon > 0$ it holds that

$$\lim_{x\to\infty} \psi(x) e^{(\tilde{\delta}-\varepsilon)x} = 0. \tag{3.7}$$

■

We now use Theorems 3.1 and 3.3 to derive several existing Lundberg bounds.

Example 3.4 Assume that $b(t,x) = r_t x + c$ (i.e., $a_t \equiv 0$, $\rho(t,x) \equiv c$), and S_t is compound Poisson (i.e., $f(t,x,\cdot) = x$) with $\nu(dx) = \lambda F(dx)$. Then we have the following cases.

(i) $r \equiv 0$ and $h \equiv 0$, then (1.1) is reduced to the *classical* Cramér-Lundberg model. Applying Theorem 3.3 we see that $\psi(x) \le e^{-\widetilde{\delta} x}$, for all x, where $\widetilde{\delta} = \sup\left\{\delta : \int_0^\infty (e^{\delta x} - 1)\lambda F(dx) - c\delta \le 0\right\}$. This is the well-known Cramér-Lundberg inequality (see, for example, Grandall[6]), and $\widetilde{\delta}$ is called *Lundberg exponent.*

(ii) r_t is deterministic, and $h \equiv 0$. This is the so-called *Discounted risk reserve* model. Theorem 3.3 then leads to that

$$\psi(x) \le e^{-\widetilde{\delta} x} \sup_{t \ge 0} \exp(\widetilde{K}_t^{\widetilde{\delta}}),$$

where $\widetilde{K}_t^\delta = \int_0^t \left\{ \int_0^\infty [\exp(\delta e^{\beta_s} x) - 1]\lambda F(dx) - c e^{\beta_s} \right\} ds$, and $\widetilde{\delta} = \sup\{\delta \ge 0 : \sup_{t \ge 0} \widetilde{K}_t^\delta < \infty\}$. This coincides with Theorem 11.4.1 of Rolski et al[16].

(iii) $r_t \equiv 0$, $h_t \equiv \varepsilon$. This is the *Perturbed risk reserve* model. In this case we see that $\widetilde{K}_t^\delta = t(-c\delta + \frac{1}{2}\delta^2\varepsilon^2 + \int_0^\infty (e^{\delta x} - 1)\lambda F(dx))$, and we can show that the Lundberg exponent $\widetilde{\delta} \stackrel{\triangle}{=} \sup\{\delta > 0 : k(\delta) = 0\} < \infty$, provided the set $\{\delta > 0 : k(\delta) = 0\}$ is not empty, where

$$k(\delta) \stackrel{\triangle}{=} -c\delta + \frac{1}{2}\delta^2\varepsilon^2 + \int_0^\infty (e^{\delta x} - 1)\lambda F(dx). \tag{3.8}$$

This also recovers the standard Lundberg bound and Lundberg exponent in such a case(cf. e.g., Rolski et al.[16]).

We remark that a sufficient condition for the function $k(\cdot)$ in (3.8) to have positive root is the following "*net profit condition*" (see, e.g., Asmussen-Nielsen[1]):

$$c > \lambda E[U_1], \tag{3.9}$$

where U_1 is the jump size random variable of the compound Poisson process S. Such a condition is also useful in next example. ■

Example 3.5 (*Asmussen-Nielsen*) Assume that $h \equiv 0$, $a \equiv 0$, $r_t \equiv r$, $\rho(t,x) \equiv \rho(x)$ with $\rho'(x) > 0$, $f(t,x,\cdot) = x$, and $v(dx) = \lambda F(dx)$. Then, letting $p(x) \stackrel{\triangle}{=} rx + c(1+\rho(x))$, the net profit condition becomes $\inf_{x\geq 0} p(x) > \lambda E[U_1]$.

Letting $I(x) = \int_0^x \gamma(y)dy$ with $\gamma(\cdot)$ being non-decreasing, we have

$$K_t^I(X^+) \leq \int_0^t \Big\{-\gamma(X_s^+)p(X_s^+) + \int_{\mathbb{R}^+} [e^{\gamma(X_s^+)x} - 1]\lambda F(dx)\Big\}ds, \qquad (3.10)$$

for all $t \geq 0$. Thus if there exists a positive, non-decreasing solution $\gamma(\cdot)$ of the following Lundberg equation (γ is called the "*local adjustment coefficient*" by Asmussen and Nielsen[1]):

$$-\gamma p(y) + \int_{\mathbb{R}^+} [e^{\gamma x} - 1]\lambda F(dx) = 0, \qquad y \geq 0, \qquad (3.11)$$

Then one extend it to a rate function I satisfying Hypothesis A, and $I(x) \leq 0$ for all $x \leq 0$. Thus Corollary 3.2 implies that the Lundberg inequalities

$$\psi(x,T) \leq e^{-I(x)} \quad \text{and} \quad \psi(x) \leq e^{-I(x)}, \qquad x \geq 0.$$

which are the same as the results of Asmussen and Nielsen[1]. (We remark that the existence of the solution to (3.11) does exist under the net profit condition and the assumption that $\rho(\cdot)$ is non-decreasing.)

We observe that the inequality in (3.10) can be improved if we allow the rate function I to take a more general form. For example, let us assume that $\rho(x) \equiv 0$, and $F(x) = 1 - e^{-\theta x}$. Then $K^I(X^+)$ becomes

$$K_t^I(X^+) = \int_0^t \Big\{-I'(X_s^+)(rX_s^+ + c) + \int_0^\infty [e^{I(X_s^+) - I(X_s^+ - x)} - 1]\lambda\theta e^{-\theta x}dx\Big\}ds.$$

Thus if we let I be a solution to the following integro-differential equation

$$-I'(y)[ry + c] + \int_0^\infty [\exp\{I(y) - I(y-x)\} - 1]\lambda\theta e^{-\theta x}dx = 0. \qquad (3.12)$$

and modify it so that Corollary 3.2 can be applied, then we would still have the bound $\psi(x) \leq e^{-I(x)}$, $x \geq 0$. A direct computation then shows that the following function

$$I(y) = -\log\left\{\frac{\int_y^\infty e^{-\theta z}(1+\frac{rz}{c})^{(\frac{\lambda}{r})-1}dz}{\frac{c}{\lambda} + \int_0^\infty e^{-\theta z}(1+\frac{rz}{c})^{(\frac{\lambda}{r})-1}dz}\right\} \qquad (3.13)$$

is a solution to (3.12) for $y \geq 0$, thus the Lundberg bound $e^{-I(x)}$ is exactly the expression inside "$\log\{\cdots\}$" in (3.13). It is worth noting that this bound is indeed the sharpest, as it is the "true" ruin probability in this case (see Segerdahl[17]). Such a bound, however, does not seem to be amendable using the local adjustment coefficient method. ■

Example 3.6 We now change Example 3.5 slightly by allowing $h_t = \alpha X_t$ where $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$ is a constant vector with $\alpha_i \geq 0$, $\forall i$, such that $|\alpha| \triangleq \sum_i \alpha_i > 0$ (this corresponds to the so-called "Proportional investments" model). Let us assume that $r > \frac{1}{2}|\alpha|^2 > 0$ ("*compatibility condition*").

Similar to Examples 3.5, we can show that in this case

$$K_t^I(X^+) = \int_0^t \Big\{ - I'(X_s^+)p(X_s^+) + \frac{1}{2}(I'^2 - I'')(X_s^+)|\alpha|^2[X_s^+]^2 + \int_{\mathbb{R}+} [e^{I(X_s^+)-I(X_s^+-x)} - 1]\lambda\theta e^{-\theta x} dx \Big\} ds, \quad t \geq 0. \tag{3.14}$$

Therefore we would like to find a rate function $I \in C^2(\mathbb{R})$ that satisfies the following 2^{nd}-order integro-differential inequality:

$$G^I(y) \triangleq -I'(y)\{ry + C\} + \frac{1}{2}(I'(y)^2 - I''(y))y^2|\alpha|^2 + \int_{R+} [e^{I(y)-I(y-x)} - 1]\lambda\theta e^{-\theta x} dx \leq 0, \quad y \geq 0, \tag{3.15}$$

and we require that $I(y) \sim k \ln y + C$ for some constant k, C, as $y \to +\infty$.

It turns out (see Ma and Sun[9]) that this can be done by following the so-called "Principle of Smooth Fit": that is, one first chooses constants k and β so that the function

$$I_{\beta,k}(y) = k(\ln(y + \beta) - \ln 2\beta)1_{[\beta,\infty)}(y),$$

is C^2 and satisfies $G^I(y) \leq 0$ for $y \in [\beta, \infty)$, and then extend it to a C^2 function I such that $G^I(y) \leq 0$ for all y. We note that such a rate function will produce the so-called "power ruin probaility" as is seen in Nyrhinen[12] and Kalashnikov and Norberg[7]. ■

4 Relation to Storage Processes, and Large Deviation Results

Given the risk reserve process X (1.1), we consider the following storage-type process, defined by a special stochastic differential equation with discontinuous paths and reflections (SDEDR):

$$Y_t = -\int_0^t b(T-s,\cdot,Y_s)ds + \xi_t + K_t, \tag{4.1}$$

where $\xi_t \triangleq -\Lambda_T + \Lambda_{T-t} + S_T - S_{T-t}$, $S_t = \int_0^{t+}\int_{\mathbb{R}_+} f(\cdots)dN_p(dtdx)$, for all $t \geq 0$; and K_t is a *reflecting* process. The procise definition of the solution to SDEDR (4.1) is the following.

Let $\mathbb{D}_T$ be the space of all càdlàg functions on $[0,T]$, and $\mathbb{D}_T^0$ be the subspace of $\mathbb{D}_T$ consisting of all $y \in \mathbb{D}_T$ such that $y(0) = 0$. Let $\mathbb{D} = \mathbb{D}_\infty$ and $\mathbb{D}^0 = \mathbb{D}_\infty^0$.

Definition 4.1 *A pair of processes (Y,K) is called a solution to SDEDR (4.1), if it satisfies the following properties P-almost surely:*

i) $(Y,K) \in \mathbb{D}^2$ and (Y,K) satisfies (4.1);

ii) $Y_t \geq 0$, $\forall t \geq 0$;

iii) $\int_0^\infty Y_s dK_s = 0$;

iv) $\Delta K_t = |Y_t + \Delta\xi_t^\pi|$, $\forall t \in S_K \triangleq \{t \geq 0 : \Delta K_t \neq 0\}$.

Note that since the SDEDR (4.1) does not involve any stochastic integrals, there is no "adaptedness" requirement for the solution. In fact, it can be understood pathwisely as an ordinary differential equation with reflection. We have the following result (see Ma and Sun[10]).

Theorem 4.2 *Assume (A1) and (A2). Then*

(i) the SDEDR (4.1) has a (pathwise) unique solution;

(ii) the refelcting process K is continuous;

(iii) define the ruin time $\tau = \inf\{t : X_t < 0\}$, and assume that $X_0 = x > 0$, then for each $T > 0$, it holds that

$$\psi(x,T) = P\{\tau < T\} = P\{Y_T > x\}. \tag{4.2}$$

■

To make use of the relation (4.2), let us recall the asymptotic estimate (3.7) from Theorem 3.3 (the finite horizon case):

$$\lim_{x\to\infty} \psi(x,T)e^{(\tilde{\delta}_T - \varepsilon)x} = 0, \tag{4.3}$$

where $\widetilde{\delta} = \sup\{\delta \in \mathcal{D} : E\sup_{t\geq 0}\exp(\widetilde{K}_\delta^t) < \infty\}$. In order to show that $\tilde{\delta}_T$ is indeed the adjustment coefficient, one needs also show that for any $\varepsilon > 0$,

$$\lim_{x\to\infty} \psi(x,T)e^{(\tilde{\delta}_T + \varepsilon)x} = \infty. \tag{4.4}$$

Combining (4.3) and (4.4) we see that this is equivalent to

$$\lim_{x\to\infty}\frac{1}{x}\log\psi(x,T)=-\tilde{\delta}_T. \tag{4.5}$$

Namely it becomes a large deviation type problem on $\psi(\cdot,T)$! The relation (4.5) then turns the problem into a large deviation result for the storage process Y, which is much easier to handle.

Let us be more specific. Assume that S is compound Poisson, and the risk reserve follows an SDE:

$$X_t=x+\int_0^t b(s,x,X_s)ds+\int_0^t\langle h_s(x),dW_s\rangle-S_t. \tag{4.6}$$

In other words, we consider the case where process h_t and the "premium" rate process b depends on the initial reserve. (As examples we can consider the case of "proportional investments", or the case where the insurance company simply holds a constant portfolio throughout, based on the amount of initial reserve.)

Let us modify the assumption (A1) to the following.

(A3) The premium rate function $b(t,x,y)$ is uniformly Lipschitz in y, uniformly in (t,x,ω); and if we denote $b^\varepsilon=\varepsilon b(t,1/\varepsilon,y/\varepsilon)$, then there exists $\widetilde{b}\in C([0,T]\times\mathbb{R})$ such that

$$\lim_{\varepsilon\to0}\sup_{t,y}|\widetilde{b}(t,y)-b^\varepsilon(t,y)|=0.$$

(A4) For some constant $q>0$ and some function $\widetilde{\sigma}$, it holds that

$$\lim_{x\to\infty}\frac{h_t(x)}{x^q}=\widetilde{\sigma}_t,\qquad\text{uniformly in } t.$$

Denoting the solution of (4.1) to be $(Y(x),K(x))$, and setting $x=1/\varepsilon$. Then Theorem 4.2 tells us that (4.5) now becomes

$$\lim_{\varepsilon\to0}\varepsilon\log\psi(1/\varepsilon,T)=\lim_{\varepsilon\to0}\varepsilon\log P\{\varepsilon Y_T(1/\varepsilon)>1\}=-\widetilde{\delta}_T. \tag{4.7}$$

This is clearly a large deviation problem for the process $Y_T^\varepsilon\triangleq\varepsilon Y_T(1/\varepsilon)$.

Finite Horizon Case. Let us denote $H_1=\{g(t)=\int_0^t f(s)ds,\ t\in[0,T]:\ f\in L^2[0,T]\}$, with $\|g\|_{H_1}=\{\int_0^T|\dot{g}(t)|^2dt\}^{\frac{1}{2}}$, and denote $F:\mathbb{D}_T^0\mapsto\mathbb{D}_T^0$ to be the solution mapping of the following ODE with reflection: for $g\in\mathbb{D}_T^0$,

$$f(t)=F[g](t)=-\int_0^t\widetilde{b}(T-s,f(s))ds+g(t)+k(t),\quad t\geq0.$$

Let the moment generating function of S_t be $M_t = E[e^{rS_t}] = e^{\lambda t g(r)}$, where $g(r) = \int_0^\infty (e^{ry}-1)F(dy)$; and let $G(x) \stackrel{\triangle}{=} \sup_{r\in\mathbb{R}}[rx - g(r)]$ be the *Fenchel-Legendre transformation* of g. Then, using the relation (4.2) we can prove the following result (Ma and Sun[10]).

Theorem 4.3 *Assume (A3); and that(A4) holds for* $q \in \left(\frac{1}{2}, 1\right)$. *Then*

(i) it holds that

$$-\inf_{x\in A} I(x) \le \liminf_{\varepsilon\to 0} \varepsilon^{2(1-q)} \log \psi(1/\varepsilon, T)$$
$$\le \limsup_{\varepsilon\to 0} \varepsilon^{2(1-q)} \log \psi(1/\varepsilon, T) \le -\inf_{x\in\bar{A}} I(x).$$

where $A = \{x \in D_0[0,T] : x(T) > 1\}$, $\bar{A}$ *is the closure of* A, *and*

$$I(x) = \inf_{\{y\in H_1, x=F(-\int_{T-\cdot}^{T}(\widetilde{\sigma}_s, \dot{y}(s)ds))\}} \frac{1}{2}\int_0^t |\dot{y}(t)|^2 dt.$$

(ii) If in addition $\lim_{x\to\infty} G(x)/x = \infty$, *then the estimate is true for* $q = 1/2$.

(iii) If the matrix function $\widetilde{\sigma}_t$ *has a column that is non-zero for all* t, *then* A *is a "continuity set" of* I, *that is,* $\inf_{x\in A} I(x) = \inf_{x\in\bar{A}} I(x)$. *In this case,* $\lim_{\varepsilon\to 0} \varepsilon^{2(1-q)} \log \psi(1/\varepsilon, T) = -\inf_{x\in A} I(x)$. ■

We remark that a sufficient condition for (ii) of Theorem 4.3 is that all the jump sizes of S are uniformly bounded. Further, if (ii) and (iii) both hold, then the adjustment coefficient $\widetilde{\delta}_T = \inf_{x\in A} I(x)$.

Infinite Horizon Case. In light of Djehiche[5] and Martin-Löf[11], in this case we would like to estimate $\varepsilon \log \psi(1/\varepsilon, T/\varepsilon)$ as $\varepsilon \to 0$. Assume that $b(t,x,y) = b(x,y)$, and $h_t \equiv c$. Let $Y_\cdot(x)$ be the corresponding storage process. From Theorem 4.2 we deduce that $\psi(1/\varepsilon, T/\varepsilon) = P\{\varepsilon Y_{T/\varepsilon}(1/\varepsilon) > 1\}$. Now let $Y_t^\varepsilon = \varepsilon Y_{t/\varepsilon}(1/\varepsilon)$, and suppose that $b^\varepsilon(y) \stackrel{\triangle}{=} b(1/\varepsilon, 1/\varepsilon y)$ has a uniform limit $\widetilde{b}(y)$. Let $F : C_T^0 \times \mathbb{D}_T^0 \mapsto \mathbb{D}_T^0$ be the solution mappings of the reflected ODE:

$$f(t) = F[g_1, g_2](t) = -\int_0^t \widetilde{b}(f(s))ds - g_1(s) + g_2(s) + \widetilde{k}_t, \quad t \ge 0.$$

Then, we have the following result (see, Ma and Sun[10]):

Theorem 4.4 *Assume (A5). Then the large deviation principle holds for* Y^ε *in* $\mathbb{D}_T^0$, *with the good rate function good rate fuction*

$$I'(x) = \inf\{I(y_1, y_2) : y_1 \in C_T^0,\ y_2 \in D_T^0,\ x = F(y_1, y_2)\},$$

where

$$I(y_1, y_2) = \begin{cases} \frac{1}{2c^2} \int_0^T |\dot{y}_1(t)|^2 dt + \int_0^T h(\dot{y}_2(t))dt, & y_1 \in H_1, y_2 \in H_1 \\ \infty & otherwise \end{cases}$$

More precisely, it holds that $\lim_{\varepsilon \to 0} \varepsilon \log \psi(1/\varepsilon, T/\varepsilon) = -\inf_{x \in A} I'(x)$, *where* A *is the same as that in theorem (4.3).* ■

We note that if $c = 0$, then we recover the result of Djehiche[5].

1. Asmussen, S. and Nielsen, H.M. (1995), *Ruin probabilities via local adjustment coefficients*, *Journal of Applied Probability*, **20**, 913-916.
2. Asmussen, S. and Petersen, S. S. (1988) *Ruin probabilities expressed in terms of storage processes*, *Adv. Appl. Prob.*, **20**,**4**,, 913-916.
3. Brémaud, P. (1981), *Point Processes and Queues —Matingale Dynamics*, Springer.
4. Dupuis, P. and Ishii, H. (1991), *On Lipschitz continuity of the solution mapping to the Skorohod problem with applications*, *Stochastics and Stochastic reports*, **35**, 31-62.
5. Djehiche, B. (1993), *A large deviation estimate for ruin probabilities*, *Scand. Actuarial J.* **1**, 42-59.
6. Grandell, J. (1991), *Aspects of Risk Theory*, Springer.
7. Kalashnikov, V. and Norberg, R. (2000) *Power tailed ruin probabilities in the presence of risky investments*, Working paper #159, Laboratory of Actuarial Mathematics, University of Copenhagen.
8. Ma, J. (1993), *Discontinuous Reflection, and a Class of Singular Stochastic Control Problems for Diffusions. Stochastic and Stochastics Reports*, Vol.44, 225–252.
9. Ma, J. and Sun, X. (2000), *Ruin Probabilities for Insurance Models Involving Investments*, Preprint.
10. Ma, J. and Sun, X. (2001), *Asymptotic Estimates of Ruin Probabilities for A General Insurance Model*, Preprint.
11. Martin-Löf, A. (1986) *Entropy, a useful concept in risk theory*, *Scan. Actuarial J.* **69** , 223-235.
12. Nyrhinen, H. (1999), *Finite and infinite time ruin probabilities in a stochastic economic environment*, Preprint 237, Department of Mathematics, University of Helsinki.
13. Petersen, S. S. (1990), *Calculation of ruin probabilities when the premium depends on the current reserve*, *Scand. Actuarial J.* 147-159.
14. Prabhu, N.U. (1980) *Stochastic Storage Processes – Queues, Insurance Risk, and Dams*, Springer-Verlag.

15. Protter, P. (1990), *Integration and Stochastic Differential Equations – A New Approach*, Springer-Verlag.
16. Rolski, T, Schmidli, H, Schmidt, V, and Teugels, J. (1999), *Stochastic Processes for Insurance and Finance*, Wiley Series in Probablility and Statistics, John Wiley & Sons.
17. Segerdahi, C. O. (1942). Uber einige risikotheoretische Fragestellungen. *Skand. Akt. Tidskr.* **25**, 43-83.

Risk-Sensitive Optimal Investment Problems with Partial Information on Infinite Time Horizon

Hideo Nagai[a]
Department of Mathematical Science
Graduate School of Engineering Science
Osaka University, Toyonaka, 560-8531, Japan
E-mail: nagai@sigmath.es.osaka-u.ac.jp

Shige Peng
Department of Mathematics
Shangdong University, Jinan 250100, China
E-mail: peng@sdu.edu.cn

1 Introduction

Bielecki and Pliska [2] have recently treated a factor model where the mean returns of individual securities are explicitly affected by economic factors defined as ergodic Gauss-Markov processes. For such model they considered an optimal investment problem maximizing the risk-sensitized expected growth rate per unit time of the value of the capital the investor possess under the condition that security prices and factors have independent randomness, which has been improved in the works by Bielecki- Pliska [3], Fleming-Sheu [5],[6] and Kuroda-Nagai [7] later. In these works as well the investment strategies are assumed to be chosen by observing all the past informations of the factor processes as well as the secutrity prices, while in the previous work [9] we relaxed the measurability conditions for the investment strategies with no constraint as the ones to be selected without using informations of factor processes but by using only past informations of security prices in the case of a finite time horizon. Then the problem is formulated as a kind of risk-sensitive stochastic control with partial information. Indeed we can formulate our problem by regarding the factor processes as system processes and security prices observation processes in terms of stochastic control. Under such setting up we have constructed the optimal strategies for the optimal investment problem on a finite time horizon, which are explicitly represented by the solutions of the ordinary differential equations with the Riccati equations concerning filter and the value function. Here we shall discuss the optimal investment problem on infinite time horizon

[a]Research supported in part by Grant-in-Aid for Scientific Research No. 13440033 and 12874017 the Ministry of Education, Science and Culture

under such formulation with partial information. To consider such problem it is necessary to study asymptotic behavior of the solution of a inhomogeneous Riccati differential equation. We shall present the results on new feature of the solution and construction of the optimal strategy for the problem from the solutions of the limit equations of the Riccati equation and filter.

2 Setting up

We consider a market with $m+1 \geq 2$ securities and $n \geq 1$ factors. We assume that the set of securities includes one bond, whose price is defined by ordinary differential equation:

$$dS^0(t) = r(t)S^0(t)dt, \quad S^0(0) = s^0, \tag{2.1}$$

where $r(t)$ is a deterministic function of t. The other secutity prices and factors are assumed to satsfy the following stochastic differential equations:

$$\begin{aligned} &dS^i(t) = S^i(t)\{(a + AX_t)^i dt + \textstyle\sum_{k=1}^{n+m} \sigma_k^i dW_t^k\}, \\ &S^i(0) = s^i, \quad i = 1, ..., m \end{aligned} \tag{2.2}$$

and

$$dX_t = (b + BX_t)dt + \Lambda dW_t, \qquad X(0) = x \in R^n, \tag{2.3}$$

where $W_t = (W_t^k)_{k=1,..,(n+m)}$ is a $m+n$ dimensional standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$. Here A, B, Λ are respectively $m \times n$, $n \times n$, $n \times (m+n)$ constant matrices and $a \in R^m$, $b \in R^n$. The constant matrix $(\sigma_k^i)_{i=1,2..,m;\ k=1,2..,(n+m)}$ will be often denoted by Σ in what follows. We always assume that

$$\Sigma\Sigma^* > 0, \tag{2.4}$$

where Σ^* stands for the transposed matrix of Σ.

Let us denote investment strategy to i-th security $S^i(t)$ by $h^i(t)$, $i = 0, 1, ..., m$ and set

$$S(t) = (S^1(t), S^2(t), ..., S^m(t))^*,$$

$$h(t) = (h^1(t), h^2(t), ..., , h^m(t))^*$$

and

$$\mathcal{G}_t = \sigma(S(u); u \le t).$$

Here S^* stands for transposed matrix of S.

Definition 2.1 $(h^0(t), h(t)^*)_{0\le t\le T}$ *is said an invetment strategy if the following conditions are satisfied*

i) $h(t)$ is a R^m valued $\mathcal{G}_t$ progressively measurable stochastic process such that

$$\sum_{i=1}^{m} h^i(t) + h^0(t) = 1 \tag{2.5}$$

ii)

$$P(\exists c(\omega) s.t. |h(s)| \le c(\omega),\ \ 0 \le s \le T) = 1.$$

The set of all investment strategies will be denoted by $\mathcal{H}(T)$. When $(h^0(t), h(t)^*)_{0\le t\le T} \in \mathcal{H}(T)$ we will often write $h \in \mathcal{H}(T)$ for simplicity since h^0 is determined by (2.5).

For given $h \in \mathcal{H}(T)$ the process $V_t = V_t(h)$ representing the investor's capital at time t is determined by the stochastic differential equation:

$$\begin{aligned}
\frac{dV_t}{V_t} &= \textstyle\sum_{i=0}^{m} h^i(t)\dfrac{dS^i(t)}{S^i(t)} \\
&= h^0(t)r(t)dt + \textstyle\sum_{i=1}^{m} h^i(t)\{(a + AX_t)^i dt + \sum_{k=1}^{m+n} \sigma^i_k dW^k_t\} \\
V_0 &= v.
\end{aligned}$$

Then, taking (2.5) into account it turns out to be a solution of

$$\begin{aligned}
\frac{dV_t}{V_t} &= r(t)dt + h(t)^*(a + AX_t - r(t)\mathbf{1})dt + h(t)^*\Sigma dW_t, \\
V_0 &= v,
\end{aligned} \tag{2.6}$$

where $\mathbf{1} = (1, 1, ..., 1)^*$.

We first consider the following problem. For a given constant $\theta > -2$, $\theta \neq 0$ maximize the following risk-sensitized expected growth rate up to time horizon T:

$$J(v, x; h; T) = -\frac{2}{\theta} \log E[e^{-\frac{\theta}{2}\log V_T(h)}], \tag{2.7}$$

where h ranges over the set $\mathcal{A}(T)$ of all investment strategies defined later. Then we consider the problem maximizing the risk-sensitized expected growth rate per unit time

$$J(v,x;h) = \limsup_{T\to\infty}\left(\frac{-2}{\theta T}\right)\log E[e^{-\frac{\theta}{2}\log V_T(h)}], \tag{2.8}$$

where h ranges over the set of all investment straregies such that $h \in \mathcal{A}(T)$ for each T. Note that in our problem a strategy h is to be chosen as $\sigma(S(u); u \leq t)$ measurable process, different from the case of Bielecki-Pliska where it is $\sigma((S(u), X_u), u \leq t)$ measurable. Namely, in our case the strategy is to be selected without using past informations of the factor process X_t.

Since V_t satisfies (2.6) we have

$$\begin{aligned} V_t^{-\theta/2} &= v^{-\theta/2}\exp\{\tfrac{\theta}{2}\textstyle\int_0^t \eta(X_s, h_s, r(s))ds \\ &\quad - \tfrac{\theta}{2}\textstyle\int_0^t h_s^*\Sigma dW_s - \tfrac{1}{2}(\tfrac{\theta}{2})^2\int_0^t h_s^*\Sigma\Sigma^* h_s ds\}, \end{aligned}$$

where

$$\eta(x,h,r) = \frac{1}{2}\left(\frac{\theta}{2}+1\right)h^*\Sigma\Sigma^*h - r - h^*(a + Ax - r\mathbf{1}).$$

Therefore, if $\theta > 0$ (resp. $-2 < \theta < 0$) our problem maximizing $J(v,x;h;T)$ is reduced to the one minimizing (resp. maximizing) the following criterion:

$$\begin{aligned} I(x,h;T) &= v^{-\theta/2}E[\exp\{\tfrac{\theta}{2}\textstyle\int_0^t \eta(X_s, h_s, r(s))ds \\ &\quad - \tfrac{\theta}{2}\textstyle\int_0^t h_s^*\Sigma dW_s - \tfrac{1}{2}(\tfrac{\theta}{2})^2\int_0^t h_s^*\Sigma\Sigma^* h_s ds\}]. \end{aligned} \tag{2.9}$$

Now we reformulate the problem as one of partially observable stochastic control. Set

$$Y_t^i = \log S^i(t),$$

then $Y_t = (Y_t^1, ..., Y_t^m)^*$ satisfies the following stochastic diferential

$$dY_t^i = \{a^i - \frac{1}{2}(\Sigma\Sigma^*)^{ii} + (AX_t)^i\}dt + \sum_{k=1}^{m+n}\sigma_k^i dW_t^k, \tag{2.10}$$

$i = 1, ..., m,$. So, setting $d = (d^i) \equiv (a^i - \frac{1}{2}(\Sigma\Sigma^*)^{ii})$, we have

$$dY_t = (d + AX_t)dt + \Sigma dW_t, \tag{2.11}$$

which we regard as the SDE defining the observation process. On the other hand, X_t defined by (2.3) is regarded as a system process. System noise ΛdW_t and observation noise ΣdW_t are correlated in general. $\sigma(Y_u,;u \le t) = \sigma(S(u);u \le t)$ holds since log is a strictly increasing function, so our problem is to minimize (or maximize) the criterion (2.9) while looking at the observation process Y_t and choosing a $\sigma(Y_u,;u \le t)$ measurable strategy $h(t)$. Though there is no control in the SDE (2.3) defining system process X_t criterion $I(x,h;T)$ is defined as a functional of the strategy $h(t)$ measurable with respect to observation and the problem is the one of stochstic control with partial observation.

Now let us introduce a new probability measure $\widehat{P}$ on $(\Omega, \mathcal{F})$ defined by

$$\left.\frac{d\widehat{P}}{dP}\right|_{\mathcal{F}_T} = \rho_T,$$

where

(2.12)

$$\rho_t = \exp\{-\int_0^t (d+AX_s)^*(\Sigma\Sigma^*)^{-1}\Sigma dW_s - \frac{1}{2}\int_0^t (d+AX_s)^*(\Sigma\Sigma^*)^{-1}(d+AX_s)ds\}.$$

We see that $\widehat{P}$ is a probability measure since it can be seen by standard arguments (cf. [1]) that ρ_t is a martingale and $E[\rho_T] = 1$. Moreover, according to Girsanov theorem,

$$\widehat{W}_t = W_t + \int_0^t \Sigma^*(\Sigma\Sigma^*)^{-1}(d+AX_s)ds \tag{2.13}$$

turns out to be a standard Brownian motion process under the probability measure $\widehat{P}$ and we have

$$dY_t = \Sigma d\widehat{W}_t \tag{2.14}$$

$$dX_t = \{b + BX_t - \Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}(d+AX_t)\}dt + \Lambda d\widehat{W}_t. \tag{2.15}$$

We rewrite our criterion $I(x,h;T)$ by new probability measure $\widehat{P}$.

$$I(x,h;T) = v^{-\theta/2}\widehat{E}[\widehat{E}[\exp\{\frac{\theta}{2}\int_0^T \eta(X_s,h_s;r(s))ds\}\Psi_T|\mathcal{G}_T]], \tag{2.16}$$

where

$$\Psi_t = \exp\{\int_0^t Q(X_s,h_s)^* dY_s - \frac{1}{2}\int_0^t Q(X_s,h_s)^*(\Sigma\Sigma^*)Q(X_s,h_s)ds\}$$

and

$$Q(x,h) = (\Sigma\Sigma^*)^{-1}(Ax+d) - \frac{\theta}{2}h = (\Sigma\Sigma^*)^{-1}\{(Ax+d) - \frac{\theta}{2}(\Sigma\Sigma^*)h\}.$$

Set

$$q^h(t)(\varphi(t)) = \widehat{E}[\exp\{\frac{\theta}{2}\int_0^t \eta(X_s,h_s;r(s))ds\}\Psi_t\varphi(t,X_t)|\mathcal{G}_t], \tag{2.17}$$

then (2.16) reads

$$I(x,h;T) = v^{-\theta/2}\widehat{E}[q^h(T)(1)] \tag{2.18}$$

Hence, if $\theta > 0$ (resp. $-2 < \theta < 0$) our problem is reduced to minimize (resp. maximize) I of (2.18) when taking h over $\mathcal{H}(T)$.

3 Finite time horizon case

Let us set

$$L\varphi = \frac{1}{2}(\Lambda\Lambda^*)^{ij}D_{ij}\varphi + (b+Bx)^i D_i\varphi, \tag{3.1}$$

Then, the following proposition can be obtained by using Ito calculus in a standard way.

Proposition 3.1 *$q(t)(\varphi(t)) \equiv q^h(t)(\varphi(t))$ satisfies the following stochastic partial differential equation (SPDE):*

(3.2)

$$q(t)(\varphi(t)) = q(0)(\varphi(0)) + \int_0^t q(s)(\tfrac{\partial\varphi}{\partial t}(s,\cdot) + L\varphi(s,\cdot) - \tfrac{\theta}{2}h_s^*\Sigma\Lambda^* D\varphi(s,\cdot)$$

$$+ \tfrac{\theta}{2}\eta_s(\cdot)\varphi(s,\cdot))ds + \int_0^t q(s)(\varphi(s,\cdot)Q(\cdot,h_s))dY_s$$

$$+ \int_0^t q(s)((D\varphi)^*\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1})dY_s,$$

where $\eta_s(\cdot) = \eta(\cdot,h_s;r(s))$.

Now let us give the explicit representation to the solution of SPDE (3.2). For that let us introduce matrix Riccati equation

$$
\begin{aligned}
&\dot{\Pi} + (\Pi A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\Pi + \Sigma\Lambda^*) - \Lambda\Lambda^* - B\Pi - \Pi B^* = 0, \\
&\Pi(0) = 0.
\end{aligned}
\tag{3.3}
$$

and stochastic differential equation:

$$
\begin{aligned}
&d\gamma_t = \{B\gamma_t + b - (\Pi A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\gamma_t + d)\}dt \\
&\quad + (\Pi A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}dY_t \\
&\gamma_0 = x.
\end{aligned}
\tag{3.4}
$$

Theorem 3.2 *The solution of SPDE (3.2) with $q(0)(\varphi(0)) = \varphi(0,x)$ has the following representation:*

$$
q(t)(\varphi(t)) = \alpha_t \int \varphi(t, \gamma_t + \Pi_t^{\frac{1}{2}} z)\frac{1}{(2\pi)^{n/2}} e^{-\frac{|z|^2}{2}} dz,
$$

where

$$
\begin{aligned}
\alpha_t = \exp\{&\textstyle\int_0^t Q(\gamma_s, h_s)^* dY_s - \frac{1}{2}\int_0^t Q(\gamma_s, h_s)^*(\Sigma\Sigma^*)Q(\gamma_s, h_s)ds \\
&+ \textstyle\frac{\theta}{2}\int_0^t \eta(\gamma_s, h_s; r(s))ds\}.
\end{aligned}
$$

Remark It is known that (3.3) has a unique solution.

criteBecause of Theorem 3.2 (2.18) reads

$$
I(x, h; T) = v^{-\theta/2}\widehat{E}[\alpha_T],
\tag{3.5}
$$

so we shall consider the problem minimizing (resp. maximizing) the criterion represented by (3.5) if $\theta > 0$ (resp. $-2 < \theta < 0$). Let us introduce the following

$n \times n$ matrix Riccati differential equation.

$$
\begin{aligned}
&\dot{U} + U\{B - \tfrac{\theta}{\theta+2}(\Pi A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}A\} \\
&+\{B^* - \tfrac{\theta}{\theta+2}A^*(\Sigma\Sigma^*)^{-1}(A\Pi + \Sigma\Lambda^*)\}U \\
&-\tfrac{2\theta}{\theta+2}U(\Pi A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\Pi + \Sigma\Lambda^*)U \\
&+\tfrac{1}{\theta+2}A^*(\Sigma\Sigma^*)^{-1}A = 0, \\
&U(T) = 0.
\end{aligned}
\tag{3.6}
$$

When we have a solution U of (3.6) we get a solution g of the following linear differential equation on R^n.

$$
\begin{aligned}
&\dot{g} + B^*g - \tfrac{\theta}{\theta+2}A^*(\Sigma\Sigma^*)^{-1}(A\Pi + \Sigma\Lambda^*)g \\
&-\tfrac{2\theta}{\theta+2}U(\Pi A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\Pi + \Sigma\Lambda^*)g \\
&+\tfrac{1}{\theta+2}\{A - \theta(A\Pi + \Sigma\Lambda^*)U\}^*(\Sigma\Sigma^*)^{-1}(a - r(t)\mathbf{1}) + Ub = 0, \\
&g(T) = 0.
\end{aligned}
\tag{3.7}
$$

Furthermore, for given solutions U of (3.6) and g of (3.7) we have a solution k of the following differential equation.

$$
\begin{aligned}
&\dot{k} + r(t) + \mathrm{tr}[U(\Pi A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\Pi + \Sigma\Lambda^*)] \\
&-\theta g^*(\Pi A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\Pi + \Sigma\Lambda^*)g \\
&+2g^*b + \tfrac{1}{\theta+2}c_t^*(\Sigma\Sigma^*)^{-1}c_t = 0, \\
&k(T) = 0,
\end{aligned}
\tag{3.8}
$$

where

$$c_t = a - r(t)\mathbf{1} - \theta(A\Pi + \Sigma\Lambda^*)g.$$

Let us denote by $\mathcal{A}(T)$ the set of all investment strategy satisfying

$$\widehat{E}[\exp\{\int_0^T \Xi_s^*(h)dY_s - \frac{1}{2}\int_0^T \Xi_s(h)\Sigma\Sigma^*\Xi_s(h)ds\}] = 1$$

where

$$\Xi_t^* = [(\gamma_t^* A^* + d^*) - \theta(\gamma_t^* U + g^*)(\Pi A^* + \Lambda\Sigma^*) - \frac{\theta}{2}h_t^*(\Sigma\Sigma^*)](\Sigma\Sigma^*)^{-1}.$$

Theorem 3.3 *If (3.6) has a solution U, then there exists an optimal strategy $\widehat{h} \in \mathcal{A}(T)$ maximizing the criterion (2.7) and it is explicitly represented as*

$$(3.9)\quad \widehat{h}_t = \frac{2}{\theta+2}(\Sigma\Sigma^*)^{-1}[a - r(t)\mathbf{1} - \theta(A\Pi + \Sigma\Lambda^*)g + \{A - \theta(A\Pi + \Sigma\Lambda^*)U\}\gamma_t]$$

where g is a solution of (3.7) and Π (resp. γ_t) is the one of (3.3) (resp. 3.4). Moreover

$$(3.10)\quad \begin{aligned} J(v, x; \widehat{h}; T) &= \sup_{h\in\mathcal{A}(T)} J(v, x; h; T) \\ &= \log v + x^* U(0)x + 2g^*(0)x + k(0) \end{aligned}$$

where k is a solution of (3.8).

Remark It is known that (3.6) has a unique solution if $\theta > 0$.

4 Stability of filter

Now we study symptotic behavior of the solution $\Pi(t)$ of (3.3) as $t \to \infty$. We always assume that $r(t)$ is a constant r in what follows.
Lemma 4.1 *Asuume that*

$$(4.1)\qquad G := B - \Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}A \quad \textit{is stable,}$$

then $\Pi(t) \to \bar{\Pi} \geq 0$, $t \to \infty$, where $\bar{\Pi}$ is a unique nonnegative definite solution of the algebraic Riccati equation

$$(4.2)\qquad G\bar{\Pi} + \bar{\Pi}G^* - \bar{\Pi}A^*(\Sigma\Sigma^*)^{-1}A\bar{\Pi} + \Lambda(I_{m+n} - \Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma)\Lambda^* = 0.$$

Moreover, $G - \bar{\Pi}A^*(\Sigma\Sigma^*)^{-1}A$ is stable.

Remark. Moreover we can see that $\Pi(t)$ converges exponentially fast to $\bar{\Pi}$.

5 Asymptotics of Inhomogeneous Riccati Equations

To study asymptotics of the solution of (3.6) we first consider the equation

$$
(5.1) \quad \begin{aligned}
&\dot{\widetilde{U}} + \widetilde{U}(B - \tfrac{\theta}{\theta+2}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}A) \\
&+(B - \tfrac{\theta}{\theta+2}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}A)^*\widetilde{U} \\
&-\tfrac{2\theta}{\theta+2}\widetilde{U}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\bar{\Pi} + \Sigma\Lambda^*)\widetilde{U} \\
&+\tfrac{1}{\theta+2}A^*(\Sigma\Sigma^*)^{-1}A = 0 \\
&\widetilde{U}(T) = 0
\end{aligned}
$$

and we obtain the following lemma.

Lemma 5.1 *Under assumption (4.1) $\widetilde{U}(t;T)$ converges to $\bar{U} \geq 0$ as $T \to \infty$, where $\bar{U}$ is a unique nonnegative definite solution of algebraic Riccati equation*

$$
(5.2) \quad \begin{aligned}
&\bar{U}(B - \tfrac{\theta}{\theta+2}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}A) \\
&+(B - \tfrac{\theta}{\theta+2}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}A)^*\bar{U} \\
&-\tfrac{2\theta}{\theta+2}\bar{U}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\bar{\Pi} + \Sigma\Lambda^*)\bar{U} \\
&+\tfrac{1}{\theta+2}A^*(\Sigma\Sigma^*)^{-1}A = 0
\end{aligned}
$$

and the following matrix is stable:

(5.3)

$$
B - \frac{\theta}{\theta+2}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}A - \frac{2\theta}{\theta+2}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\bar{\Pi} + \Sigma\Lambda^*)\bar{U}
$$

To study specific equation (3.6), we shall see general feature on asymptotics of the solutions of more general inhomogeneous Riccati differential equations as follows. For given continuous matrix valued functions $C(t)$, $D(t)$ and $R(t) \geq 0$ and a constant matrix $N > 0$ we consider the inhomogeneous Riccati equation

$$
(5.4) \quad \begin{aligned}
&0 = \dot{K}_T + C(t)^*K_T + K_TC(t) - K_TD(t)N^{-1}D(t)^*K_T + R(t)^*R(t) \\
&K_T(T) = 0
\end{aligned}
$$

On asymptotics of the solution of this equation we have the following lemma

Lemma 5.2 *Assume that $C(t)$, $D(t)$ and $R(t)$ converge exponentially fast to $\bar{C}, \bar{D}, \bar{R}$ respectively as $t \to \infty$ and that $(\bar{C}, \bar{D})$ is stabilizable and $(\bar{R}, \bar{C})$ is* *detectable. Then there exists $\kappa > 0$, $\beta > 0$ and $T_* > 0$ such that for each $T > T_0 > T_*$ the solution of (5.4) on $[T_0, T]$ satisfies*

$$(5.5) \quad \widetilde{K}_T(t) + \kappa e^{-\beta T_0} K_T^-(t) \le K_T(t) \le \widetilde{K}_T(t) + \kappa e^{-\beta T_0} K_T^+(t), \quad t \in [T_0, T]$$

where $\widetilde{K}_T(t)$, $t \in [T_0, T]$ is the solution of

$$(5.6) \quad \begin{aligned} &0 = \dot{\widetilde{K}}_T + \bar{C}^* \widetilde{K}_T + \widetilde{K}_T \bar{C} - \widetilde{K}_T \bar{D} N^{-1} \bar{D}^* \widetilde{K}_T + \bar{R}^* \bar{R} \\ &\widetilde{K}_T(T) = 0 \end{aligned}$$

and $K_T^-(t)$ and $K_T^+(t)$ is the ones of
(5.7)

$$\begin{aligned} &0 = \dot{K}_T^- + \bar{C}^* K_T^- + K_T^- \bar{C} - K_T^- \bar{D} N^{-1} \bar{D}^* \widetilde{K}_T - \widetilde{K}_T \bar{D} N^{-1} \bar{D}^* K_T^- - I_n \\ &K_T^-(T) = 0 \end{aligned}$$

and
(5.8)

$$\begin{aligned} &0 = \dot{K}_T^+ + \bar{C}^* K_T^+ + K_T^+ \bar{C} - K_T^+ \bar{D} N^{-1} \bar{D}^* \widetilde{K}_T - \widetilde{K}_T \bar{D} N^{-1} \bar{D}^* K_T^+ + I_n \\ &K_T^+(T) = 0 \end{aligned}$$

respectively.
By using the above lemma we shall obtain the following theorem concerning aysmptotics of the solutions of (3.6), (3.7) and (3.8).
Theorem 5.3 *Assume (4.1). Then for the solutions $U(t;T)$, $g(t;T)$ and $k(t;T)$ of the equations (3.6), (3.7) and (3.8) respectively it follows that*

$$(5.14) \quad \lim_{T-t\to\infty,\ t\to\infty} U(t;T) = \bar{U}$$

$$(5.15) \quad \lim_{T-t\to\infty,\ t\to\infty} g(t;T) = \bar{g}$$

$$(5.16) \quad -\lim_{T-t\to\infty,\ t\to\infty} \dot{k}(t;T) = \rho(\theta),$$

where $\bar{U} \ge 0$ is the solution of (5.2), $\bar{g}$ the one of

$$(5.17) \quad \begin{aligned} &\{B - \tfrac{2\theta}{\theta+2}(\bar{\Pi} A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\bar{\Pi} + \Sigma\Lambda^*)\bar{U} \\ &-\tfrac{\theta}{\theta+2}(\bar{\Pi} A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}A\}^* \bar{g} \\ &+\bar{U}b + \tfrac{1}{\theta+2}\{A - \theta(A\bar{\Pi} + \Sigma\Lambda^*)\bar{U}\}^*(\Sigma\Sigma^*)^{-1}(a - r1) = 0 \end{aligned}$$

and $\rho(\theta)$ *is defined by*

$$\begin{array}{ll}(5.18) & \begin{array}{l}\rho(\theta) = tr[\bar{U}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\bar{\Pi} + \Sigma\Lambda^*)] \\ -\theta\bar{g}^*(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\bar{\Pi} + \Sigma\Lambda^*)\bar{g} \\ + \frac{1}{\theta+2}c^*(\Sigma\Sigma^*)^{-1}c + r + 2\bar{g}^*b\end{array}\end{array}$$

where $c = a - r\mathbf{1} - \theta(A\bar{\Pi} + \Sigma\Lambda^*)\bar{g}$

6 Infinite time horizon case

Theorem 6.1 *i) Under the assumptions of Theorem 5.3*

$$\sup_h J(v, x; h) \le \rho(\theta), \tag{6.1}$$

where $\rho(\theta)$ *is a constant defined by (5.18).*
ii) Besides above conditions we assume that

$$\bar{U}(\bar{\Pi}A^* + \Lambda\Sigma^*)(\Sigma\Sigma^*)^{-1}(A\bar{\Pi} + \Sigma\Lambda^*)\bar{U} < \frac{1}{\theta^2}A^*(\Sigma\Sigma^*)^{-1}A, \tag{6.2}$$

then

$$\bar{h}_t = \frac{2}{\theta+2}(\Sigma\Sigma^*)^{-1}\{a - r\mathbf{1} - \theta(A\bar{\Pi} + \Sigma\Lambda^*)\bar{g} + [A - \theta(A\bar{\Pi} + \Sigma\Lambda^*)\bar{U}]\gamma_t\}$$

is optimal:

$$J(v, x; \bar{h}) = \sup_h J(v, x; h) = \rho(\theta).$$

7 Appendix

Definition 7.1 *i) The pair* (L, M) *of* $n \times n$ *matrix* L *and* $n \times l$ *matrix* M *is said stabilizable if there exists* $l \times n$ *matrix* K *such that* $L - MK$ *is stable.*

ii) The pair (L, F) *of* $l \times n$ *matrix* L *and* $n \times n$ *matrix* F *is called detectable if* (F^*, L^*) *is stabilizable*

Let us consider the Riccati differential equation:

$$\begin{array}{ll}(7.1) & \begin{array}{l}\dot{P} + K_1^*P + PK_1 - P\Lambda N^{-1}\Lambda^*P + C^*C = 0, \\ P(T) = 0.\end{array}\end{array}$$

Then, the following theorem would be well known in engineering.

Theorem 7.1 *(Wonham[11], Kucera[8]) Assume that* $N > 0$ *and* (K_1, Λ) *is stabilizable, then for the solution of (7.1)* $\exists \lim_{T\to\infty} P(t;T) \equiv \lim_{T\to\infty} P(t) \equiv \widetilde{P}$ *and* $\widetilde{P}$ *satisfies the algebraic Riccati equation:*

$$K_1^* \widetilde{P} + \widetilde{P} K_1 - \widetilde{P} \Lambda N^{-1} \Lambda^* \widetilde{P} + C^* C = 0 \tag{7.2}$$

Moreover, if (C, K_1) *is detectable* $K_1^* - \widetilde{P}\Lambda N^{-1}\Lambda^*$ *is stable and the nonnegative definite solution* $\widetilde{P}$ *of (7.2) is unique.*

Definition 7.2 *i) The pair* (K, L) *of* $n \times n$ *matrix* K *and* $n \times l$ *matrix* L *is said controllable if* $n \times nl$ *matrix* $(L, KL, K^2L, ..., K^{n-1}L)$ *has rank* n.

ii) The pair (L, K) *of* $l \times n$ *matrix and* $n \times n$ *matrix is said observable if* (K^*, L^*) *is controllable.*

It is known that if the pair (K, L) of matrices is controllable (resp. observable) then it is stabilizable (resp. detectable) (cf. Wonham[11]).

References

1. A. Bensoussan and J.H. Van Schuppen, "Optimal control of partially observable stochastic systems with an exponential-of integral performance index", *SIAM J. Cont. Optim.*, vol. 23 (1985) 599-613

2. T.R. Bielecli and S.R. Pliska, "Risk-Sensitive Dynamic Asset Management",*Appl. Math. Optim.* vol. 39 (1999) 337-360

3. T.R. Bielecki and S.R. Pliska, "Risk-Sensitive Intertemporal CAPM, With Application to Fixed Income Management", preprint

4. W.H. Fleming and S.J. Sheu, "Optimal Long Term Growth Rate of Expected Utility of Wealth" preprint, (1998)

5. W.H. Fleming and S.J. Sheu, "Risk-sensitive control and an optimal investment model" Mathematical Finance, 10 (2000)197-213

6. W.H. Fleming and S.J. Sheu, "Risk-sensitive control and an optimal investment model (II)", preprint

7 K. Kuroda and H. Nagai "Risk-sensitive portfolio optimization on infinite time horizon", to appear in Stochastics and Stochastics Reports

8. V. Kucera "A contribution to matrix quadratic equations", IEEE Trans. Automat. Control, Vol. AC-17 (1972) 344-347

9. H. Nagai " Risk-Sensitive Dynamic Asset Management with Partial Information", Stochastics in finite and infinite dimension, a volume in honor of G. Kallianpur, Eds. Rajput et al., Birkhäuser, 321-340 (2000)

10. H. Nagai and S. Peng "Risk-sensitive dynamic portfolio optimization with partial information on infinite time horizon", to appear in Annals of Applied Prob.

11. W.M. Wonham, "On a Matrix Riccati Equation of Stochastic Control", *SIAM J. Control Optim.*, vol. 6 (1968) 681-697

Filtration Consistent Nonlinear Expectations

François Coquet, Ying Hu, Jean Mémin
IRMAR, U.M.R. C.N.R.S. 6625, Campus de Beaulieu,
Université Rennes 1, 35042 Rennes, Cédex, France.

Shige Peng[a]
Department of Mathematics
Shangdong University, Jinan 250100, China

Abstract

From a general definition of nonlinear expectations, viewed as operators preserving monotonicity and constants, we derive, under rather general assumptions, the notions of conditional nonlinear expectation and nonlinear martingales. We prove that any such nonlinear martingale can be represented as the solution of a backward stochastic equation, and in particular admits continuous paths. In other words, it is a g-martingale.

Mathematic Subject Classification: primary 60H10

Keywords: BSDE, nonlinear expectation, g-martingale, nonlinear martingale, Doob-Meyer decomposition

1 Introduction

A fundamental problem in financial ecomimics is how to evaluate risky assets. The notion of nonlinear expectations and non–additive probabilities is considered as a key concept. A (possibly nonlinear) expectation on a probability space $(\Omega, \mathcal{F}, P)$ is a map

$$\mathcal{E} : L^2(\Omega, \mathcal{F}, P) \longmapsto R$$

which satisfies the following properties:

$$\begin{aligned} &\text{if} \quad X_1 \geq X_2 \quad a.s., \quad \mathcal{E}[X_1] \geq \mathcal{E}[X_2], \qquad \text{and} \\ &\text{if} \quad X_1 \geq X_2 \quad a.s., \quad \mathcal{E}[X_1] = \mathcal{E}[X_2] \quad \Longleftrightarrow \quad X_1 = X_2 \quad a.s. \end{aligned}$$

$$\mathcal{E}[c] = c, \quad \text{for each constant } c.$$

[a]This author's work is supported in part by The National Science Foundation of China, Probability Program and CNRS, France. He would like to thank IRMAR, Université Rennes 1 for its warm hospitality during his visit.

In particular, if $\mathcal{E}[\cdot]$ is linear, then it becomes a classic expectation under the probability measure defined by $P_{\mathcal{E}}(A) = \mathcal{E}[1_A]$, $A \in \mathcal{F}$. In fact, there is a one-to-one correspondence between the set of linear expectations and that of σ-additive probability measures on $(\Omega, \mathcal{F})$. But in the nonlinear case this one-to-one correspondence no longer holds true: a nonlinear expectation can always induce a, generally non-additive, 'probability measure' by $P(A) = \mathcal{E}[1_A]$. But, in general, a (possibly non-additive) probability measure can not characterize a nonlinear expectation. For example, if E is the classical linear expectation defined by the probability measure P, and f denotes a strictly increasing continuous function on $\mathbf{R}$ such that $f(x) = x$ whenever $0 \leq x \leq 1$, $\mathcal{E}^f[X] = f^{-1}\Big(E[f(X)]\Big)$ defines a non linear expectation (unless f is a linear mapping). But clearly, any such expectation induces the same probability measure, that is P itself: $P(A) = E[1_A] = \mathcal{E}^f[1_A]$.

A nonlinear expectation is said to be filtration-consistent under a given filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if, for each $t \geq 0$, the corresponding conditional expectation $\mathcal{E}[X|\mathcal{F}_t]$ of X under $\mathcal{F}_t$, characterized by

$$\mathcal{E}\Big[\mathcal{E}[X|\mathcal{F}_t]1_A\Big] = \mathcal{E}[X1_A], \quad \forall A \in \mathcal{F}_t,$$

exists.

A type of filtration-consistent nonlinear expectations, under a Brownian filtration, was introduced in [12], under the name "g-expectation" (see Section 2. for details). These g-expectations can be considered as a nonlinear extension of the well-known Girsanov transformations. It is a nonlinear mapping, but it preserves almost all other properties of the classical linear expectations. For more detailed views on this topic, we refer to [12], [5], [13], or [1] where some special cases are studied in depth, including the y-independent case, which will turn out to be the natural setting behind the present work.

A very interesting problem is: is this notion of g-expectation general enough to represent all "enough regular" filtration-consistent nonlinear expectations? Answering this question is the main objective of the present paper. We will give this theorem in Section 4 (Theorem 4.1) and prove it in Section 7: if for a large enough $\mu > 0$, a nonlinear expectation $\mathcal{E}[\cdot]$ is dominated by the '$\mu|z|$–expectation' $\mathcal{E}^{\mu}[\cdot]$ (that is, the g-expectation defined by $g(z) = \mu|z|$), and if $\mathcal{E}[X+\eta|\mathcal{F}_t] = \mathcal{E}[X|\mathcal{F}_t]+\eta$ for all $\mathcal{F}_t$-measurable η (i.e., the nonlinearity of $\mathcal{E}[\cdot]$ is only from to the risk), then, there exists a unique g such that $\mathcal{E}[\cdot]$ is the nonlinear expectation defined by g, still according to the definition of [12]. Our main tool will be the decomposition theorem for g-supermartingales proved in [13], developed here along a new version suitable for continuous $\mathcal{E}$-supermartingales, which we prove in Section 6. Basic definitions about g-expectations are given in Section 2. Sections 3. and 4. give the general framework of non-linear expectations, while Section 5. is devoted to martingales defined under non-linear expectations. The omitted proofs in this paper can be find in [7]

2 Basic notations and results on g-expectations

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $(B_t)_{t\geq 0}$, be a d-dimensional standard Brownian motion on this space such that $B_0 = 0$. Let $(\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by this Brownian motion:

$$\mathcal{F}_t = \sigma\{B_s, s \in [0,t]\} \vee \mathcal{N},$$

where $\mathcal{N}$ is the set of all P-null subsets. Let $T > 0$ be a given number. Without loss of generality, in this paper, we always work in the space $(\Omega, \mathcal{F}_T, P)$, and only consider processes indexed by $t \in [0,T]$.

$L^2_{\mathcal{F}}(0,T;E)$ will denote the space of all E-valued, $(\mathcal{F}_t)_{t\leq T}$-adapted processes ϕ such that

$$E\int_0^T |\phi(s)|^2 ds < \infty.$$

We will shorten this notation by putting $L^2_{\mathcal{F}}(0,T)= L^2_{\mathcal{F}}(0,T;\mathbf{R})$.

We first recall the notion of g-expectations, defined in [12], from which most basic material of this section is taken . We are given a function g:

$$g(\omega,t,y,z) : \Omega \times [0,T] \times R \times R^d \longmapsto R$$

satisfying

$$\begin{cases} \text{(i)} \;\; g(\cdot,y,z) \in L^2_{\mathcal{F}}(0,T), \quad \text{for each } y \in R,\; z \in R^d; \\ \text{(ii)} \;\; g(\cdot,y,0) \equiv 0, \quad \text{for each } y \in R; \\ \text{(iii)} \;\; \exists C_0, \mu > 0 \quad \text{such that} \quad \forall y_1,\; y_2 \in R, \quad z_1, z_2 \in R^d, \\ \qquad |g(t,y_1,z_1) - g(t,y_2,z_2)| \leq C_0|y_1 - y_2| + \mu|z_1 - z_2|. \end{cases} \tag{1}$$

For each given $X \in L^2(\Omega, \mathcal{F}_T, P)$, let $(y^X(\cdot), z^X(\cdot)) \in L^2_{\mathcal{F}}(0,T;R^1 \times R^d)$ be the unique solution of the following backward stochastic differential equation (BSDE):

$$\begin{aligned} -dy^X(t) &= g(t, y^X(t), z^X(t))dt - z^X(t)dB_t, \\ y^X(T) &= X. \end{aligned}$$

(We refer to [10] for definitions and basic results about BSDEs; it will be enough here to remember that, provided that g satisfies (1), there is a unique pair $(y^X(\cdot), z^X(\cdot))$ of adapted processes solving the equation above).

Definition 2.1 (**g-expectation**) *The g-expectation $\mathcal{E}_g[\cdot] : L^2(\Omega, \mathcal{F}, P) \longmapsto R$ is defined by*

$$\mathcal{E}_g[X] = y^X(0).$$

Definition 2.2 (**conditional g-expectation**) *The conditional g-expectation of X with respect to $\mathcal{F}_t$ is defined by*

$$\mathcal{E}_g[X|\mathcal{F}_t] = y^X(t).$$

If $\tau \leq T$ is a stopping time, we define similarly

$$\mathcal{E}_g[X|\mathcal{F}_\tau] = y^X(\tau).$$

g-expectations and conditional g-expectations are in general not linear. However, they meet the following basic properties of usual expectations (see [12] for proofs):

Proposition 2.1 *(i) (preserving of constants): For each constant c, $\mathcal{E}_g[c] = c$;*
(ii)(monotonicity): If $X_1 \geq X_2$ a.s., then $\mathcal{E}_g[X_1] \geq \mathcal{E}_g[X_2]$;
(iii)(strict monotonicity): If $X_1 \geq X_2$ a.s., and $P(X_1 > X_2) > 0$, then

$$\mathcal{E}_g[X_1] > \mathcal{E}_g[X_2].$$

Proposition 2.2 *(i) If X is $\mathcal{F}_t$-measurable, then $\mathcal{E}_g[X|\mathcal{F}_t] = X$;*

(ii) For all stopping times τ and $\sigma \leq T$, $\mathcal{E}_g\left[\mathcal{E}_g[X|\mathcal{F}_\tau]|\mathcal{F}_\sigma\right] = \mathcal{E}_g[X|\mathcal{F}_{\tau\wedge\sigma}]$;

(iii) If $X_1 \geq X_2$ a.s., then $\mathcal{E}_g[X_1|\mathcal{F}_t] \geq \mathcal{E}_g[X_2|\mathcal{F}_t]$; if, moreover, $P(X_1 > X_2) > 0$, then $P\Big(\mathcal{E}_g[X_1|\mathcal{F}_t] > \mathcal{E}_g[X_2|\mathcal{F}_t]\Big) > 0$;

(iv) For each $B \in \mathcal{F}_t$, $\mathcal{E}_g[1_B X|\mathcal{F}_t] = 1_B \mathcal{E}_g[X|\mathcal{F}_t]$.

Proposition 2.3 *$\mathcal{E}_g[X|\mathcal{F}_t]$ is the unique random variable η in $L^2(\Omega, \mathcal{F}_t, P)$ such that*

$$\mathcal{E}_g[1_A X] = \mathcal{E}_g[1_A \eta] \quad \textit{for all} \quad A \in \mathcal{F}_t. \tag{2}$$

Definition 2.3 (**g-martingales**) *A process $(Y_t)_{0\leq t\leq T}$ such that $E[Y_t^2] < \infty$ for all t is a g-martingale (resp. g-supermartingale, g-submartingale) iff*

$$\mathcal{E}_g[Y_t|\mathcal{F}_s] = Y_s, \quad (\textit{resp.} \leq Y_s, \geq Y_s), \quad \forall s \leq t \leq T.$$

In the following proposition, $\|\cdot\|_p$ denotes the norm of $L^p(\Omega, \mathcal{F}_T, P)$.

Proposition 2.4 *Let $g(\omega, t, y, z) : \Omega \times [0,T] \times R \times R^d \longmapsto R$ be a given function satisfying (1). Then for every ε such that $0 < \varepsilon \leq 1$, there exists a constant C_ε such that, for every X,*

$$|\mathcal{E}_g[X]| \leq C_\varepsilon ||X||_{1+\varepsilon} \tag{3}$$

□

We shall often have to assume that

$$g \text{ does not depend on } y. \tag{4}$$

The importance of this special setting follows from the following lemma, which is proven in [1], subsection 4.2:

Lemma 2.1 *Let $g(\omega, t, y, z) : \Omega \times [0,T] \times R \times R^d \longmapsto R$ be a given function satisfying (1). Then*

$$\mathcal{E}_g[X + \eta|\mathcal{F}_t] = \mathcal{E}_g[X|\mathcal{F}_t] + \eta, \quad \forall \eta \in L^2(\Omega, \mathcal{F}_t, P) \tag{5}$$

if and only if g satisfies (4)

We will always write in the sequel $\mathcal{E}^\mu[X] \equiv \mathcal{E}_g[X]$ for $g = \mu|z|$ and $\mathcal{E}^{-\mu}[X] = \mathcal{E}_g[X]$ for $g \equiv -\mu|z|$. Note that

$$\forall C > 0, \quad \mathcal{E}^\mu[CX|\mathcal{F}_t] = C\mathcal{E}^\mu[X|\mathcal{F}_t] \tag{6}$$

and

$$\forall C < 0, \quad \mathcal{E}^\mu[CX|\mathcal{F}_t] = -C\mathcal{E}^\mu[-X|\mathcal{F}_t].$$

Next lemma will be useful later.

Lemma 2.2 *We have for all $\mu > 0$ and $X \in L^2(\Omega, \mathcal{F}_T, P)$,*

$$E\left[\mathcal{E}^\mu[X|\mathcal{F}_t]^2\right] \leq e^{\mu^2(T-t)} E[X^2].$$

□

Next Proposition of Doob-Meyer's type is obtained in [13].

Proposition 2.5 *Assume that g satisfies (1) and (4), and that (Y_t) is a right-continuous g-supermartingale on $[0,T]$ such that $E[\sup_{t \leq T} Y_t^2] < \infty$. Then there exists a unique pair (M, A) of processes such that*

M is a g-martingale;

A is an increasing càdlàg process;

$$Y_t = M_t - A_t, \quad \forall t \in [0,T].$$

More specifically, Y is the unique solution of the BSDE

$$Y_t = Y_T + \int_t^T g(s, Z_s)ds + (A_T - A_t) - \int_t^T Z_s dB_s, \quad t \in [0, T].$$

We end this Section by giving an appropriate version of a downcrossing inequality given in [6] as Theorem 6.

Proposition 2.6 *Let g satisfy (1) and (Y_t) be a g-supermartingale on $[0, T]$. Let $0 = t_0 < t_1 < \cdots < t_n = T$, and $a < b$ be two constants. Then the number $D_a^b[Y, n]$ of downcrossings of $[a, b]$ by $\{X_{t_j}\}_{0 \le j \le n}$ satisfies*

$$\mathcal{E}^{-\mu}\left[D_a^b[Y, n]\right] \le \frac{1}{b-a}\mathcal{E}^{\mu}[Y_0 \wedge b - Y_T \wedge b].$$

Remark 2.1 *Contrarily to Theorem 6 in [6], we need not assume that Y is positive: indeed, as $g(\cdot, y, 0) = 0$, one checks easily that the proof given in [6] can be carried over for every g-supermartingale.*

Remark 2.2 *This proposition allows us to prove, by classical means, that a g-supermartingale (Y_t) admits a càdlàg modification if and only if the mapping $t \longrightarrow \mathcal{E}_g(Y_t)$ is right-continuous. More details on this topic will be given in Lemma 5.2.*

3 Filtration-Consistent Nonlinear Expectations

We give the basic notions and properties of $\mathcal{F}_t$-consistent nonlinear expectations. The omitted proof, good excercises to students, can be found in [7].

Definition 3.1 *A nonlinear expectation is a functional:*

$$\mathcal{E}[\cdot] : L^2(\Omega, \mathcal{F}_T, P) \longmapsto R$$

which satisfies the following properties:

(i) Strict monotonicity:

$$\text{if} \quad X_1 \ge X_2 \quad a.s., \quad \mathcal{E}[X_1] \ge \mathcal{E}[X_2], \qquad \text{and}$$
$$\text{if} \quad X_1 \ge X_2 \quad a.s., \quad \mathcal{E}[X_1] = \mathcal{E}[X_2] \quad \Longleftrightarrow \quad X_1 = X_2 \quad a.s.$$

(ii) preserving of constants:

$$\mathcal{E}[c] = c, \quad \text{for each constant } c.$$

Lemma 3.1 *Let $t \leq T$ and $\eta_1, \eta_2 \in L^2(\Omega, \mathcal{F}_t, P)$. If*

$$\mathcal{E}[\eta_1 1_A] = \mathcal{E}[\eta_2 1_A], \quad \forall A \in \mathcal{F}_t,$$

then $\eta_2 = \eta_1$, a.s. □

Definition 3.2 *For the given filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$, a nonlinear expectation is called $\mathcal{F}$-consistent expectation(or $\mathcal{F}$-expectation) if for each $X \in L^2(\Omega, \mathcal{F}_T, P)$ and for each $t \in [0, T]$ there exists a random variable $\eta \in L^2(\Omega, \mathcal{F}_t, P)$, such that*

$$\mathcal{E}[X 1_A] = \mathcal{E}[\eta 1_A], \quad \forall A \in \mathcal{F}_t.$$

From Lemma 3.1 above, such an η is uniquely defined. We denote it by $\eta = \mathcal{E}[X|\mathcal{F}_t]$. $\mathcal{E}[X|\mathcal{F}_t]$ is called the conditional $\mathcal{F}$-expectation of X under $\mathcal{F}_t$. It is characterized by

$$\mathcal{E}[X 1_A] = \mathcal{E}\Big[\mathcal{E}[X|\mathcal{F}_t] 1_A\Big], \quad \forall A \in \mathcal{F}_t. \tag{7}$$

Remark that, if f is a continuous, strictly increasing function on $\mathbf{R}$ such that $f(0) = 0$, $\mathcal{E}[X] = f^{-1}\Big(E[f(X)]\Big)$ defines an $\mathcal{F}$-expectation. Indeed, it is readily seen that $\mathcal{E}[X|\mathcal{F}_t] := f^{-1}\Big(E[f(X)|\mathcal{F}_t]\Big)$ satisfies (7).

The following lemma is obvious:

Lemma 3.2 *Let $g(\omega, t, y, z) : \Omega \times [0, T] \times R \times R^d \longmapsto R$ be a function satisfying (1), then the related g-expectation $\mathcal{E}_g[\cdot]$ is an $\mathcal{F}$-expectation.*

Lemma 3.3 *We have, for each $0 \leq s \leq t \leq T$,*

$$\mathcal{E}\Big[\mathcal{E}[X|\mathcal{F}_t]|\mathcal{F}_s\Big] = \mathcal{E}[X|\mathcal{F}_s] \quad a.s. \tag{8}$$

In particular,

$$\mathcal{E}\Big[\mathcal{E}[X|\mathcal{F}_t]\Big] = \mathcal{E}[X]. \tag{9}$$

□

Lemma 3.4 *We have a.s.*

$$\mathcal{E}[X 1_A|\mathcal{F}_t] = \mathcal{E}[X|\mathcal{F}_t] 1_A, \quad \forall A \in \mathcal{F}_t. \tag{10}$$

□

Lemma 3.5 *For any $X, \zeta \in L^2(\Omega, \mathcal{F}_T, P)$ and for each $t \in [0, T]$ and $A \in \mathcal{F}_t$ we have*

$$\mathcal{E}[X 1_A + \zeta 1_{A^C}|\mathcal{F}_t] = \mathcal{E}[X|\mathcal{F}_t] 1_A + \mathcal{E}[\zeta|\mathcal{F}_t] 1_{A^C}$$

□

Lemma 3.6 *For any X, $Y \in L^2(\Omega, \mathcal{F}_T, P)$, if $X \leq Y$ a.s., then we have for each $t \in [0,T]$,*

$$\mathcal{E}[X|\mathcal{F}_t] \leq \mathcal{E}[Y|\mathcal{F}_t] \quad a.s.$$

If moreover $\mathcal{E}[X|\mathcal{F}_t] = \mathcal{E}[Y|\mathcal{F}_t]$ a.s. for some $t \geq 0$, then $X = Y$ a.s. □

4 An $\mathcal{E}^\mu$-Dominated Expectation is a g–Expectation

We will state our main result. Recall that we have defined $\mathcal{E}^\mu[X] = \mathcal{E}_g[X]$ for $g \equiv \mu|z|$ and $\mathcal{E}^{-\mu}[X] = \mathcal{E}_g[X]$ for $g \equiv -\mu|z|$. We first study $\mathcal{F}$-expectations dominated by $\mathcal{E}^\mu$, for some large enough $\mu > 0$:

Definition 4.1 ($\mathcal{E}^\mu$-domination) *Given $\mu > 0$, we say that an $\mathcal{F}$-expectation $\mathcal{E}$ is dominated by $\mathcal{E}^\mu$ if*

$$\mathcal{E}[X + \eta] - \mathcal{E}[X] \leq \mathcal{E}^\mu[\eta], \ \forall X, \eta \in L^2(\Omega, \mathcal{F}_T, P) \tag{11}$$

Remark 4.1 *For any g satisfying (1) and (4), the associated g-expectation is dominated by $\mathcal{E}^\mu$, where μ is the Lipschitz constant in (1).*

Lemma 4.1 *If $\mathcal{E}$ is dominated by $\mathcal{E}^\mu$ for some $\mu > 0$, then*

$$\mathcal{E}^{-\mu}[\eta] \leq \mathcal{E}[X + \eta] - \mathcal{E}[X] \leq \mathcal{E}^\mu[\eta]. \tag{12}$$

□

Lemma 4.2 *If $\mathcal{E}$ is dominated by $\mathcal{E}^\mu$ for some $\mu > 0$, then $\mathcal{E}[\cdot]$ is, for all $\varepsilon \in]0,1]$ a continuous operator on $L^{1+\varepsilon}(\Omega, \mathcal{F}_T, P)$ in the following sense:*

$$\exists C > 0, \quad |\mathcal{E}[\xi_1] - \mathcal{E}[\xi_2]| \leq C\,\|\xi_1 - \xi_2\|_{L^{1+\varepsilon}}, \quad \forall \xi_1, \xi_2 \in L^2(\Omega, \mathcal{F}_T, P). \tag{13}$$

Proof The claim follows easily from Lemma 4.1 above and Proposition 2.4. □

Remark 4.2 *Note that Lemma 4.2 provides easy examples of $\mathcal{F}$-expectations that are not μ-dominated : just take $\mathcal{E}[X] = f^{-1}\Big(E[f(X)]\Big)$ with $f(x) = x^{\frac{5}{3}}$ and $\varepsilon = 1/2$ for instance.*

Until the end of the paper, we will deal with $\mathcal{F}$-expectations $\mathcal{E}[\cdot]$ also satisfying the following condition:

$$\mathcal{E}[X + \eta|\mathcal{F}_t] = \mathcal{E}[X|\mathcal{F}_t] + \eta, \quad \forall X \in L^2(\Omega, \mathcal{F}_T, P) \quad \text{and} \quad \eta \in L^2(\Omega, \mathcal{F}_t, P) \tag{14}$$

Remark 4.3 *The meaning of this assumption is: the nonlinearily of this expectation is due to the risk. Recall that, when $\mathcal{E}[\cdot]$ is a g-expectation, (14) means that g satisfies (4). We observe that an expectation $E_Q[\cdot]$ under a Girsanov transformation $\frac{dQ}{dP}$ satisfies this assumption.*

Our first result connected to (14) will consist in deducing '$\mathcal{E}^{\mu}$-domination at time t' from (11). This will be correctly stated and proved in Lemma 4.4, but we need first to introduce some new notation.

For a given $\zeta \in L^2(\Omega, \mathcal{F}_T, P)$, we consider the mapping $\mathcal{E}_\zeta[\cdot]$ defined by

$$\mathcal{E}_\zeta[X] = \mathcal{E}[X+\zeta] - \mathcal{E}[\zeta] : L^2(\Omega, \mathcal{F}_T, P) \longmapsto R. \tag{15}$$

Lemma 4.3 *If $\mathcal{E}[\cdot]$ is an $\mathcal{F}$-expectation satisfying (11) and (14), then the mapping $\mathcal{E}_\zeta[\cdot]$ is also an $\mathcal{F}$-expectation satisfying (11) and (14). Its conditional expectation under $\mathcal{F}_t$ is*

$$\mathcal{E}_\zeta[X|\mathcal{F}_t] = \mathcal{E}[X+\zeta|\mathcal{F}_t] - \mathcal{E}[\zeta|\mathcal{F}_t]. \tag{16}$$

□

Lemma 4.4 *Let $\mathcal{E}[\cdot]$ be an $\mathcal{F}$-expectation satisfying (11) and (14). Then, for each $t \le T$, we have a.s.*

$$\mathcal{E}^{-\mu}[X|\mathcal{F}_t] \le \mathcal{E}[X|\mathcal{F}_t] \le \mathcal{E}^{\mu}[X|\mathcal{F}_t],\ \forall X \in L^2(\Omega, \mathcal{F}_T, P).$$

This lemma is a simple consequence of the following one, whose proof is inspired by [1].

Lemma 4.5 *Let $\mathcal{E}_1[\cdot]$ and $\mathcal{E}_2[\cdot]$ be two $\mathcal{F}$-expectations satisfying (11) and (14). If*

$$\mathcal{E}_1[X] \le \mathcal{E}_2[X], \quad \forall X \in L^2(\Omega, \mathcal{F}_T, P),$$

then a.s. and for all t,

$$\mathcal{E}_1[X|\mathcal{F}_t] \le \mathcal{E}_2[X|\mathcal{F}_t], \quad \forall X \in L^2(\Omega, \mathcal{F}_T, P).$$

□

Lemma 4.6 *If $\mathcal{E}$ meets (11) and (14), there exists a positive constant C such that, for all X and η in $L^2(\Omega, \mathcal{F}_T, P)$, and for all $t \ge 0$,*

$$\mathcal{E}\Big[\mathcal{E}[X+\eta|\mathcal{F}_t] - \mathcal{E}[X|\mathcal{F}_t]\Big] \le C\|\eta\|_{L^2}.$$

□

We now state the main result of this talk: **a filtration-consistent nonlinear expectation is in fact a g–expectation.**

Theorem 4.1 (**Main theorem**) *We assume that an $\mathcal{F}$-expectation $\mathcal{E}[\cdot]$ satisfies (11) and (14) for some $\mu > 0$. Then there exists a function $g = g(t, z)$: $\Omega \times [0,T] \times R^d$ satisfying (1) and (4) such that*

$$\mathcal{E}[X] = \mathcal{E}_g[X], \quad \forall X \in L^2(\Omega, \mathcal{F}_T, P).$$

In particular, every $\mathcal{E}$-martingale is continuous a.s.

Moreover, we have $|g(t,z)| \leq \mu|z|$ for all $t \in [0,T]$.

5 Nonlinear Martingales

In order to prove the above main theorem, we need to study nonlinear martingales.

Definition 5.1 *A process $(X_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(0,T)$ is called an $\mathcal{E}$-martingale (resp. $\mathcal{E}$-supermartingale, -submartingale) if for each $0 \leq s \leq t \leq T$*

$$X_s = \mathcal{E}[X_t|\mathcal{F}_s], \ (resp. \ \geq \mathcal{E}[X_t|\mathcal{F}_s], \leq \mathcal{E}[X_t|\mathcal{F}_s]).$$

Lemma 5.1 *An $\mathcal{E}^{\mu}$-supermartingale (ξ_t) is both an $\mathcal{E}$-supermartingale and $\mathcal{E}^{-\mu}$-supermartingale. An $\mathcal{E}^{-\mu}$-submartingale (ξ_t) is both an $\mathcal{E}$-submartingale and $\mathcal{E}^{\mu}$-submartingale. An $\mathcal{E}$-martingale (ξ_t) is an $\mathcal{E}^{-\mu}$-supermartingale and an $\mathcal{E}^{\mu}$-submartingale.*

Proof It comes simply from the fact that, for each $0 \leq s \leq t \leq T$,

$$\mathcal{E}^{-\mu}[\xi_t|\mathcal{F}_s] \leq \mathcal{E}[\xi_t|\mathcal{F}_s] \leq \mathcal{E}^{\mu}[\xi_t|\mathcal{F}_s].$$

□

We will now prove throught two lemmas that every $\mathcal{E}$-martingale admits continuous paths.

Lemma 5.2 *For each $X \in L^2(\Omega, \mathcal{F}_T, P)$ the process $\mathcal{E}[X|\mathcal{F}_t]$, $t \in [0,T]$ admits a unique modification with a.s. càdlàg paths.*

Proof We can deduce from Lemma 5.1 that the process $\mathcal{E}[X|\mathcal{F}_t]$, $t \in [0,T]$, is an $\mathcal{E}^{-\mu}$-supermartingale. The proof is standard with an application of the downcrossing inequality recalled in Proposition 2.6

□

The following property is important for proving the main theorem.

Lemma 5.3 *For each $X \in L^2(\Omega, \mathcal{F}_T, P)$, let*

$$y(t) = \mathcal{E}[X|\mathcal{F}_t].$$

Then there exists a pair $(g(\cdot), z(\cdot)) \in L^2_{\mathcal{F}}(0,T; R \times R^d)$ *with*

$$|g(t)| \le \mu |z(t)| \tag{17}$$

such that

$$y(t) = X + \int_t^T g(s)ds - \int_t^T z(s)dB_s. \tag{18}$$

In particular, y admits a.s. continuous paths.

Furthermore, take $X' \in L^2(\Omega, \mathcal{F}_T, P)$, *put* $y'(t) = \mathcal{E}[X'|\mathcal{F}_t]$, *and let* $(g'(\cdot), z'(\cdot)) \in L^2_{\mathcal{F}}(0,T; R \times R^d)$ *be the corresponding pair. Then we have*

$$|g(t) - g'(t)| \le \mu |z(t) - z'(t)| \tag{19}$$

Proof Since

$$y(t) = \mathcal{E}[X|\mathcal{F}_t], \quad 0 \le t \le T,$$

is an $\mathcal{E}$-martingale, and since it is cadlag, it is a right-continous $\mathcal{E}^\mu$-submartingale (resp. $\mathcal{E}^{-\mu}$-supermartingale) and we know from the g-supermartingale decomposition theorem (Proposition 2.5) that there exist (z^μ, A^μ) and $(z^{-\mu}, A^{-\mu})$ in $L^2_{\mathcal{F}}([0,T]; R \times R^d)$ with A^μ and $A^{-\mu}$ càdlàg and increasing such that $A^\mu(0) = 0$, $A^{-\mu}(0) = 0$ and

$$y(t) = y(T) + \int_t^T \mu |z^\mu(s)| ds - A^\mu(T) + A^\mu(t) - \int_t^T z^\mu(s) dB_s.$$

$$y(t) = y(T) - \int_t^T \mu |z^{-\mu}(s)| ds + A^{-\mu}(T) - A^{-\mu}(t) - \int_t^T z^{-\mu}(s) dB_s.$$

Hence,

$$z^\mu(t) \equiv z^{-\mu}(t),$$
$$-\mu |z^\mu(t)| dt + dA^\mu(t) \equiv \mu |z^\mu(t)| dt - dA^{-\mu}(t),$$

whence

$$2\mu |z^\mu(t)| dt \equiv dA^\mu(t) + dA^{-\mu}(t).$$

It follows that A^μ and $A^{-\mu}$ are both absolutely continuous and we can write:

$$dA^\mu(t) = a^\mu(t) dt, \quad dA^{-\mu}(t) = a^{-\mu}(t) dt$$

with

$$0 \le a^\mu(t), \quad 0 \le a^{-\mu}(t).$$

We also have

$$a^{\mu}(t) + a^{-\mu}(t) \equiv 2\mu|z^{\mu}(t)|,$$

so, if we define

$$\begin{aligned} z(t) &= z^{\mu}(t) \\ g(t) &= \mu|z(t)| - a^{\mu}(t), \end{aligned}$$

we get (18) and (17).

Now, we prove (19). We have

$$\begin{aligned} y(t) - y'(t) &= \mathcal{E}[X|\mathcal{F}_t] - \mathcal{E}[X'|\mathcal{F}_t] \\ &= \mathcal{E}[X - X' + X'|\mathcal{F}_t] - \mathcal{E}[X'|\mathcal{F}_t] \\ &= \mathcal{E}_{X'}[X - X'|\mathcal{F}_t] \end{aligned}$$

Recall (Lemma 4.3 in Section 4) that $\mathcal{E}_{X'}[\cdot]$ is another $\mathcal{F}$-expectation satisfying (11) and (14). Thus there also exists a pair $(\tilde{g}(\cdot), \tilde{z}(\cdot)) \in L^2_{\mathcal{F}}([0,T]; R \times R^d)$ with

$$|\tilde{g}(t)| \leq \mu|\tilde{z}(t)| \tag{20}$$

such that the $\mathcal{E}_{X'}$-martingale $y(t) - y'(t)$ satisfies

$$y(t) - y'(t) = X - X' + \int_t^T \tilde{g}(s)ds - \int_t^T \tilde{z}(s)dB_s.$$

On the other hand, we have

$$y(t) - y'(t) = X - X' + \int_t^T [g(s) - g'(s)]ds - \int_t^T [z(s) - z'(s)]dB_s.$$

It follows then that

$$\tilde{g}(t) \equiv g(t) - g'(t), \quad \text{and} \quad \tilde{z}(t) \equiv z(t) - z'(t).$$

This with (20) yields (19). The proof is complete. □

Let us note the following easy consequence of Lemma 5.3 :

Lemma 5.4 *Let $\mathcal{E}[\cdot]$ be an $\mathcal{F}$-expectation satisfying (11) and (14). Then for each $X \in L^2(\Omega, \mathcal{F}_T, P)$ and $g \in L^2_{\mathcal{F}}(0,T)$ the process $\mathcal{E}[X + \int_t^T g(s)ds|\mathcal{F}_t]$, $t \in [0,T]$ is a.s. continuous.*

Proof Indeed, we can write

$$\begin{aligned}\mathcal{E}[X+\int_t^T g(s)ds|\mathcal{F}_t] &= \mathcal{E}[X+\int_0^T g(s)ds - \int_0^t g(s)ds|\mathcal{F}_t] \\ &= \mathcal{E}[X+\int_0^T g(s)ds|\mathcal{F}_t] - \int_0^t g(s)ds\end{aligned}$$

because of (14). The claim follows then easily from Lemma 5.3. □

To end this section, it is useful to remark that, by the same way as in Lemma 5.2, we can prove the following optimal sampling theorem for $\mathcal{E}$-martingales (resp. supermartingales, submartingales):

Lemma 5.5 *Let the process X be an $\mathcal{E}$-martingales (resp. supermartingale, submartingale), and let σ and τ be two stopping times such that $\sigma \leq \tau$ a.s.. Then*

$$\mathcal{E}[X_\tau|\mathcal{F}_\sigma] = X_\sigma \quad (\text{resp.} \leq, \geq).$$

6 $\mathcal{E}$-Supermartingale Decompositions

The above application of g–supermartingale decompositions of[13] is a key step. But it is not enough to prove the main theorem, We have to introduce $\mathcal{E}$-supermartingale decompositions.

Let a function f be given

$$f(\omega, t, y) : \Omega \times [0,T] \times R \longmapsto R$$

satisfying, for some constant $C_1 > 0$,

$$\begin{cases} \text{(i)} \ \ f(\cdot, y) \in L^2_{\mathcal{F}}(0,T), \quad \text{for each } y \in R; \\ \text{(ii)} \ \ |f(t,y_1) - f(t,y_2)| \leq C_1|y_1 - y_2|, \quad \forall y_1,\ y_2 \in R. \end{cases} \tag{21}$$

For a given terminal data $X \in L^2(\Omega, \mathcal{F}_T, P)$, we consider the following type of equation:

$$Y(t) = \mathcal{E}[X + \int_t^T f(s, Y(s))ds|\mathcal{F}_t] \tag{22}$$

Theorem 6.1 *We assume (21). Then there exists a unique process $Y(\cdot)$ solution of (22). Moreover, $Y(\cdot)$ admits continuous paths.*

The proof of this theorem is based on the following lemma and applications of the usual fix point theorem.

Lemma 6.1 *Define a mapping* $\Phi(y(\cdot)) : L^2_{\mathcal{F}}(0,T) \longmapsto L^2_{\mathcal{F}}(0,T)$ *by*

$$\Phi(y(\cdot))(t) = \mathcal{E}\Big[X + \int_t^T f(s, y(s))ds|\mathcal{F}_t\Big].$$

Then we have for all t:

$$E\Big[|\Phi(y_1(\cdot))(t) - \Phi(y_2(\cdot))(t)|^2\Big] \leq C_1^2 e^{\mu^2 T}(T-t)E\Big[\int_t^T |y_1(s) - y_2(s)|^2 ds\Big].$$

□

Theorem 6.2 (Comparison Theorem). *Let* Y *be the solution of (22) and let* Y' *be the solution of*

$$Y'(t) = \mathcal{E}[X' + \int_t^T [f(s, Y'(s)) + \phi(s)]ds|\mathcal{F}_t]$$

where $X' \in L^2(\Omega, \mathcal{F}_T, P)$ *and* $\phi \in L^2_{\mathcal{F}}(0,T)$. *If*

$$X' \geq X, \quad \phi(t) \geq 0, \quad dP \times dt\text{-a.e.}, \tag{23}$$

then we have

$$Y'(t) \geq Y(t), \quad dP \times dt\text{-a.e.} \tag{24}$$

(24) becomes equality if and only if (23) become equalities.

□

Our next result generalizes the decomposition theorem for g-supermartingales of[13] to continuous $\mathcal{E}$-supermartingales. The proof uses mainly arguments from [13].

Theorem 6.3 *(***Decomposition theorem for** $\mathcal{E}$**-supermartingales***) Let* $\mathcal{E}[\cdot]$ *be an* $\mathcal{F}$*-expectation satisfying (11) and (14), and let* (Y_t) *be a related continuous* $\mathcal{E}$*-supermartingale with*

$$E[\sup_{t\in[0,T]} |Y(t)|^2] < \infty.$$

Then there exists an $A(\cdot) \in L^2_{\mathcal{F}}(0,T;R)$ *such that* $A(\cdot)$ *is continuous and increasing with* $A(0) = 0$, *and such that* $Y(t) + A(t)$ *is an* $\mathcal{E}$*-martingale.*

□

7 Proof of Theorem 4.1

We are now ready to prove our main result.

Proof of Theorem 4.1 For each given $z \in R^d$, we consider the following forward equation

$$\begin{cases} dY^z(t) = -\mu|z|dt + zdB_t, \\ Y^z(0) = 0. \end{cases}$$

We have $E[\sup_{t\in[0,T]} |Y^z(t)|^2] < \infty$. It is also clear that Y^z is an $\mathcal{E}^\mu$-martingale, thus an $\mathcal{E}[\cdot]$-supermartingale. Indeed, we can write $Y^z(t) = \mathcal{E}^\mu[Y^z(T)|\mathcal{F}_t]$. From Theorem 6.3, we know the existence of an increasing process $A^z(\cdot)$ with $A^z(0) = 0$ and $E[A^z(T)^2] < \infty$, such that

$$Y^z(t) = \mathcal{E}[Y^z(T) + A^z(T) - A^z(t)|\mathcal{F}_t].$$

Or

$$Y^z(t) + A^z(t) = \mathcal{E}[Y^z(T) + A^z(T)|\mathcal{F}_t], \quad t \in [0,T].$$

Then, from Lemma 5.3, there exists $(g(z,\cdot), Z^z(\cdot)) \in L^2_{\mathcal{F}}(0,T; R \times R^d)$ with $|g(z,t)| \leq \mu|Z^z(t)|$ such that

$$Y^z(t) + A^z(t) = Y^z(T) + A^z(T) + \int_t^T g(z,s)ds - \int_t^T Z^z(s)dB_s.$$

We also have

$$|g(z,t) - g(z',t)| \leq \mu|Z^z(t) - Z^{z'}(t)|. \tag{25}$$

But on the other hand, since

$$Y^z(t) = Y^z(T) + \int_t^T \mu|z|ds - \int_t^T zdB_s,$$

it follows that

$$A^z(t) \equiv \mu|z|t - \int_0^t g(z,s)ds$$
$$Z^z(t) \equiv z$$

In particular, (25) becomes

$$|g(z,t) - g(z',t)| \leq \mu|z - z'|. \tag{26}$$

Moreover,

$$Y^z(t) + A^z(t) = Y^z(r) + A^z(r) - \int_r^t g(z,s)ds + \int_r^t zdB_s, \quad 0 \leq r \leq t \leq T,$$

and $Y^z(t) + A^z(t)$ is an $\mathcal{E}$-martingale. But with the assumption (14) one has, for each $z \in R^d$ and $r \le t$

$$\mathcal{E}[-\int_r^t g(z,s)ds + \int_r^t zdB_s|\mathcal{F}_r] = \mathcal{E}[Y^z(t) + A^z(t) - (Y^z(r) + A^z(r))|\mathcal{F}_r],$$

i.e.

$$\mathcal{E}[-\int_r^t g(z,s)ds + \int_r^t zdB_s|\mathcal{F}_r] = 0 \quad 0 \le r \le t \le T \tag{27}$$

Now let $\{A_i\}_{i=1}^N$ be a $\mathcal{F}_r$-measurable partition of Ω (i.e., A_i are disjoint, $\mathcal{F}_r$-measurable and $\cup A_i = \Omega$) and let $z_i \in R^d$, $i = 1, 2, \cdots, N$. From Lemma 3.5, and the fact that $g(0,s) \equiv 0$, it follows that

$$\begin{aligned} &\mathcal{E}[-\int_r^t g(\sum_{i=1}^N z_i 1_{A_i}, s)ds + \int_r^t \sum_{i=1}^N z_i 1_{A_i} dB_s|\mathcal{F}_r] \\ &= \mathcal{E}[\sum_{i=1}^N 1_{A_i}\left(-\int_r^t g(z_i,s)ds + \int_r^t z_i dB_s\right)|\mathcal{F}_r] \\ &= \sum_{i=1}^N 1_{A_i}\mathcal{E}[-\int_r^t g(z_i,s)ds + \int_r^t z_i dB_s|\mathcal{F}_r] \\ &= 0 \end{aligned}$$

(because of (27)). In other words, for each simple function $\eta \in L^2(\Omega, \mathcal{F}_r, P)$,

$$\mathcal{E}[-\int_r^t g(\eta,s)ds + \int_r^t \eta dB_s|\mathcal{F}_r] = 0.$$

From this, the continuity of $\mathcal{E}[\cdot]$ in L^2 given by (13) and the fact that g is Lipschitz in z, it follows that the above equality holds for $\eta(\cdot) \in L^2_{\mathcal{F}}(0,T;R^d)$:

$$\mathcal{E}\left[-\int_r^t g(\eta(s),s)ds + \int_r^t \eta(s)dB_s|\mathcal{F}_r\right] = 0. \tag{28}$$

We just have to prove now that

$$\mathcal{E}_g[X] = \mathcal{E}[X], \quad \forall X \in L^2(\Omega, \mathcal{F}_T, P).$$

To this end we first solve the following BSDE

$$\begin{aligned} -dy(s) &= g(t, z(s))ds - z(s)dB_s, \\ y(T) &= X. \end{aligned}$$

Since g is Lipschitz in z, there exists a unique solution $(y(\cdot), z(\cdot)) \in L^2_{\mathcal{F}}(0,T;R\times R^d)$. By the definition of g-expectation,

$$\mathcal{E}_g[X] = y(0).$$

On the other hand, using (28), one finds

$$\begin{aligned}\mathcal{E}[X] &= \mathcal{E}[y(0) - \int_0^T g(z(s),s)ds + \int_0^T z(s)dB_s] \\ &= y(0) + \mathcal{E}[-\int_0^T g(z(s),s)ds + \int_0^T z(s)dB_s] \\ &= y(0) = \mathcal{E}_g[X].\end{aligned}$$

It follows that this g-expectation $\mathcal{E}_g[\cdot]$ coincides with $\mathcal{E}[\cdot]$ and we are finished. □

1. P. Briand, F. Coquet, Y. Hu, J. Mémin and S. Peng, A converse comparison theorem for BSDEs and related properties of g-expectations, Electron. Comm. Probab, **5** (2000).
2. Z. Chen, A property of backward stochastic differential equations, C.R. Acad. Sci. Paris Math. **326** (1998), no 4, 483–488.
3. Z. Chen and L. Epstein, Ambiguity, Risk and Asset Returns in Continuous Time, Preprint (1999).
4. N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng and M.-C. Quenez, Reflected Solutions of Backward SDE and Related Obstacle Problems for PDEs, Ann. Probab. **25** (1997), no 2, 702–737.
5. Z. Chen and S. Peng, Continuous Properties of g-martingales, Preprint (1999).
6. Z. Chen and S. Peng, A general downcrossing inequality for g-martingales, Statist. Probab. Lett. **46** (2000), no. 2, 169–175.
7. F. Coquet, Y. Hu and J. Memin and S. Peng, Filtrition Consistent Nonlinear Expectations and Related g-Expectations, Prepublication 00–02, Jan. 2000, of IRMAR to appear in PTRF.
8. D. Duffie and L. Epstein, Stochastic differential utility, Econometrica **60** (1992), no 2, 353–394.
9. N. El Karoui, S. Peng and M.-C. Quenez, Backward stochastic differential equation in finance, Math. Finance **7**(1997), no 1, 1–71.
10. E. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation, Systems and Control Letters **14** (1990), no 1, 55-61.

11. S. Peng, A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation, Stochastics stoch. reports **38** (1992), no 2, 119–134.
12. S. Peng, BSDE and related g-expectation, in Pitman Research Notes in Mathematics Series, no.364, "Backward Stochastic Differential Equation", Ed. by N. El Karoui & L. Mazliak, (1997), 141–159.
13. S.Peng, Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob-Meyer's type, Prob. Theory Rel. Fields **113** (1999), no 4, 473-499.

Pricing and Hedging of Index Derivatives under an Alternative Asset Price Model with Endogenous Stochastic Volatility

David Heath

Eckhard Platen

University of Technology Sydney,
School of Finance & Economics and
Department of Mathematical Sciences,
PO Box 123, Broadway, NSW, 2007, Australia

Abstract

The paper discusses a financial market model that generates stochastic volatility using a minimal number of factors. These factors model the dynamics of different denominations of a benchmark portfolio. Asset prices are specified as transformations of square root processes. Numerical results for the pricing and hedging of standard derivatives on indices for this class of models are documented. This includes cases where the standard risk neutral pricing methodology fails but a form of arbitrage still exists. However, payoffs can be perfectly hedged. In addition, the term structure of implied volatilities is documented.

1 Introduction

Although the theoretical and practical importance of the well-known Black-Scholes model (BSM) cannot be underestimated, it is far from being satisfactory. The BSM assumes that geometric Brownian motion generates the asset price dynamics. The resulting deterministic volatility does not match historically observed stochastic volatility. For instance, in the context of option pricing, practitioners have to correct for implied volatility skews and smiles due to stochastic volatility. The at-the-money short term implied volatility of an index has typically a negative correlation with the index itself and leads to a negatively skewed implied volatility term structure.

This paper studies the problem of derivative pricing for a specific diffusion model, the *minimal market model* (MMM), proposed in [4] and [5]. It uses a minimal number of factors. These factors are modeled as square root processes under the real world probability measure. The basic building blocks of the MMM are the different denominations of a benchmark portfolio measured in units of the different primary assets. They determine key financial quantities including short rates, volatilities and risk premia. The MMM generates endogenously stochastic volatility without using any additional stochastic volatility process. In this paper it will be demonstrated how to price and hedge basic index derivatives under the MMM without relying on the standard risk neutral pricing methodology. In addition, we describe the implied volatility term structure that arises for European options on indices. Cases are studied where payoffs can be perfectly hedged but a form of arbitrage still exists. We tolerate a form of arbitrage because markets may be imperfect following shocks or turbulences and they differ in the degree of sophistication and maturity.

The paper is organized as follows. Section 2 describes the MMM. In Section 3 the prices of European style contingent claims are derived. Some numerical results are discussed in Section 4.

2 Minimal Market Model

2.1 Savings Accounts and Growth Optimal Portfolio

Let us define a *primary asset* as an income or loss producing tradeable asset, for instance, a stock or currency. We consider in a market the evolution of the prices of $d+1$ primary assets, $d \in \{1, 2, \ldots\}$, that are modeled on a filtered probability space $(\Omega, \mathcal{A}_T, \underline{\mathcal{A}}, P)$. Here the filtration $\underline{\mathcal{A}} = (\mathcal{A}_t)_{t\in[0,T]}$ fulfills the usual conditions with $\mathcal{A}_0$ being trivial, see [2].

We assume that each primary asset has its own time value. The time value of the domestic currency is expressed via the corresponding *savings account process* $B^0 = \{B^0(t),\, t \in [0, T]\}$, where

$$dB^0(t) = B^0(t)\, f^0(t)\, dt$$

for $t \in [0, T]$, $T \in (0, \infty)$ with $B^0(0) = 1$. This savings account accumulates interest continuously according to the *domestic short rate process* $f^0 = \{f^0(t),\, t \in [0, T]\}$, which describes the shortest forward rate for holding the domestic currency. Here the 0th primary asset is interpreted as the domestic currency. The time value of the jth primary asset is similarly modeled by the jth *savings account process* $B^j = \{B^j(t),\, t \in [0, T]\}$, where

$$dB^j(t) = B^j(t)\, f^j(t)\, dt \tag{1}$$

for $t \in [0, T]$ with $B^j(0) = 1$, $j \in \{1, 2, \ldots, d\}$. The jth *short rate process* $f^j = \{f^j(t),\, t \in [0, T]\}$ can be, for instance, a dividend rate or foreign interest rate. In summary, the jth savings account measures accumulated income or loss generated by the jth asset in units of the jth asset, $j \in \{0, 1, \ldots, d\}$.

The i,jth *exchange price* $X^{i,j}(t)$ is the price of one unit of the jth asset at time t measured in units of the ith asset. The jth *savings account price* $S^{i,j}(t)$ at time t, when measured in units of the ith primary asset, is given by

$$S^{i,j}(t) = X^{i,j}(t)\, B^j(t) \tag{2}$$

for $t \in [0, T]$ and $i, j \in \{0, 1, \dots, d\}$.

We assume that there exists a *growth optimal portfolio* (GOP) which is a strictly positive $\underline{\mathcal{A}}$-adapted, self-financing portfolio that when used as numeraire makes any benchmarked price process an $(\underline{\mathcal{A}}, P)$-local martingale. It can be shown that such a market is locally arbitrage free, see[1],[6]. Also, in cases where the risk neutral pricing methodology can be applied, the GOP can be shown, see[6], to be the inverse of the state price density, see[3]. Let the process $D^j = \{D^j(t),\ t \in [0, T]\}$ denote the jth denomination of the GOP, when it is measured in units of the jth primary asset, $j \in \{0, 1, \dots, d\}$. This means that $D^j(t)$ has the representation

$$D^j(t) = \sum_{\ell=0}^{d} \delta^\ell(t)\, S^{j,\ell}(t) \tag{3}$$

for $t \in [0, T]$ and $j \in \{0, 1, \dots, d\}$. Here $\delta^\ell(t)$ denotes the number of units of the ℓth savings account held at time t in the GOP. In the stylized version of the MMM considered here, the jth denomination $D^j(t)$ of the GOP at time t is specified as a transformation of the form

$$D^j(t) = (Y^j(t))^{q_j}\, \xi^j(t), \tag{4}$$

with jth *average GOP*

$$\xi^j(t) = \xi^j(0)\, \exp\left\{ \int_0^t \eta^j(s)\, ds \right\} \tag{5}$$

for $t \in [0, T]$ and $j \in \{0, 1, \dots, d\}$. The jth *growth rate* $\eta^j = \{\eta^j(t),\ t \in [0, T]\}$ governs, according to (5), the jth average GOP. We assume that η^j is a deterministic function of time. The jth *exponent*

$$q_j \in (0, \infty) \tag{6}$$

is constant, $j \in \{0, 1, \dots, d\}$. The jth *square root process* $Y^j = \{Y^j(t),\ t \in [0, T]\}$, that appears in (4), is characterized by the stochastic differential equation (SDE)

$$dY^j(t) = \frac{\nu^j}{4}\, \varphi^j(t)\, (1 - Y^j(t))\, dt + \sum_{k=1}^{d} \gamma^{j,k}(t)\, \sqrt{Y^j(t)}\, dW^k(t) \tag{7}$$

with jth *diffusion parameter*

$$\varphi^j(t) = \sum_{k=1}^{d} (\gamma^{j,k}(t))^2 \tag{8}$$

for $t \in [0,T]$ and initial value $Y^j(0) > 0$, $j \in \{0,1,\dots,d\}$. Here $W^1,\dots,W^d$ are independent standard Wiener processes. The jth *dimension* $\nu^j \in (2,\infty)$ is constant and the j,kth *volatility parameter* $\gamma^{j,k} : [0,T] \to (-\infty,\infty)$ is a deterministic function of time for $j \in \{0,1,\dots,d\}$, $k \in \{1,2,\dots,d\}$. Obviously, the square root process Y^j fluctuates around its reference level of one. The diffusion parameter φ^j controls the time scale of its evolution and the dimension ν^j the magnitude of extreme fluctuations. For larger dimension ν^j extreme fluctuations are less likely. Note that the SDE for this square root process has a unique solution with an explicitly known transition density. Since $\nu^j > 2$ this process remains strictly positive w.p.1, see [2].

2.2 Asset Price Dynamics

The jth *benchmarked savings account process* $\hat{S}^j = \{\hat{S}^j(t),\, t \in [0,T]\}$ is formed by the ratio

$$\hat{S}^j(t) = \frac{B^j(t)}{D^j(t)} \tag{9}$$

for $t \in [0,T]$ and $j \in \{0,1,\dots,d\}$. By application of the Itô formula we obtain from (9), (1), (4) and (7) for the jth benchmarked savings account the SDE

$$\begin{aligned} d\hat{S}^j(t) &= \hat{S}^j(t)\left[f^j(t) - r^j(t) + q_j\,\frac{\nu^j}{4}\,\varphi^j(t) + \frac{\varphi^j(t)\,q_j}{Y^j(t)}\left(\frac{q_j+1}{2} - \frac{\nu^j}{4}\right)\right]dt \\ &\quad - \hat{S}^j(t)\sum_{k=1}^{d}\sigma^{j,k}(t)\,dW^k(t) \end{aligned} \tag{10}$$

for $t \in [0,T]$ with initial value $\hat{S}^j(0) = \frac{1}{D^j(0)}$. Here the j,kth *volatility* of the jth benchmarked savings account has the form

$$\sigma^{j,k}(t) = \frac{q_j\,\gamma^{j,k}(t)}{\sqrt{Y^j(t)}} \tag{11}$$

for $j \in \{0,1,\dots,d\}$, $k \in \{1,2,\dots,d\}$ and $t \in [0,T]$, see [5]. To avoid redundant assets we assume that the volatility matrix $v(t) = [v^{k,i}(t)]_{k,i=0}^{d}$ with

$$v^{k,i}(t) = \begin{cases} 1 & \text{for} \quad k = 0 \\ \sigma^{i,k}(t) & \text{for} \quad k \in \{1,2,\dots,d\} \end{cases} \tag{12}$$

is, for all $t \in [0,T]$, invertible.

By assumption, all benchmarked savings account processes must be $(\underline{\mathcal{A}}, P)$-local martingales. Thus $\hat{S}^j$ has no drift in its SDE (10) and the jth short rate fulfills the relation

$$f^j(t) = \eta^j(t) + q_j\,\varphi^j(t)\left\{\frac{1}{Y^j(t)}\left[\frac{\nu^j}{4} - \frac{q_j+1}{2}\right] - \frac{\nu^j}{4}\right\} \tag{13}$$

for $t \in [0,T]$ and $j \in \{0,1,\dots,d\}$, see [5]. From (9), (10) and (1) we obtain by application of the Itô formula the dynamics of the jth denomination of the GOP in the form

$$\begin{aligned} dD^j(t) &= d\left(\frac{B^j(t)}{\hat{S}^j(t)}\right) \\ &= D^j(t)\left[\left(f^j(t) + \sum_{k=1}^{d}(\sigma^{j,k}(t))^2\right)dt + \sum_{k=1}^{d}\sigma^{j,k}(t)\,dW^k(t)\right] \end{aligned} \tag{14}$$

for $t \in [0,T]$ and $j \in \{0,1,\dots,d\}$. Using the ith and jth denominations of the GOP, the i,jth exchange price can be expressed by the ratio

$$X^{i,j}(t) = \frac{D^i(t)}{D^j(t)}.$$

Thus we obtain by the Itô formula, (14), (2) and (1) for the jth savings account, when measured in units of the ith asset, the SDE

$$dS^{i,j}(t) = S^{i,j}(t)\left[f^i(t)\,dt + \sum_{k=1}^{d}(\sigma^{i,k}(t) - \sigma^{j,k}(t))\left\{\sigma^{i,k}(t)\,dt + dW^k(t)\right\}\right] \tag{15}$$

for $t \in [0,T]$ with $S^{i,j}(0) = X^{i,j}(0)$ and $i,j \in \{0,1,\dots,d\}$. For $i=0$ equation (15) describes the dynamics of the jth savings account expressed in units of the domestic currency.

In this paper we consider the case $\nu_j = 3$, $q_j = \frac{1}{2}$, where the jth short rate is deterministic, see (13). The jth benchmarked savings account $\hat{S}^j$ is then the inverse of a three-dimensional Bessel process. This process is known to be a strict $(\underline{\mathcal{A}}, P)$-local martingale, see [8]. The standard risk neutral pricing measure would have the Radon-Nikodym derivative $\Lambda(t) = \frac{\hat{S}^0(t)}{\hat{S}^0(0)}$. However, in this case the process $\Lambda = \{\Lambda(t),\, t \in [0,T]\}$ is not a martingale. Therefore, for this example, the standard risk neutral approach cannot be applied.

3 Pricing and Hedging of Derivatives

In the following we apply the benchmark pricing methodology, proposed by Platen [6], which works also in many cases, where the well-known risk neutral approach fails. Let $H^0_{\bar{T}} = H^0_{\bar{T}}(Y^0(\bar{T}), \ldots, Y^d(\bar{T})) \in [0,\infty)$ be a *payoff* at the maturity date $\bar{T} \in (0,T]$, measured in units of the domestic currency. The corresponding nonnegative *benchmarked payoff* $H_{\bar{T}}$ is given by

$$H_{\bar{T}} = H_{\bar{T}}(Y^0(\bar{T}), \ldots, Y^d(\bar{T})) = \frac{H^0_{\bar{T}}}{D^0(\bar{T})}, \tag{16}$$

where we assume $E(|H_{\bar{T}}|) < \infty$.

As shown in [7], for this example and certain payoffs, a class of benchmarked derivative prices that allow perfect replication via a corresponding self-financing hedging strategy exists. A corresponding benchmarked pricing function $u : [0,\bar{T}] \times (0,\infty)^{d+1} \to [0,\infty)$ has to satisfy the partial differential equation (PDE)

$$\begin{aligned} &\Bigg(\frac{\partial}{\partial t} + \sum_{\ell=0}^{d} \frac{\nu^\ell}{4}\, \varphi^\ell(t)\,(1 - Y^\ell(t))\, \frac{\partial}{\partial Y^\ell} \\ &\quad + \frac{1}{2} \sum_{\ell,r=0}^{d} \sum_{k=1}^{d} \gamma^{\ell,k}(t)\gamma^{r,k}(t)\,\sqrt{Y^\ell(t)\,Y^r(t)}\,\frac{\partial^2}{\partial Y^\ell\,\partial Y^r}\Bigg) u(t, Y^0, \ldots, Y^d) \\ &\quad = 0 \end{aligned} \tag{17}$$

for $(t, Y^0, \ldots, Y^d) \in [0,\bar{T}) \times (0,\infty)^{d+1}$ with terminal condition

$$u(\bar{T}, Y^0, \ldots, Y^d) = H_{\bar{T}}(Y^0, \ldots, Y^d) \tag{18}$$

for $(Y^0, \ldots, Y^d) \in (0,\infty)^{d+1}$. For our example and certain payoffs there exist several solutions of this PDE.

Now we introduce for a benchmarked pricing function u the vector $c_u(t) = (c^0_u(t), \ldots, c^d_u(t))^\top$ with $c^0_u(t) = 1$ and

$$c^k_u(t) = -\sum_{\ell=0}^{d} \gamma^{\ell,k}(t)\,\sqrt{Y^\ell(t)}\,\frac{\partial \log(u(t, Y^0(t), \ldots, Y^d(t)))}{\partial Y^\ell} \tag{19}$$

for $t \in [0,\bar{T}]$ and $k \in \{1, 2, \ldots, d\}$. For such a function u the vector of proportions $\pi_u(t) = (\pi^0_u(t), \ldots, \pi^d_u(t))^\top$ of the values to be held in the savings accounts for a corresponding self-financing hedging portfolio, which replicates the payoff must satisfy the equation

$$\pi_u(t) = v^{-1}(t)\, c_u(t) \tag{20}$$

for $t \in [0, \bar{T}]$, see [7].

4 MMM for Index Derivatives

4.1 Dynamics of the GOP

We have according to (14), (8) and (11) the SDE

$$dD^0(t) = D^0(t) \left[\left(f^0(t) + \frac{(q_0)^2\, \varphi^0(t)}{Y^0(t)} \right) dt + \frac{q_0}{\sqrt{Y^0(t)}} \sqrt{\varphi^0(t)}\, d\bar{W}^0(t) \right], \tag{21}$$

where

$$d\bar{W}^0(t) = \frac{1}{\sqrt{\varphi^0(t)}} \sum_{k=1}^{d} \gamma^{0,k}(t)\, dW^k(t) \tag{22}$$

for $t \in [0, T]$. Note that since

$$d\langle \bar{W}^0 \rangle_t = \frac{\sum_{k=1}^{d} (\gamma^{0,k}(t))^2}{\varphi^0(t)}\, dt = dt \tag{23}$$

for $t \in [0, T]$, see (22) and (8), it follows by Lévy's theorem that $\bar{W}^0$ is a standard Wiener process. Combining (7) and (22) we can therefore write

$$dY^0(t) = \frac{\nu^0}{4}\, \varphi^0(t)\, (1 - Y^0(t))\, dt + \sqrt{\varphi^0(t)\, Y^0(t)}\, d\bar{W}^0(t) \tag{24}$$

for $t \in [0, T]$, where $Y^0(0) > 0$.

The value $D^0(t)$ of the GOP when dominated in units of the domestic currency at time t is given by the sum (3). For large d we can therefore interpret the GOP as a well diversified portfolio and thus as a market index, see [6]. In the next subsection we consider some basic derivatives on the GOP.

4.2 Zero Coupon Bond

One of the simplest index derivatives is formed by a zero coupon bond that pays one unit of the domestic currency at a maturity date $\bar{T} \in [0, T]$. Its benchmarked payoff at time $\bar{T}$ is given by $H_{\bar{T}} = \frac{1}{D^0(\bar{T})}$. Since the short rate is deterministic the benchmarked traded zero coupon bond price $\hat{P}^0_*(t,\bar{T})$ at time t can be expressed in the form

$$u_*(t, Y^0(t)) = \hat{P}^0_*(t,\bar{T}) = \frac{1}{D^0(t)} \frac{B^0(t)}{B^0(\bar{T})} = \frac{\hat{S}^0(t)}{B^0(\bar{T})} \tag{25}$$

for $t \in [0, T]$, where u_* solves the PDE (17) - (18). The process $\hat{P}^0_*(\cdot, \bar{T})$ is an $(\underline{\mathcal{A}}, P)$-supermartingale. The corresponding zero coupon bond price $P^0_*(t,\bar{T})$ measured in units of the domestic currency is then given by

$$P^0_*(t,\bar{T}) = D^0(t)\,\hat{P}^0_*(t,\bar{T}) = \frac{B^0(t)}{B^0(\bar{T})} \tag{26}$$

for $t \in [0,\bar{T}]$. An alternative benchmarked pricing function u, which also solves the PDE (17)–(18) and hedges perfectly the benchmarked payoff $\frac{1}{D^0(\bar{T})}$, is given by the conditional expectation

$$u(t, Y^0(t)) = \hat{P}^0(t,\bar{T}) = E\left(\frac{1}{D^0(\bar{T})}\,\middle|\,\mathcal{A}_t\right) \tag{27}$$

for $t \in [0,\bar{T}]$. This function can be explicitly computed using the well-known transition density of the square root process. In domestic currency the corresponding alternative bond price $P^0(t,\bar{T})$ is given by $P^0(t,\bar{T}) = D^0(t)\,\hat{P}^0(t,\bar{T})$.

Figure 1 shows the alternative zero coupon bond price $P^0(t,\bar{T})$ in domestic currency as a function of time t and initial value $D^0(t)$ of the GOP. We use in this paper the default parameter values $\bar{T} = 10$, $\nu^0 = 3$, $\varphi^0 = 0.04$, $\xi^0(0) = 1$, $\eta^0(t) = 0.065$, $t \in [0,\bar{T}]$ and $q_0 = 0.5$, see (13). This yields a domestic short rate of $f^0(t) = 0.05$. The traded zero coupon bond price, see (26), is for large values of the GOP similar to the alternative zero coupon bond price. However it is independent of the GOP value. For our example the bond prices $P_*(t,\bar{T})$ and $P(t,\bar{T})$ are therefore different. We call their difference the corresponding arbitrage amount.

4.3 European Call Option

For a European call option on the GOP with maturity date $\bar{T}$ the payoff expressed in units of the domestic currency is $(D^0(\bar{T}) - K)^+$, where K is the strike price. This means that the benchmarked payoff $\left(1 - \frac{K}{D^0(\bar{T})}\right)^+$ is bounded.

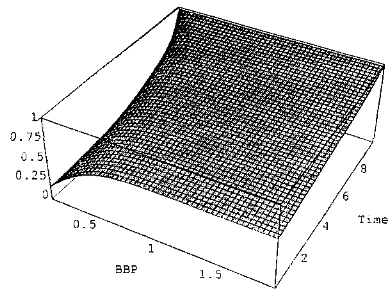


Figure 1: Alternative zero coupon bond price as a function of $D^0(t)$ and time t.

Consequently, it can be shown that the benchmarked call option price $\hat{c}^0_{\bar{T},K}(t, Y^0(t))$ is uniquely given by the conditional expectation

$$\hat{c}^0_{\bar{T},K}(t, Y^0(t)) = E\left(\left(1 - \frac{K}{D^0(\bar{T})}\right)^+ \Bigg| \mathcal{A}_t\right) \tag{28}$$

for $t \in [0, \bar{T}]$, see [7]. The function $\hat{c}^0_{\bar{T},K}$ satisfies the PDE (17)–(18). The corresponding price $c^0_{\bar{T},K}(t, Z^0(t))$ for this option, expressed in units of the domestic currency, is then $c^0_{\bar{T},K}(t, Y^0(t)) = D^0(t)\,\hat{c}^0_{\bar{T},K}(t, Y^0(t))$ for $t \in [0, \bar{T}]$.

The European put option price can be obtained by put-call parity. Note that since we have at least two different bond prices that satisfy the PDE (17) and (18), then different European put prices can be obtained, which all allow perfect hedging prescriptions.

The implied volatility term structure for European call options is typically used to asses the deviation of option prices from those obtained from the Black-Scholes model. In Figure 2 an implied volatility term structure for European calls is displayed using different values of time t and strike K. The implied volatility surface in Figure 2 shows a negative skew. This arises endogenously under the MMM without the need of using an external stochastic volatility process. In addition, the above model forms a complete market model. In summary, the above benchmark framework tolerates some form of arbitrage but insists on the perfect hedging of contingent claims, see [7].

This paper is intended to stimulate discussions on the degree to which arbitrage can be tolerated. In addition, it applies a simple one factor stochastic volatility model, which reproduces the typically observed negative skew in implied volatility.

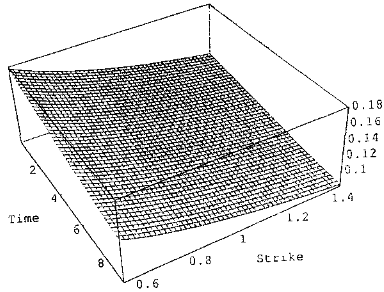


Figure 2: Implied volatilities for European calls.

References

1. Jamshidian, F. (1997). LIBOR and swap market models and measures. *Finance and Stochastics* 1(4), 293–330.
2. Karatzas, I. & S. E. Shreve (1988). *Brownian Motion and Stochastic Calculus.* Springer.
3. Karatzas, I. & S. E. Shreve (1998). *Methods of Mathematical Finance*, Volume 39 of *Appl. Math.* Springer.
4. Platen, E. (2000a). A minimal financial market model. Technical report, Humboldt University Berlin. Discussion Paper 91.
5. Platen, E. (2000b). Risk premia and financial modelling without measure transformation. Technical report, University of Technology, Sydney. QFRG Research Paper 45.
6. Platen, E. (2001a). Benchmark pricing model. University of Technology Sydney, (working paper).
7. Platen, E. (2001b). Derivative pricing with benchmark. University of Technology Sydney, (working paper).
8. Protter, P. (1990). *Stochastic Integration and Differential Equations.* Springer.

Risk Sensitive Asset Management With Constrained Trading Strategies [a]

Tomasz R. Bielecki
Department of Mathematics, The Northeastern Illinois University
5500 North St. Louis Avenue, Chicago, IL 60625-4699, USA
e-mail: T-Bielecki@neiu.edu

Daniel Hernandez-Hernandez
Centro de Investigación en Matemáticas
Callejón Jalisco s/n, AP 402, Guanajuato GTO, MEXICO
e-mail: dher@cimat.mx

Stanley R. Pliska
Department of Finance, University of Illinois at Chicago
601 S. Morgan Street, Chicago, IL 60607-7124, USA
e-mail: srpliska@uic.edu

This paper presents an application of risk sensitive control theory in financial decision making. A variation of Merton's continuous-time intertemporal capital asset pricing model is investigated where the investor's infinite horizon objective is to maximize the portfolio's risk adjusted growth rate. In earlier studies it was assumed either that the residuals associated with the assets are uncorrelated with the residuals associated with the factors or that there are no exogenous constraints like short selling restrictions. Here we develop computational procedures for the case where both of these assumptions are removed. Our approach is to first approximate the continuous time problem with a discrete time controlled Markov chain. We then solve the latter using the method of successive approximations for risk sensitive Markov decision chains. We show by numerical example that our approach is feasible, at least for cases where there are only a few factors. Our results suggest that when the hedging term is dominated by the myopic term for the corresponding unconstrained problem, then the optimal strategy computed for the constrained problem differs very little from the optimal myopic strategy for the same constrained problem.

1 Introduction

In a recent series of papers ([3], [4], [7], [5], [6], [1]), Bielecki and Pliska developed a variation of Merton's [13] intertemporal capital asset pricing model where, instead of maximizing expected utility of wealth at a fixed planning horizon, the investor's objective is to maximize his or her risk adjusted growth rate. This

[a]The research of the three authors was partially supported by NSF Grant DMS-9971307, CONACYT Grant 27933E, and NSF DMS-9971424, respectively.

criterion can be viewed as being analogous to the classical Markowitz single-period approach except that, instead of trading off single-period criteria, the investor is trading off the long run growth rate (which by itself is maximized by the growth optimal portfolio) versus risk as measured by the asymptotic variance of the portfolio. This risk adjusted growth rate objective emerges naturally from the application of recent mathematical results on risk sensitive control theory. A principal benefit of this objective is that it is an infinite horizon criterion and therefore, as with most control problems in general, the ICAPM is more tractable than if a finite horizon criterion is used.

The aforementioned work by Bielecki and Pliska focused on models with Gaussian factors and with assets having constant volatilities and appreciation rates that are affine functions of the factor levels. The model they considered will also be the subject of this paper. In particular, denoting by $S_i(t)$ the price of the i-th security and by $X_j(t)$ the level of the j-th factor at time t, they considered the following market model for the dynamics of the security prices and factors:

$$\frac{dS_0(t)}{S_0(t)} = (a_0 + A_0 X(t))dt, \quad S_0(0) = s_0. \tag{1}$$

$$\frac{dS_i(t)}{S_i(t)} = (a + AX(t))_i dt + \sum_{k=0}^{N} \sigma_{ik} dW_k(t), \quad S_i(0) = s_i, \quad i = 1, 2, \cdots, n \tag{2}$$

$$dX(t) = (b + BX(t))dt + \Lambda dW(t), \quad X(0) = x, \tag{3}$$

where $W(t) = (W_1(t), \ldots, W_N)'$ is a R^N valued standard Brownian motion process, $X(t) = (X_1(t), \ldots, X_m)'$ is the R^m valued factor process, the market parameters a_0, A_0, a, A, $\Sigma := [\sigma_{ij}]$, b, B, $\Lambda := [\Lambda_{ij}]$ are matrices of appropriate dimensions, and $(a + Ax)_i$ denotes the $i - th$ component of the vector $a + Ax$. Asset S_0 is interpreted as a bank account corresponding to the risk-less interest rate $a_0 + A_0 X(t)$.

The trading strategies h are n dimensional adapted processes with $h(t)$ always taking values in a specified set $\mathbf{U}$. The value of $h_i(t)$ is interpreted as the proportion of time-t wealth that is invested in asset i, so $1-(h_1(t)+\ldots+h_n(t))$ is interpreted as the proportion of time-t wealth invested in the bank account. If there are no exogenous restrictions on these proportions, that is, if $\mathbf{U} = \mathbf{R}^n$, then the model is referred to as the *unconstrained case*. On the other hand, if there are also short selling restrictions (e.g., $h_i(t) \geq 0$), borrowing restrictions (e.g., $h_1(t) + \ldots + h_n(t) \leq 1$), and the like, then $\mathbf{U} \neq \mathbf{R}^n$ and the model is referred to as the *constrained case*.

With $h(t)$ an admissible (there are also some technical requirements for admissibility, as specified in the aforementioned references) investment process, by standard results there exists a unique, strong, and almost surely positive solution $V(t)$ to the following equation:

$$\begin{aligned} dV(t) &= [a_0 + A_0 X(t)]V(t)dt \\ &\quad + \textstyle\sum_{i=1}^{m} h_i(t)V(t)\Big(\big[\mu_i(X(t)) - a_0 - A_0X(t)\big]dt \\ &\quad + \textstyle\sum_{k=0}^{N} \sigma_{ik} dW_k(t)\Big), \\ V(0) &= v > 0; \end{aligned} \tag{4}$$

where $\mu_i(x)$ is the i-th coordinate of the vector $a + Ax$ for $x \in R^{m+1}$. This solution is given by

$$\begin{aligned} V(t) = V(0)\exp\Big\{&\textstyle\int_0^t h'(s)[a + AX(s)]ds + \int_0^t [1 - h'(s)1][a_0 + A_0X(s)]ds \\ &-\tfrac{1}{2}\textstyle\int_0^t ||h'(s)\Sigma||^2 ds + \int_0^t h'(s)\Sigma dW(s)\Big\} \end{aligned} \tag{5}$$

The process $V(t)$, called the *value process* or *wealth process*, represents the investor's capital at time t, and $h_i(t)$ represents the proportion of capital that is invested in security i, so that $h_i(t)V(t)/S_i(t)$ represents the number of shares invested in security i, just as in, for example, Section 3 of[10].

In this paper, as well as in the aforementioned references [3], [4] etc., the investor seeks to find a trading strategy which solves the following risk sensitized optimal investment problem, labeled as $\mathcal{P}_\theta$:

for $\theta \in (0, \infty)$, maximize the risk sensitized expected growth rate

$$J_\theta(v, x; h(\cdot)) := \liminf_{t\to\infty} \Big(\frac{-2}{\theta}\Big)t^{-1} ln\ \mathbf{E}^{h(\cdot)}\ [e^{-(\frac{\theta}{2}) ln V(t)} | V(0) = v, X(0) = x] \tag{6}$$

over the class of all admissible investment processes $h(\cdot)$, subject to (1)-(4),

where $\mathbf{E}^{h(\cdot)}$ is the expectation with respect to $\mathbf{P}$. The notation $\mathbf{E}^{h(\cdot)}$ emphasizes that the expectation is evaluated for process $V(t)$ generated by (4) under the investment strategy $h(t)$. Here θ is a parameter that characterizes the investor's attitude toward risk; the bigger the value of θ, the more risk averse the investor.

The theoretical foundations of this model were initiated by Bielecki and Pliska in [3], and its applications to asset allocation problems were explored by Bielecki, Pliska, and Sherris in [7]. Both constrained and unconstrained cases were considered in these two papers. But while the model studied there fully allows for there to be correlations between asset returns and movements of the factor levels, it has a critical shortcoming: the partial correlations between asset returns and movements of the factor levels must be zero. In other words, the residuals of the asset returns must be independent of the residuals of the factors, that is, one must have $\Sigma\Lambda' = 0$. While this assumption is reasonable for some applications, such as where the risky assets are stocks and all the factors are macroeconomic variables, it is unacceptable for other applications such as where some factors are interest rates and some assets are fixed income securities.

For the purpose of some issues that come later in this paper, it is now convenient to repeat a result from [3]: the optimal trading strategy h can simply be obtained by solving a parametric quadratic program. In particular, if $X(t) = x \in \mathbf{R}^m$ then the optimal value of $h(t)$ is given by the solution of

$$\inf_{h\in \mathbf{U}} \left[\frac{1}{2}\Big(\frac{\theta}{2}+1\Big)h'\Sigma\Sigma' h - h'(a + Ax - a_0\mathbf{1} - A_0 x\mathbf{1})\right], \tag{7}$$

where $\mathbf{1}$ denotes an n-dimensional column vector of ones. This solution is called a *myopic* solution, in accordance with terminology adopted by Merton [13]. Even in the constrained case, quadratic programs like this can readily be solved with commercial software. Moreover, in the unconstrained case it is easy to see that the optimal value of $h(t)$ is explicitly given by

$$h(t) = \Big(\frac{\theta}{2}+1\Big)^{-1}[\Sigma\Sigma']^{-1}(a + Ax - a_0\mathbf{1} - A_0 x\mathbf{1}). \tag{8}$$

So with $\Sigma\Lambda' = 0$ it is easy to compute the optimal trading strategy, but with $\Sigma\Lambda' \neq 0$ matters are much more difficult. The unconstrained case when $\Sigma\Lambda' \neq 0$ was studied by Bielecki and Pliska first in [5] and then more extensively in [6]; the application of this kind of model to fixed income management was explored in [1]. Fleming and Sheu [9] and Kuroda and Nagai [11] have done closely related work. Suffice it to say that in the unconstrained case one can still obtain an explicit solution for the optimal trading strategy, but now there will be a second term (called the *hedging* component, in accordance with Merton's [13] terminology), in addition to the myopic component given by (7). Explicit results can be obtained for this situation because, in the unconstrained case, the minimizing selector in the Hamilton-Jacobi-Bellman equation can be specified explicitly.

This brings us to the subject of this paper: the constrained case when $\Sigma\Lambda' \neq 0$. Unfortunately, we do not know of any explicit results for this situation. Therefore, in this paper we develop a numerical approach. Roughly speaking, our approach is to approximate the continuous time problem with a controlled, discrete time Markov chain. In order to make this approximation we find it convenient to first transform our problem into an equivalent continuous-time problem that has a "classical" form in stochastic control theory; this is explained in Section 2. Then, following the general approach articulated by Kushner and Dupuis [12], we construct the approximating Markov decision chain. This Markov decision chain is classical, except that it still involves a risk sensitive optimality criterion; its construction and corresponding dynamic programming equation are presented in Section 3. To solve this discrete time, risk sensitive Markov decision chain we use the method of successive approximations that was developed by Bielecki, Hernandez-Hernandez, and Pliska [2]. This all is illustrated in Section 4 where we present a numerical example. This example is of independent interest because it is based upon market date and involves one factor (the bank account's interest rate) and two popular mutual funds. We conclude with some brief remarks in Section 5.

2 Transformation to a Classical, Continuous-Time, Risk Sensitive Control Problem

Throughout the remainder of this paper we make the following assumption about the set of admissible proportions for the risky assets:

Assumption. The set $\mathbf{U} \in \mathbf{R}^n$ is compact.

In view of the investor's objective (6) we are interested in using (5) to compute

$$E_x^{h(\cdot)}\left[e^{-\frac{\theta}{2}\ln V(t)}\right] = V(0)^{-\frac{\theta}{2}} E_x^{h(\cdot)} \exp\Big\{-\frac{\theta}{2}\Big[\int_0^t h'(s)[a + AX(s)]ds$$

$$+\int_0^t [1-h'(s)\mathbf{1}][a_0+A_0X(s)]ds - \frac{1}{2}\int_0^t ||h'(s)\Sigma||^2 ds + \int_0^t h'(s)\Sigma dW(s)\Big]\Big\}. \quad (9)$$

Denote

$$M(\theta, t) = \exp\Big\{-\frac{\theta^2}{8}\int_0^t ||h'(s)\Sigma||^2 ds - \frac{\theta}{2}\int_0^t h'(s)\Sigma dW(s)\Big\};$$

this is a martingale because $h(\cdot)$ is bounded and $E[M(\theta, t)] = 1$. For $T > 0$ define the probability measure

$$\tilde{P}_T^{h(\cdot),\theta}(A) := E^{h(\cdot)}[1_A M(\theta, T)], \quad A \in \mathcal{F}_T.$$

Under $\tilde{P}_T^{h(\cdot),\theta}$ the process

$$\tilde{W}(t) := W(t) + \frac{\theta}{2}\int_0^t (h'(s)\Sigma)' ds$$

is an $\mathcal{F}_T$-Brownian motion. Thus (9) can be written as

$$\mathbf{E}_x^{h(\cdot)}\left[e^{-\frac{\theta}{2}\ln V(t)}\right] = V(0)^{-\frac{\theta}{2}}\tilde{\mathbf{E}}_{T,x}^{h(\cdot),\theta}\exp\Big\{-\frac{\theta}{2}\int_0^t h'(s)[a+AX(s)]ds$$

$$-\frac{\theta}{2}\int_0^t [1-h'(s)\mathbf{1}][a_0+A_0X(s)]ds + \frac{\theta}{4}\left(\frac{\theta}{2}+1\right)\int_0^t ||h'(s)\Sigma||^2 ds\Big\}, \qquad (10)$$

where $\tilde{\mathbf{E}}_{T,x}^{h(\cdot),\theta}$ denotes the expectation operator under $\tilde{P}_T^{h(\cdot),\theta}$. Moreover, the dynamics of the factor process under the new Brownian motion are

$$dX(t) = (b+BX(t))dt - \frac{\theta}{2}\Lambda\Sigma' h(t)dt + \Lambda d\tilde{W}(t), \quad X(0)=x. \qquad (11)$$

It is now apparent that the investor's objective, to be maximized, can be written as

$$J_\theta(v,x;h(\cdot)) := \liminf_{t\to\infty}\frac{-2}{\theta}t^{-1}ln\ \tilde{\mathbf{E}}_{T,x}^{h(\cdot),\theta}\ \exp\Big\{-\frac{\theta}{2}\int_0^t R\big(X(s),h(s)\big)ds\Big\}, \quad (12)$$

where we have introduced the function

$$R(x,h) := [1-h\mathbf{1}][a_0+A_0x] + h'(a+Ax) - \frac{1}{2}\left(\frac{\theta}{2}+1\right)||h'\Sigma||^2.$$

We thus have transformed the investor's original problem $\mathcal{P}_\theta$ to an equivalent new one having classical, continuous-time dynamics for the state equation given by (11). Moreover, the investor seeks to maximize the classical risk sensitive objective function (12), where $R(x,h)$ can be interpreted as a "reward rate" function. Thus, as will be seen in the next section, we have arrived at a continuous time control problem that is in a convenient form for approximation by a discrete time, controlled Markov chain.

3 Approximation by a Discrete Time Markov Chain

In this section we explain how to approximate the continuous-time, risk sensitive, control problem presented at the end of the preceding section by a discrete-time, risk sensitive, controlled Markov chain. Moreover, at the end of this section we provide the dynamic programming equation for this risk sensitive Markov decision chain. To get there we follow the approximation approach described in Kushner and Dupuis [12].

The first step is to impose a boundary on the state space of the factor process X so that the new state space, denoted G, will be a compact subset of

$\mathbf{R}^m$. Changing from an unbounded to a compact state space is not expected to produce inaccurate computational results because we make the following

Assumption. The matrix B is stable.

This same assumption was made in earlier work by Bielecki and Pliska[3], and as a consequence the factor process X is ergodic with a point of stability (the so-called *mean reversion level*). Hence by making this assumption and choosing the boundaries of G so the stable point is well within the interior of G and far from its boundaries, one can anticipate the boundary behavior will have little bearing on the computational results. For convenience we take G to be a compact rectangle with simple normal reflection on its boundaries.

The second step is to impose a Cartesian grid on the state space G. We let δ denote its mesh size. Without loss of generality we assume the points on the "boundary" of this grid coincide with the boundary of G. Following Kushner and Dupuis[12] we now introduce the normalizing constant

$$Q := 2 \max_{x \in G, h \in U} \Big\{ \sum_i c_{ii} - \frac{1}{2} \sum_{i \neq j} |c_{ij}| + \delta \sum_i |(b + Bx - \frac{\theta}{2} \Lambda \Sigma' h)_i| \Big\},$$

where we have introduced the notation c_{ij} for the elements of the covariance matrix $(c_{ij}) := \Lambda\Lambda'$. The normalizing constant Q needs to be strictly positive, so we need to impose the following:

Assumption. $c_{ii} - \sum_{j \neq i} |c_{ij}| > 0$, for all i.

Remark. This assumption is somewhat challenging but not severe. It says that while the residuals of the factors can be correlated they cannot be highly correlated. Thus, for example, it might not be possible to take two interest rates as factors, although perhaps one could take one interest rate plus a spread.

The third step is to choose the time step for the approximating Markov chain. Denoting this by Δt, we follow Kushner and Dupuis[12] and take $\Delta t = \delta^2/Q$.

The fourth step is to set up the transition probabilities for the Markov chain. Throughout we let $p(x, y|h)$ denote the conditional, one-step transition probability that the next state is y given the current state is x and the current control action is h. Moreover, we let e_i denote an m-dimensional column vector consisting of all zeros except for a one as the ith component. Furthermore, for arbitrary real-number z we denote $z^+ := \max\{0, z\}$ and $z^- := \max\{0, -z\}$.

For x in the interior of the state space we then follow Kushner and Dupuis [12] and take

$$p(x, x \pm e_i\delta|h) = \frac{\frac{1}{2}c_{ii} - \frac{1}{2}\sum_{j\neq i}|c_{ij}| + \delta\big(b + Bx - \frac{\theta}{2}\Lambda\Sigma' h\big)^{\pm}}{Q},$$

$$p(x, x + e_i\delta + e_j\delta|h) = p(x, x - e_i\delta - e_j\delta|h) = \frac{c_{ij}^{+}}{2Q},$$

$$p(x, x - e_i\delta + e_j\delta|h) = p(x, x + e_i\delta - e_j\delta|h) = \frac{c_{ij}^{-}}{2Q},$$

$$p(x, y|h) = 0, \quad \text{for all other } y \neq x,$$

and

$$p(x, x|h) = 1 - \sum_{y\neq x} p(x, y|h).$$

For x on the boundary of the state space grid we take

$$p(x, x|h) = \min_{y\in \mathrm{int}G, h\in U} p(y, y|h)$$

and $p(x, y|h) = 1 - p(x, x|h)$ for y the closest neighbor to x in the interior of the grid. Note that, by our assumptions, each of these two probabilities will be strictly between zero and one, for every x on the boundary of G.

Recall the assumption we made earlier in this section that the matrix B is stable and thus the factor process X is ergodic. It follows from this and our other assumptions that the transition matrix for the discrete time Markov chain is irreducible. This is important because irreducibility is a requirement for convergence of our value iteration computational approach.

We now have our approximating risk sensitive Markov decision chain. Denoting this by $\{X_k; k = 0, 1, \ldots\}$ and the control by $\{u_k; k = 0, 1, \ldots\}$, the investor's objective is to choose an adapted control with $u_k \in \mathbf{U}$ so as to maximize

$$\liminf_{k\to\infty}(-2/\theta)\frac{1}{k\Delta t} ln\ \mathbf{E}\ \exp\Big\{-\frac{\theta}{2}\Delta t\sum_{i=1}^{k} R\big(X_i, u_i\big)\Big\}.$$

According to standard results (see, for example, Bielecki, Hernandez-Hernandez, and Pliska [2]), the dynamic programming equation for this discrete time problem is

$$e^{-\frac{\theta}{2}\lambda\Delta t + w(x)} = \inf_{u\in\mathbf{U}}\Big\{e^{-\frac{\theta}{2}R(x,u)\Delta t}\sum_{y\in G} e^{w(y)} p(x, y|u)\Big\}. \tag{13}$$

The solution consists of the scalar λ, which will turn out to correspond to the optimal objective value, and the bias function $w(\cdot)$. We know there will exist a solution, and this solution will be unique up to an additive constant for the bias function. A solution can be computed with the method of successive approximations (also called "value iteration"). The optimal control will be stationary and given by the minimizing selector in the dynamic programming equation. All this will be illustrated in the following section, where a numerical example is presented.

4 A Numerical Example: An Interest Rate Factor With Two Mutual Funds

In this section we present a numerical example that is based on eight recent years of daily price data for three funds at the College Retirement Equities Fund (CREF), a North American financial organization primarily for academics. These funds and the corresponding assets in our example are:

- S_0: CREF money market fund
- S_1: CREF stock fund
- S_2: CREF global equities fund

We used the money market price data to back out an interest rate, and we used the latter as the single factor in this example. Using regression and other simple statistical methods we then estimated the various model parameters. We found, for example that the interest rate process is stable and ergodic with a mean reversion level of 4.81%. Moreover, low interest rates are bullish for the global equities fund but mildly bearish for the stock fund. And, most importantly, the residuals for the interest rate factor are moderately correlated with the residuals for the two risky assets, so this example cannot simply be solved by using expression (7) for the optimal myopic strategy.

To set up the approximating Markov chain we chose $G = [0, .12]$ for the state space with $\delta = .0012$, thereby giving 101 states in the grid. For the set of admissible proportions we took $\mathbf{U} = \{(u_1, u_2) : u_1 \geq 0, u_2 \geq 0, u_1 + u_2 \leq 1\}$. Finally, for the risk aversion parameter we took $\theta = 10$.

In order to make a comparison and help validate our results, we next used the formulas of Bielecki and Pliska [6] to compute the optimal strategy and optimal objective value assuming there are no constraints on the proportions. The optimal risk adjusted growth rate in this unconstrained case is 12.07%.

Finally, using a standard spreadsheet, we implemented the method of successive approximations, starting with initial bias function $w_0(\cdot) = 0$, to solve the dynamic programming equation (13). After only 25 iterations we had $max_{x \in G}|w_{25}(x) - w_{24}(x)| = 0.000232$ and a corresponding risk adjusted growth rate of 13.26%, an unsatisfactory state of affairs. However, after 800 iterations we had $max_{x \in G}|w_{800}(x) - w_{799}(x)| = 0.000021$ and a corresponding risk adjusted growth rate of 10.17%. The corresponding optimal strategy is shown in Figure 1. It is interesting to note that the level of the interest rate seems to have a profound effect on the optimal proportions.

5 Concluding Remarks

We have presented a numerical approach for solving the Bielecki-Pliska risk sensitive asset management model in the constrained case where the residuals associated with the factors are correlated with the residuals associated with the assets (i.e., where $\Sigma\Lambda' \neq 0$). As demonstrated by our numerical example, our computational approach is feasible, at least when there are only a few factors. In particular, a rather coarse approximation of the state space seemed to give rather accurate results. On the other hand, since computation time is at least linear in the number of states, the overall computation time is at least exponential in the number of factors. Hence our approach appears to be impractical when there are more than, say, four or five factors.

Nevertheless, our numerical example does suggest an approximation that appears useful for some situations, even when there are many factors. We noted that the computed optimal strategy shown in Figure 1 is nearly identical (up to two if not three significant digits) to the optimal myopic strategy given by (7). We think this is related to the fact that, in the unconstrained case, the myopic component of the optimal strategy dominates the hedging component. We therefore conjecture for general problems that if in the unconstrained case the myopic term dominates the hedging term, then in the constrained case the optimal myopic strategy given by (7) (which is easy to compute!) will be a good approximation to the overall optimal strategy. Intuitively, if the hedging component is unimportant in the unconstrained case, then one can assume $\Sigma\Lambda' = 0$, in which case the myopic strategy given by (7) will apply. We must admit, however, that when there are many factors it might not be so easy to verify that the hedging term is dominated by the myopic term in the unconstrained case.

Another important point to emphasize is that our "approximation to optimality" results may not hold true for the whole range of values of the risk sensitivity parameter θ. One reason for this is that for some values of $\theta \in (0, \infty)$ there may not exist any optimal solution for the continuous time problem $\mathcal{P}_\theta$.

Finally, since we followed the Kushner and Dupuis [12] Markov chain approximation approach and since our numerical experiments produced logical and consistent results, there is every reason to believe that our approximation approach is an accurate one. But we are not certain of this. In particular, we would like to know that as our mesh parameter δ converges to zero (and as the boundary of the state space G expands in a suitable fashion) the optimal strategy for the risk sensitive Markov decision chain converges in a suitable way to the optimal strategy for the original, continuous time problem. Fitzpatrick and Fleming[8] demonstrated convergence results like this for a somewhat different kind of problem. Proving analogous convergence results for our model is a worthy topic for future research.

1. Bielecki, T.R., Harris, A., Li, J., and Pliska, S.R., (2001), "Risk Sensitive Asset Management: Two Empirical Examples," *Proceedings of the October 2000 Conference on Mathematical Finance in Konstanz, Germany*, edited by M. Kohlmann and S. Tang, Birkhäuser, Basel, Switzerland.
2. Bielecki, T. R., Hernandez-Hernandez, D., and Pliska, S.R., (1999), "Risk sensitive control of finite state Markov chains in discrete time, with applications to portfolio management," *Math. Meth. Oper. Res.*, vol. 50, pp. 167-188.
3. Bielecki, T.R., and Pliska, S.R., (1999), "Risk Sensitive Dynamic Asset Management," *Appl. Math. Optim.*, vol. 39., pp. 337-360.
4. Bielecki, T. R., and Pliska, S.R., (2000), "Risk-sensitive dynamic asset management in the presence of transaction costs," *Finance and Stochastics*, vol. 4, pp. 1-33.
5. Bielecki, T.R., and Pliska, S.R., (2001), "Risk Sensitive Intertemporal Control With Applications to Fixed Income Management," *European Congress of Mathematics, Barcelona, 10-14 July, 2000*, edited by C. Casacuberta, R.M. Miró-Roig, J. Verdera, and S. Xambó, Birkhäuser, Basel, Switzerland.
6. Bielecki, T.R., and Pliska, S.R., (2001), "A Risk Sensitive Intertemporal CAPM, With Application to Fixed Income Management," *working paper.*
7. Bielecki, T. R., Pliska, S. R., and M. Sherris (2000), "Risk Sensitive Asset Allocation," *Journal of Economic Dynamics and Control*, vol. 24, pp. 1145-1177.
8. Fitzpatrick, B.B., and Fleming, W.H., (1991), "Numerical Methods for an Optimal Investment-Consumption Model," *Math. Opns. Res.*, vol. 16, pp. 823-841.
9. Fleming, W. H., and Sheu, S. J., (2000), "Risk Sensitive Control and an Optimal Investment Model," *Mathematical Finance*, vol. 10, pp. 197-213.
10. Karatzas I., and Kou S.G., (1996), "On the pricing of contingent claims under constraints," *Annals of Applied Probability*, vol. 6, pp. 321-369.

11. Kuroda, K., and Nagai, H., (2000), "Risk-sensitive portfolio optimization on infinite time horizon," working paper, Osaka University.
12. Kushner H.J., and Dupuis G.P., (1992), "Numerical Methods for Stochastic Control Problems in Continuous Time", Springer-Verlag, Berlin.
13. Merton, R. C., (1973), "An intertemporal capital asset pricing model," *Econometrica* vol. 41, pp. 866-887.

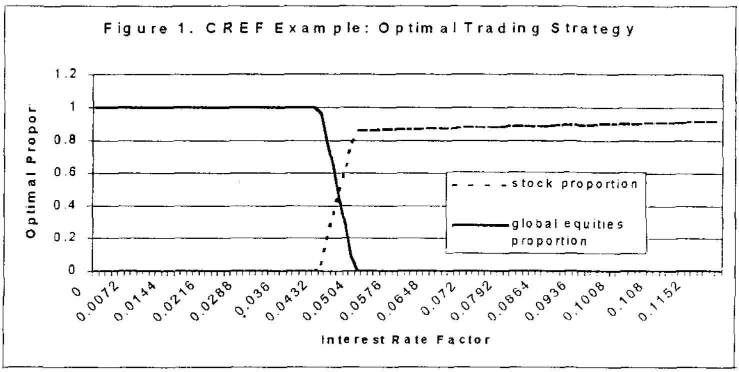


Figure 1. CREF Example: Optimal Trading Strategy

On Filtering in Markovian Term Structure Models

Carl Chiarella
School of Finance and Economics
University of Technology Sydney
PO Box 123, Broadway NSW 2007, Australia
email: carl.chiarella@uts.edu.au

Sara Pasquali
Dipartimento di Matematica
Università degli Studi di Parma
85, Via M. D'Azeglio, I - 43100 Parma, Italy
email: sara.pasquali@unipr.it

Wolfgang J. Runggaldier
Dipartimento di Matematica Pura ed Applicata
Università di Padova
7 Via Belzoni, I - 35131Padova, Italy
runggal@math.unipd.it

Abstract

We study a nonlinear filtering problem to estimate the market price of risk as well as the parameters in a given term structure model on the basis of noisy observations of forward rates. An approximation approach is described for the actual computation of the filter.

1 Introduction

The paper by Heath, Jarrow and Morton[12] (henceforth HJM) marked an important step in the development of models of the term structure of interest rates. The HJM model had been presaged by the simpler (and less general) Ho-Lee[13] model. The HJM model distinguished itself from previous term structure models, which were essentially based conceptually on the approach of Vasicek[21], by providing a pricing framework which is consistent with the currently observed yield curve and whose major input is a function specifying the volatility of forward interest rates. To this extent it can be viewed as the complete analogue, in the world of stochastic interest rates, to the Black-Scholes model of the deterministic interest rate world that prices derivatives consistently with respect to the price of the underlying asset (of which the currently observed yield curve is the analogue) and requires as its major input the volatility of returns of the underlying asset (to which the forward rate volatility function is the analogue).

The challenges posed in implementing the HJM model arise from the fact that in its most general form the stochastic dynamics are non-Markovian in nature. As a result most implementations of the HJM model revolve around some procedure, and/or assumptions, that allow the stochastic dynamics to be re-expressed in Markovian form - usually by employing the "trick" of expanding the state-space.

As we have stated above the major input into the HJM model is the forward rate volatility function and indeed its specification will determine the nature of the stochastic dynamics and whether and how it then can be reduced to Markovian form.

In view of finite dimensional realizations of HJM models (for a general study see [4]), Chiarella and Kwon [6] have shown that a broad, and important for applications, class of interest rate derivative models whose dynamics can be "Markovianised" can be obtained by assuming forward rate volatility functions that depend on a finite set of forward rates with given maturities as well as time to maturity.

An important practical problem faced in implementing such term structure models is the estimation of the parameters entering into the specification of the forward rate volatility function. In fact, one of the major aims of this paper is to show how this estimation problem can be approached within a filtering framework.

In section 2 we introduce our basic model that is a particular version of the Chiarella-Kwon [6] framework in which the volatility function depends on the instantaneous spot rate of interest (tenor of zero), one forward rate of fixed maturity and, time to maturity. Under the risk-neutral probability measure the stochastic dynamics of the spot rate and of the fixed maturity forward rate are given by a two-dimensional Markovian stochastic differential equation system. However as our observations occur under the so-called historical probability measure, we need to introduce also the market price of interest rate risk (that connects the two probability measures). We assume that the market price of risk forms a mean reverting process and so, under the historical measure, we are left with a three-dimensional Markovian stochastic differential system. A truncation factor is furthermore added to the coefficients thereby guaranteeing existence and uniqueness of a strong solution that takes values in a compact set. Assuming that the information comes from noisy observations of the fixed-maturity forward rate, in this same section 2 we also formulate the filtering problem, whose solution leads to the estimation of the market price of risk and of the unobserved instantaneous rates of interest and as well as of the parameters in the model.

The resulting filtering problem is highly nonlinear so that approximation methods have to be used for its solution. We shall describe a method, based on time discretization that, together with further approximations (quantization), leads to a discrete time approximating problem for which a filter of fixed finite dimension can be derived. Provided the discretization is sufficiently fine, the optimal filter for the approximating problem can be shown to be an arbitrarily good approximation to the filter for the original problem. Time and spatial discretization methods for nonlinear filtering were pioneered by H.Kushner and his co-workers (for a general exposition see [16]). Our method here differs in various respects from those in [16] and extends previous work in [10], [15], [20] (see also [18], [19] and the references in those papers).

In section 3 we discuss the time discretization and show the convergence of the time discretized filter for each observed trajectory and not merely in the mean with respect to the observations. We also mention further discretizations (quantizations) that lead to finite-dimensional approximating filters. We point out that the time discretization does not even need to be looked at as an approximation per se, since the real observations take place in discrete time only and so the true filtering problem is actually one in discrete time. In this sense the convergence of the time discretized filter can be viewed as guaranteeing the consistency of the discrete time models with the original continuous-time setup.

Due to space limitation, in this paper we only mention the main steps and main results; for technical details and proofs we refer the reader to the Working Paper [7].

2 Stochastic Dynamics and Filter Setup

Bhar, Chiarella, El-Hassan and Zheng [2] derived, within the HJM framework, a Markovian model for the term structure, where the Markovian factors are given by the unobserved short rate r_t [a] and a forward rate $f_t = f(t,\tau)$ with fixed maturity τ [b]. Given a volatility structure of the form

$$\sigma(t,T;r,f) = g(r,f)\, e^{-\lambda(T-t)} \tag{1}$$

with $0 \leq t < \tau < T$ and with $\lambda > 0$ a parameter, the dynamics of a generic forward rate $f(t,T)$, of the fixed maturity forward rate $f_t = f(t,\tau)$, and of the short rate $r_t = f(t,t)$ satisfy, according to [6], the

[a] The instantaneous short rate of interest, r_t, is treated as unobserved since the shortest rate we observe in most markets is a 30-day rate. In many empirical studies in finance this latter rate is treated as a proxy for r_t. Part of our contribution is the development of a methodology that avoids such an approximation. We should however also point out that [5] offers some justification for using the 30-day rate as a proxy for r_t.

[b] f_t is the rate we contract at t for instantaneous borrowing at time τ.

following equations

$$\begin{cases} df(t,T) = D_t(T)\,\sigma^2\,(t,T;r_t,f_t)\,dt + \sigma\,(t,T;r_t,f_t)\,d\tilde{w}_t \\ df_t = D_t\,\sigma^2\,(t,\tau;r_t,f_t)\,dt + \sigma\,(t,\tau;r_t,f_t)\,d\tilde{w}_t \\ dr_t = [A_t + B_t r_t + C_t f_t]\,dt + \sigma\,(t,t;r_t,f_t)\,d\tilde{w}_t \end{cases} \tag{2}$$

In this model, $\tilde{w}_t$ is a Wiener process on a filtered probability space $(\Omega,\mathcal{F},\mathcal{F}_t,Q)$ with Q the HJM "martingale measure". The function $g(r,f)$ in (1) is assumed to be of the form

$$g(r,f) = |\,a_0 + a_1 r + a_2 f\,|^{\delta} \tag{3}$$

for some positive parameters a_0,a_1,a_2,δ. Furthermore,

$$\begin{aligned} D_t(T) &= \lambda^{-1}\left(e^{\lambda(T-t)} - 1\right) \;;\quad D_t = D_t(\tau) \\ B_t &= -\lambda\,\left[\left(e^{-\lambda(\tau-t)} - 1\right)^{-1} + 1\right] \;; \\ C_t &= -\lambda e^{\lambda(\tau-t)}\left(e^{-\lambda(\tau-t)} - 1\right)^{-1} \;; \\ A_t &= f_T(0,t) - B_t f(0,t) - C_t f(0,\tau) \end{aligned} \tag{4}$$

where $f(0,t), f(0,\tau)$ are the initial forward rates for the maturities t and τ respectively, and $f_T(0,t)$ represents the partial derivative of $f(0,t)$ with respect to the second variable.

From model (2) we can derive the dynamics for the price $P(t,T)$ of a zero-coupon bond with generic maturity T, namely

$$dP(t,T) = P(t,T)\,[r_t dt - D_t(T)\sigma(t,T;r_t,f_t)\,d\tilde{w}_t] \tag{5}$$

We may now assume that agents can observe the prices of the available zero-coupon bonds. Since these prices have to be reconstructed from the actually accessible data, such observations have to be considered as noisy. On the other hand, there is a one-to-one correspondence between these prices and the forward rates and so we may as well assume that the agents have access to noisy observations of the latter (forward rates can be reconstructed from observable data). In order to make matters simple, we shall assume (see (11) below) that the available observations are noisy observations of the forward rate with the fixed maturity τ (one may obviously add noisy observations of forward rates with other maturities as well as of any other economic quantity, whose dynamics can be derived from (2)).

Since the observations take place under the "historical" or "real world" probability measure P, we shall also introduce the "market price of risk" process ψ_t, that corresponds to the translation of the Wiener when passing from the measure Q to P, and assume that it satisfies, under the measure P, a mean reverting diffusion model.

Denote then by X_t the "state" process

$$X_t := [f_t, r_t, \psi_t]' \tag{6}$$

and, given a (large) $H > 0$ and a (small) $\epsilon > 0$, let

$$\begin{aligned} \chi(X) &= \begin{cases} 1 & \text{if} \quad \max\{|f_t|,\, |r_t|\} \leq H \\ 0 & \text{if} \quad \min\{|f_t|,\, |r_t|\} \geq H + \epsilon\,; \\ \text{else a Lipschitz interpolation} \end{cases} \\ \bar{\chi}(\psi) &= \begin{cases} 1 & \text{if } |\psi| \leq H \\ 0 & \text{if } |\psi| > H + \epsilon \\ \frac{H+\epsilon-|\psi|}{\epsilon} & \text{if } H < |\psi| < H + \epsilon \end{cases} \end{aligned} \tag{7}$$

Under the measure P with Wiener process $w_t = \tilde{w}_t - \int_0^t \psi_s ds$, we now let the processes f_t, r_t, ψ_t satisfy the dynamics

$$\begin{cases} df_t = (D_t\, \sigma(t,\tau; r_t, f_t) + \psi_t)\, \sigma(t,\tau; r_t, f_t)\ \chi(X_t)\, dt \\ \qquad\quad + \sigma(t,\tau; r_t, f_t)\ \chi(X_t)\, dw_t \\ dr_t = [A_t + B_t r_t + C_t f_t + \psi_t \sigma(t,t; r_t, f_t)]\ \chi(X_t)\, dt \\ \qquad\quad + \sigma(t,t; r_t, f_t)\ \chi(X_t)\, dw_t \\ d\psi_t = \kappa\left(\bar{\psi} - \psi_t\right)\ \bar{\chi}(\psi_t) dt + b\, |\psi_t|^{\gamma} \bar{\chi}(\psi_t)\, dw_t \end{cases} \tag{8}$$

where the totality of the parameters is given by the vector

$$\theta := (a_0, a_1, a_2, \delta, \kappa, \bar{\psi}, b, \gamma, \lambda) \tag{9}$$

and each of them is supposed to take values in a compact subset of the positive halfline. With the vector X_t as in (6), we shall write the dynamics in (8) in compact form as

$$dX_t = F_t(X_t)\, dt + G_t(X_t)\, dw_t \tag{10}$$

where $F_t(\cdot)$ and $G_t(\cdot)$ are implicitly defined in (8) and, due also to the multiplicative factors $\chi(X_t)$ and $\bar{\chi}(\psi_t)$, satisfy a global Lipschitz property. As a consequence, equation (8) (or, equivalently, (10)) admits a unique strong solution that is furthermore bounded, i.e. $X_t \in \mathcal{X}$ with $\mathcal{X}$ a compact subset of $\mathbb{R}^3$. In what follows, the generic i–th $(i = 1, 2, 3)$ components of $F_t(\cdot)$ and $G_t(\cdot)$ will be denoted by $F_t^{(i)}(\cdot)$ and $G_t^{(i)}(\cdot)$ respectively.

Remark 2.1 *In the literature one can find results on the existence of a strong solution to equations of the form (2) with volatilities according to (1) and (3) (see e.g. [8]). These results hold however for specific ranges of the parameter δ in (3). In our application δ may take any positive value and so we preferred to introduce the Lipschitz truncation factors (7) to ensure in any case the existence of a strong and bounded solution. From a practical point of view this truncation is hardly any restriction at all.*

Model (10), resulting from (8) is a minimal Markovian model for the term structure : the dynamics of the various other forward rates $f(t,T)$ with generic maturity T (as well as the corresponding zero-coupon bond prices) can be derived from the first equation in (2) and from (5), whose dynamics depend only on the vector X_t.

In line with the foregoing, we shall assume that agents have access to noisy observations of $f_t = f(t,\tau)$. Denoting the observation process by y_t, we assume that it satisfies

$$dy_t = f_t dt + \hat{\epsilon} d\hat{w}_t \tag{11}$$

with $\hat{\epsilon} > 0$ small and $\hat{w}_t$ a $P-$Wiener, independent of w_t.

The goal here is a recursive Bayesian-type estimation of X_t and θ on the basis of the past and present observations of y_t, i.e. the combined filtering and parameter estimation of (X_t, θ), given $\mathcal{F}_t^y$, which is the filtration generated by the process y_t. The most complete solution to this problem is the recursive computation of the conditional joint distribution $p\,(X_t, \theta \mid \mathcal{F}_t^y)$. Tis is a highly nonlinear filtering problem and so in section 3

we shall compute a weak approximation to $p\,(X_t, \theta \mid \mathcal{F}_t^y)$ in the sense that we shall compute an approximation of the conditional expectation

$$E\left\{\bar{\Gamma}(X_t;\theta) \mid \mathcal{F}_t^y\right\} = \int \bar{\Gamma}(X;\theta)\, dp\,(X;\theta \mid \mathcal{F}_t^y) \tag{12}$$

where, for each θ, $\bar{\Gamma}(\cdot;\theta)$ is Lipschitz. The approximation is by discretization in time, which is motivated not only by the difficulty of computing (12) exactly, but also by the fact that, in reality, y_t is observed in discrete time. Additional possible approximations will also be mentioned in section 3

Remark 2.2 *Since the solution X_t of (10) takes values in the compact set $\mathcal{X}$, we may, without changing the value in (12), assume that $\bar{\Gamma}(X;\theta) = 0$ for $X \notin \mathcal{X}$. Notice also that from the econometric literature one has an indication of what could be possible values of the parameter vector θ. We shall thus assume that θ takes already from the outset only a finite number of possible values to which we may assign a uniform prior. This implies that the time discretization below concerns only the process X_t and, to emphasize this fact, we shall put $\Gamma_\theta(X) := \bar{\Gamma}(X;\theta)$ so that, instead of (12), we shall compute/approximate*

$$E\left\{\Gamma_\theta(X_t) \mid \mathcal{F}_t^y\right\} \tag{13}$$

Remark 2.3 *Stochastic filtering can be viewed as a dynamic generalization of Bayesian statistics. The "prior distribution" in this dynamic setup is given by the joint distribution of the (unobservable) state process*

X_t and of the parameter vector θ. This distribution is implied by the dynamic model for X_t (see (8) and (10)) and by the prior distribution on θ. This joint prior distribution is then successively updated on the basis of empirical data, namely of the noisy observations y_t of f_t. Analogously to classical Bayesian statistics, also in its dynamic generalization the "prior" is specified on the basis of extra-experimental information and/or on the basis of prior empirical information. This is also the sense in which one has to understand our double use of observations of forward rates : the one time initial observations of $f(0,t), f(0,\tau)$, $f_T(0,t)$ correspond to "prior" empirical information which is used, see (4), to determine the function A_t that is part of the dynamic model for X_t (see (8)), and thus of the "prior" for X_t. The successive noisy observations y_t of f_t on the other hand constitute the successively increasing empirical information, on which basis the prior of (X_t, θ) is being updated.

We want to point out that, in Bayesian statistics the current distributions turn out to be more informative, if one is able to assign a more informative prior. To this effect notice that, although the solution of (10) takes values in the compact set $\mathcal{X}$, there is no guarantee for the positivity of the instantaneous rates r_t and f_t. Since these rates are essentially positive, we should get more informative results if the "prior", i.e. our dynamic model for X_t guarantees positivity of these rates. For this purpose notice next that, if two quantities are in a one-to-one correspondence with each other, observing one of them or updating the distribution of one of them turns out to be equivalent to observing the other or updating its distribution respectively. We may therefore apply to the rates r_t and f_t an invertible transformation that transforms them into positive rates. For this purpose we use the $\mathcal{C}^2-$transformation

$$\bar{x} = T(x) := \begin{cases} x & \text{if } x \geq \epsilon + \eta \\ (\epsilon + \eta) + \frac{2\eta}{\pi} \arctan \left[\frac{\pi}{2\eta}(x - \epsilon - \eta)\right] & \text{if } x < \epsilon + \eta \end{cases} \tag{14}$$

where ϵ is, again, a small positive real and $0 < \eta < \epsilon$.

Define $\rho_t := T(r_t)$, $\phi_t := T(f_t)$ and notice that, with the same H as in (7), $\rho_t, \phi_t \geq T(-H-\epsilon) > \epsilon$ and, on $[\epsilon+\eta, H]$, we have $\rho_t = r_t$, $\phi_t = f_t$. Putting $\bar{X}_t := [\phi_t, \rho_t, \psi_t]'$, we may, with some abuse of notation, also write $\bar{X}_t = T(X_t)$ and, applying Ito's rule, obtain from (10)

$$d\bar{X}_t = \bar{F}_t(X_t)dt + \bar{G}_t(X_t)\,dw_t \tag{15}$$

where the $i-$th $(i = 1,2,3)$ components of $\bar{F}_t(\cdot)$ and $\bar{G}_t(\cdot)$ are

$$\begin{aligned} \bar{F}_t^{(i)}(X_t) &= \begin{cases} F_t^{(i)}(X_t) & \text{if } i = 3 \\ \dot{T}(X_t^{(i)})F_t^{(i)}(X_t) + \frac{1}{2}\ddot{T}(X_t^{(i)})(G_t^{(i)})^2(X_t) & \text{if } i = 1,2 \end{cases} \\ \bar{G}_t^{(i)}(X_t) &= \begin{cases} G_t^{(i)}(X_t) & \text{if } i = 3 \\ \dot{T}(X_t^{(i)})G_t^{(i)}(X_t) & \text{if } i = 1,2 \end{cases} \end{aligned} \tag{16}$$

and they are bounded since all the individual factors on the right in (16) are. Since $T(\cdot)$ is invertible, the Ito process $\bar{X}_t$ in (15) can be represented as solution of

$$d\bar{X}_t = \bar{F}_t(T^{-1}(\bar{X}_t))dt + \bar{G}_t(T^{-1}(\bar{X}_t))\,dw_t \tag{17}$$

that can be seen to admit a unique strong solution (see [7]).

In what follows we shall always refer to the same model (10) also in the case when we apply the transformation $T(\cdot)$. In this latter case X_t stands for $\bar{X}_t$, and the functions $F_t(X)$ and $G_t(X)$ then correspond to $\bar{F}_t\left(T^{-1}(\bar{X})\right)$ and $\bar{G}_t\left(T^{-1}(\bar{X})\right)$ respectively. Similarly, f_t in equation (11) stands for ϕ_t in case we apply the transformation $T(\cdot)$.

Notice that alternative approaches to obtain positive rates can be found in the recent literature (see e.g. [11]).

Notice finally that the filtering approach to HJM term structure models can also be seen as a possible way to overcome consistency problems in the calibration of HJM models (for the latter see e.g. the overview in [3]).

3 Time discretization and convergence results

In the following we implicitly assume that a generic value of θ has been fixed. Consider the partition of $[0,T]$ into subintervals of the same width $\Delta = \frac{T}{N}$ and perform an Euler discretization of (10), namely

$$X_{n+1}^N - X_n^N = F_n\left(X_n^N\right)\Delta + G_n\left(X_n^N\right)\Delta w_n \tag{18}$$

with $\Delta w_n = w_{(n+1)\Delta} - w_{n\Delta}$. Notice that, while the solution of the continuous-time model (10) is bounded, its discretized version (18) does not guarantee boudedness of (X_n^N). Denote by X_t^N the piecewise constant time interpolation of X_n^N, namely

$$X_t^N := \begin{cases} X_n^N & n\Delta \le t < (n+1)\,\Delta \\ X_N^N & t = T \end{cases} \tag{19}$$

and simply write X_n for $X_{n\Delta}^N$ as well as $X^{(i)}$ for the i−th component of X.

Consider next a Girsanov-type change of measure which allows us to transform the original filtering problem into one with independent state and observations. Denote by P^0 the measure under which y_t is a Wiener process, independent of X_t and thus also of X_t^N. In fact, the change of measure affects only the distribution of y_t and not also of X_t . The corresponding Radon-Nikodym derivative is

$$\frac{dP}{dP^0} = \exp\left[\frac{1}{\bar{\epsilon}^2}\int_0^T f_s dy_s - \frac{1}{2\bar{\epsilon}^2}\int_0^T f_s^2 ds\right] \tag{20}$$

Analogously, denote by P^N the measure under which y_t satisfies the equation

$$dy_t = f_t^N dt + \hat{\epsilon} dw_t^N \tag{21}$$

with w_t^N a P^N−Wiener process and where, with some abuse of notation, we denote by f_t^N the first component of X_t^N, truncated upon exit from $[-(H+\epsilon),(H+\epsilon)]$ (H and ϵ are the same as in (7)); as a consequence, in what follows f_t^N will be treated as having the same bounds as f_t. We thus have that, under P^N, y_t has the same form as under P, but as a function of the discretized state.

Applying the so-called Kallianpur-Striebel formula (see[14]), the filter in (13) can be expressed as

$$E\left\{\Gamma_\theta(X_t)\,|\,\mathcal{F}_t^y\right\} = \frac{E^0\left\{\Gamma_\theta\left(X_t\right)\frac{dP}{dP^0}\,|\mathcal{F}_t^y\right\}}{E^0\left\{\frac{dP}{dP^0}\,|\mathcal{F}_t^y\right\}} \tag{22}$$

It follows that it suffices to approximate, for each value of θ,

$$V_t\left(\Gamma_\theta;y\right) := E^0\left\{\Gamma_\theta\left(X_t\right)\frac{dP}{dP^0}\,|\mathcal{F}_t^y\right\} \tag{23}$$

(the denominator in (22) is in fact simply $V_t(1;y)$).

Define

$$z_t := E^0\left\{\frac{dP}{dP^0}|\mathcal{F}_t\right\} = \exp\left[\int_0^t \frac{1}{\hat{\epsilon}^2} f_s dy_s - \frac{1}{2\hat{\epsilon}^2}\int_0^t f_s^2 ds\right] \tag{24}$$

$$z_t^N := E^0\left\{\frac{dP^N}{dP^0}|\mathcal{F}_t\right\} = \exp\left[\int_0^t \frac{1}{\hat{\epsilon}^2} f_s^N dy_s - \frac{1}{2\hat{\epsilon}^2}\int_0^t (f_s^N)^2 ds\right] \tag{25}$$

where $\mathcal{F}_t = \mathcal{F}_t^y \vee \mathcal{F}_t^X$. By analogy to (23) define, for $N \in \mathbb{N}$,

$$V_t^N\left(\Gamma_\theta;y\right) := E^0\left\{\Gamma_\theta\left(X_t^N\right) z_T^N\,|\mathcal{F}_t^y\right\} \tag{26}$$

By the "smoothing property" of conditional expectations we have

$$\begin{aligned} V_t\left(\Gamma_\theta;y\right) &= E^0\left\{\Gamma_\theta\left(X_t\right) z_t\,|\mathcal{F}_t^y\right\}, \\ V_t^N\left(\Gamma_\theta;y\right) &= E^0\left\{\Gamma_\theta\left(X_t^N\right) z_t^N\,|\mathcal{F}_t^y\right\} \end{aligned} \tag{27}$$

We first have the following proposition, whose proof an be easily adapted from[10].

Proposition 3.1 *The processes* $\{X_t\}$ *and* $\left\{X_t^N\right\}$ *satisfy, for* $t \in [0,T]$

$$E\left\|X_t - X_t^N\right\|^4 \le K\Delta^2 \quad \text{and} \quad E^0\left\|X_t - X_t^N\right\|^4 \le K\Delta^2$$

where K *is a positive constant.*

Notice that, according to Remark 2.2, the value of $V_{n\Delta}^N(\Gamma_\theta; y)$ in (26) does not change if we change the values of X_t^N outside of $\mathcal{X}$. Consequently, we shall truncate the process X_t^N as soon as it exits from $\mathcal{X}$ and denote by X_n the so truncated process ($X_n^{(i)}$ will denote the i–th ($i = 1, 2, 3$) component of X_n and notice that for $f_n^N = f_n = X_n^{(2)}$ we have already used this truncation after (21)). The process X_n is now bounded Markov with a well-defined transition kernel $P(X_{n+1} \mid X_n)$.

We also make the following assumption, which is in line with our observation model (11)

Assumption A.1 : the actually observed trajectory (y_t) satisfies, for $n = 0, \cdots, N-1$,

$$\sup_{s,t \in [n\Delta,(n+1)\Delta]} | y_s - y_t | \leq K\Delta^{1/2}$$

Lemma 3.2 *Given an observed trajectory y_s $(s \leq t)$ satisfying A.1, we have for $t = n\Delta$*

$$E^0 \left\{ z_t^2 | \mathcal{F}_t^y \right\} \leq K(y) \quad ; \quad E^0 \left\{ (z_t^N)^2 | \mathcal{F}_t^y \right\} \leq K(y) \tag{28}$$

$$E^0 \left\{ \left| z_t - z_t^N \right| | \mathcal{F}_t^y \right\} \leq \bar{K}(y) \cdot \Delta^{\frac{1}{2}} \tag{29}$$

where $K(y)$, $\bar{K}(y)$ depend only on the obseved trajectory y_s, $s \leq t$.

Theorem 3.3 *For each $n = 0, 1, ..., N$, for $t = n\Delta$, for each observed trajectory y_s, $s \leq t$ satisfying A.1 and for each value of θ*

$$\left| V_t\left(\Gamma_\theta; y\right) - V_{n\Delta}^N\left(\Gamma_\theta; y\right) \right| \leq K_1(y)\Delta^{\frac{1}{2}}. \tag{30}$$

where $K_1(y)$ depends only on the obseved trajectory y_s, $s \leq t$.

The proofs of the Lemma and Theorem can be found in [7].

Remark 3.4 *Theorem 3.3 implies convergence of the filter for each observed trajectory. This is a stronger form of convergence than those in the traditional filtering literature (see e.g.[17]), where convergence is obtained in the mean with respect to y.*

Consider next the sequence of nonnegative measures $q_n(B; y^n)$, where B denotes the generic Borel subset of $\mathcal{X}$ and $y^n = (y_1^\Delta, \cdots, y_n^\Delta)$ with $y_n^\Delta := y_{n\Delta} - y_{(n-1)\Delta}$, and that are recursively defined by

$$\begin{aligned} q_0\left(B\right) &:= p_0\left(B\right) \\ q_{n+1}\left(B; y^{n+1}\right) &:= \int_B \int_{\mathcal{X}} \exp\left[\frac{1}{\bar{\epsilon}^2} f_n y_{n+1}^\Delta - \frac{1}{2\bar{\epsilon}^2} f_n^2 \Delta\right] \\ &\qquad \cdot P\left(X_{n+1} \left| X_n\right.\right) dq_n\left(X_n; y^n\right) dX_{n+1} \end{aligned} \tag{31}$$

where p_0 is the initial distribution and f_n corresponds to $X_n^{(1)}$, which is also the same as f_t^N in (21) and (25).

Proposition 3.5 *For any bounded function Ψ we have*

$$E^0\left[\Psi\left(X_n\right) z_T^N \left|\mathcal{F}_{n\Delta}^y\right.\right] = \int_{\mathbf{X}} \Psi\left(X\right)\, dq_n\left(X;y^n\right). \tag{32}$$

For a proof see e.g. [1].

Applying this proposition we immediately obtain (writing V_n^N for the $V_{n\Delta}^N$ according to (26))

$$V_n^N\left(\Gamma_\theta;y\right) = \int_{\mathbf{X}} \Gamma_\theta\left(X\right)\, dq_n\left(X;y^n\right) \tag{33}$$

for $n = 0, 1, ..., N$ and this also implies that, when computing $V_n^N\left(\Gamma_\theta;y\right)$, we do not lose information by considering only y^n instead of the entire filtration $\mathcal{F}_{n\Delta}^y$.

From (33) and (22) it is easily seen that the measures $q_n(B;y^n)$ can be given the interpretation of unnormalized conditional distributions. To determine the time discretized filter it suffices thus to compute the recursions (31). This is still an infinite-dimensional problem and so further approximations are needed, specifically discretizations in the spatial variable (quantization). This can be done in a variety of ways, for which we refer e.g. to [1,10,15,16,18,19,20]. In particular, for problems that are already reduced to discrete time, in [18,19,20] a specific methodology is described to arrive at a finite-dimensional approximating filter. Alternatively, always for problems already in discrete time, one could also use the recent so-called "particle approach" to nonlinear filtering, that is based on a simulation methodology (see e.g.[9]).

1. A.Bensoussan, W.Runggaldier, *An approximation method for stochastic control problems with partial observation of the state - a method for constructing $\epsilon-$optimal controls*, Acta Appl. Math., 10 (1987), 145-170.
2. R.Bhar, C.Chiarella, N.El-Hassan, X.Zheng, *The Reduction of Forward Rate Dependent Volatility HJM Models to Markovian Form: Pricing European Bond Options*, Journal of Computational Finance, 3(3), 47-72, 2000.
3. T.Björk, *A geometric view of interest rate theory*, SSE/EFI Working Paper Series in Economics and Finance No 419, Dec. 2000. To appear in *Handbook of Mathematical Finance*, Cambridge University Press.
4. T.Björk, L.Svensson, *On the existence of finite dimensional realizations for nonlinear forward rate models*, Mathematical Finance, 11 (2), 205-243, 2001.
5. D.A.Chapman, J.B.Long Jr., N.D.Pearson, *Using Proxies for the Short Rate : When Are Three Months Like an Instant ?*, Review of Financial Studies, 12, 763-806, 1999.

6. C.Chiarella, O.K.Kwon, *Classes of Interest Rate Models under the HJM Framework*, Asia Pacific Financial Markets, forthcoming 2001.
7. C.Chiarella, S.Pasquali, W.J.Runggaldier, *Filtering in Markovian Term Structure Models - an approximation approach*, Working paper, School of Finance and Economics, University of Technology, Sydney, 2001.
8. G.Deelstra, F.Delbaen, *Convergence of Discretized Stochastic (Interest Rate) Processes with Stochastic Drift Term*, Appl. Stoch. Models Data Anal. 14, 77-84, 1998.
9. P.Del Moral, *Measure valued processes and interacting particle systems. Applications to nonlinear filtering problems*, Ann.Appl.Probab., 8, 438-495, 1998.
10. G.B. Di Masi, M.Pratelli, W.J.Runggaldier, *An Approximation for the Nonlinear Filtering Problem, with Error Bound*, Stochastics, 14, 247-271, 1985.
11. B.Flesaker, L.P. Hughston, *Positive interest*, Risk 9, 46-49, 1996.
12. D.Heath, R.Jarrow, A.Morton, *Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation*, Econometrica, 60(1),77-105, 1992.
13. T.S.Y.Ho, S.B.Lee,*Term Structure Movements and Pricing Interest Rate Contingent Claims*,Journal of Finance, 1011-1029, 1986 .
14. G.Kallianpur, C.Striebel, *Estimation of stochastic processes. Arbitrary system processes with additive white noise observation error*, Ann. Math. Stat., 39 (1968), 785-801.
15. H. Korezlioglu, W.J. Runggaldier, *Filtering for Nonlinear Systems Driven by Nonwhite Noises : an Approximation Scheme*, Stochastics and Stochastics Reports, 44 (1993), 65-102.
16. H.J.Kushner, P.Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York 1992.
17. R.S. Liptser, O. Zeitouni, *Robust Diffusion Approximation for Nonlinear Filtering*, J.Math.Systems, Estimation and Control, 8 (1998), 1-22.
18. S. Pasquali, W.J. Runggaldier, *Approximations of a controlled diffusion model for renewable resource exploitation*, Kluwer Volume on Markov Processes and Controlled Markov Chains, in print.
19. W.J. Runggaldier, *On the construction of $\epsilon-$optimal strategies in partially observed MDP's*, Annals of Oper. Res., 28, 81-96, 1991.
20. W.J. Runggaldier, O. Zane, *Approximations for Discrete-Time Adaptive Control : Construction of $\epsilon-$Optimal Controls*, Math. Control Signals Systems, 4, 269-291, 1991.
21. O.Vasicek, *An Equilibrium Characterisation of the Term Structure*, Journal of Financial Economics, 5, 177-188, 1977.

A Theory of Volatility

Antoine Savine
Asia Group, BNP-Paribas
18th floor, Tokyo Sankei Building 1-7-2 Otemachi
Chiyoda-Ku, Tokyo 100-0004, Japan

Abstract

Bruno Dupire defined in 1993 an extension of Black & Scholes's model consistent with the market implied volatility smile on equity derivatives. This paper shows how Dupire's theory can be extended to take into account more realistic assumptions, such as interest rates, dividends, stochastic volatility and jumps. Moreover, we describe a similar theory in an interest rate setting, deriving a "caplet smile" formula, similar to Dupire's "local volaitlity" formula, defining an extended Vasicek model consistent with the market implied caplet smile. Our results also lead to an efficient calibration strategy for Markov interest rate models, and they are likely to considerably speed-up the calibration of non Gaussian models, with or without taking into account the volatility smile.

The first section introduces the fundamental mathematical concepts and notations, and presents a few applications. It also re-derives Dupire's formula in a different context. The second section shows how Dupire's theory can be extended to take into account realistic market considerations, such as interest rates, dividends, stochastic volatility and jumps. The last section derives similar results in an interest rate framework and shows how the Focker-Planck equation can be applied to design fast calibration algorithms. Sections 1 and 2 are a compilation of existing results presented in the framework of a unified theory. The results in the section 3 are new.

Our objective being the mathematical derivation of the results, we leave the discussions on the models themselves and the means to numerically implement them to subsequent papers. We also neglect the regularity conditions in order to focus on the concepts and the core of the calculations.

1 The mathematics of volatility and the smile theory

A) The theory of Distributions and Tanaka's formula

It is not the purpose of this paper to describe thoroughly the mathematics of Distributions. We rather go for a heuristic illustration of the theory, and jump to its applications and benefits. A proper mathematical formulation can be found in Schwartz and Hormander.

The familiar differential calculus formula:

$$df(t,x) = \frac{\partial f}{\partial t}(t,x)dt + \frac{\partial f}{\partial x}(t,x)dx$$

and its modification, known as Ito's lemma, intervening when X is a stochastic process:

$$dX_t = \mu_t dt + \sigma_t dW_t \tag{1}$$

then

$$df(t,X_t) = \frac{\partial f}{\partial t}(t,X_t)dt + \frac{\partial f}{\partial x}(t,X_t)dX + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,X_t)\sigma_t^2 dt \tag{2}$$

require the function f to be twice continuously differentiable. This is necessary in order to give a sense to the derivation sign ∂. What if it is not the case, like for the "call" function $c_K(x) = (x-K)^+$[a]?, the "digital" function $d_K(x) = 1_{\{x>k\}}$[b] or the Dirac mass $\delta_K(x)$[c]? Can we still apply (stochastic) differential calculus?

Clearly, derivatives cannot be defined in the traditional sense. The item "$\frac{\partial c_K}{\partial x}(x)$" is meaningless in the functional analysis context. The theory of distributions is an attempt to extend the notion of derivative to this type of functions, and subsequently extend differential calculus to handle them. It gives a sense to "$\frac{\partial c_K}{\partial x}(x)$" that is a natural extension of the traditional derivative. The call, digital and Dirac functions are then said to be "differentiable in the sense of distributions" (although they're not differentiable in the sense of functions), and their derivatives in this sense are defined as:

$$\frac{\partial c_K}{\partial x}(x) = d_k(x)$$

$$\frac{\partial^2 c_K}{\partial x^2}(x) = \frac{\partial d_k}{\partial x}(x) = \delta_K(x)$$

And the Dirac mass $\delta_K(x)$, defined by the property

$$\int f(x)\delta_K(x)dx = f(K)$$

[a] By $(x-K)^+$ we mean $\max\{x-K, 0\}$.
[b] By $1_{\{x>k\}}$ we mean 1 if $x > k$, 0 otherwise.
[c] There is a number of ways to define the Dirac mass. A popular approach is to consider it as the limit of a Gaussian distribution density, with expectation K, as its standard deviation decreases towards 0. A more proper mathematical approach defines it as a linear form defined on the linear space of smooth functions with compact support, by: $\delta_K(f) = \int f(x)\delta_K(x)dx = f(K)$.

has derivatives defined by defined by:

$$\int f(x)\delta_k^{(n)}(x)dx = (-1)^n f^{(n)}(K).$$

Derivatives in the sense of distributions satisfy a number of properties that make them a natural extension of derivatives in the sense of functions. Among them:

1. Any function is differentiable in the sense of distributions[d].

2. If a function is differentiable in the sense of functions, its derivative in the sense of functions and its derivative in the sense of distributions coincide.

3. If a function f is not differentiable in the sense of functions (but still satisfies the condition defined in footnote 4), then it is possible to define a series (f_n) of smooth functions, such that $f = lim_{n\to\infty} f_n$. And its derivative in the sense of distributions satisfy: $\frac{\partial f}{\partial x} = \lim_{n\to\infty} \frac{\partial f_n}{\partial x}$. The 3 most important cases as far as this paper is concerned are:
a) $\frac{\partial}{\partial x}(x-K)^+ = 1_{\{x>K\}}$
b) $\frac{\partial}{\partial x}1_{\{x>K\}} = \delta_K(x)$
c) $\delta_K^{(n)}$ is defined by: for any smooth function g with compact support

$$\int f(x)\delta_K^{(n)}(x)dx = (-1)^n f^{(n)}(K)$$

4. Derivatives in the sense of distributions are linear, i.e. for any pair of distributions f and g: $\frac{\partial}{\partial x}(\alpha f + \beta g) = \alpha\frac{\partial f}{\partial x} + \beta\frac{\partial g}{\partial x}$.

5. Derivatives in the sense of distributions satisfy the integration by parts formula: if f is a distribution and g is a smooth function with compact support, then

$$\int f(x)\frac{\partial g}{\partial x}(x)dx = -\int \frac{\partial f}{\partial x}(x)g(x)dx.$$

But the main property, as far as this paper is concerned, is that Ito's Lemma can be extended in the sense of distributions. The formula (2) holds if f is not smooth. In this case, the derivatives are to be understood in the sense of distributions and the formula is referred to as Tanaka's formula rather

[d]Actually, a technical condition is required (see *Schwartz*), but it is wide enough to handle functions of interest for us. The condition is: for every compact region K, it exists an integer p and a constant c, such that for every smooth function g, $|\int_K f(x)g(x)dx| \leq c\sup_{x\in K,|\alpha|\leq p}|g^{(\alpha)}(x)|$.

than Ito's lemma. For a proper derivation of Tanaka's formula, we refer to *Revuz-Yor* or *Karatzas-Schreve.*

B) Illustration 1: the stop-loss strategy and the local time

Let us assume an economy with no financing costs, and a call option written on a non dividend bearing stock, with quoted price S, the option maturing at time T with strike K. We assume that the stock has an absolute[e] instantaneous volatility of σ, without restrictive assumptions on the model. The volatility van be constant, time-dependent, stock price level-dependent, stochastic, etc.

Applying Tanaka's formula to $(S_t - K)^+$ we get:

$$(S_T - K)^+ = (S_0 - K)^+ + \int_0^T 1_{\{S_t > K\}} dS_t + \frac{1}{2} \int_0^T \delta_K(S_t)\sigma_t^2 dt.$$

The left member is the payoff of the call at maturity. The right member is the composition of 3 items:

1. $(S_0 - K)^+$ is the intrinsic value of the option.

2. $\int_0^T 1_{\{S_t > K\}} dS_t$ is the output of a self-financing strategy consisting in holding one unit of stock when spot is above strike, and no units when spot is below strike.

3. $\frac{1}{2} \int_0^T \delta_K(S_t)\sigma_t^2 dt$ is the mis-replication term. If the strategy above is applied in order to replicate the payoff of the option, the hedge will be down by $\frac{1}{2} \int_0^T \delta_K(S_t)\sigma_t^2 dt$. This term is called "local time at the strike", it represents the "time spent at the strike" weighted by the variance experienced at that moment. The intuitive interpretation of this term is as follows: if the spot does not cross the strike throughout the life of the option, the replication is perfect. If, on the contrary, the spot crosses the strike a large number of times, we would have to sell the stock for a price slightly below the strike, and buy it for a slightly higher price as the option comes back in the money. This phenomenon is not linked to transaction costs, but to what a Brownian Motion moves "too quickly"(the order of moves is in squared root of time) for us to be able to buy and sell exactly at the strike price as the option crosses the money.

---

[e] by absolute we mean that volatility is defined as the standard deviation of the absolute move in the asset price, and not that of its proportional move, as it is defined in general.

The local time is a random variable. Its risk-neutral expectation is equal to the time value of the option. Precisely,

$$TV = \frac{1}{2}E\Big[\int_0^T \delta_K(S_t)\sigma_t^2 dt\Big] = \frac{1}{2}E\Big[\int_0^T \delta_K(S_t)E\big(\sigma_t^2/S_t\big)dt\Big]$$

$$= \frac{1}{2}\int_0^T \int_{-\infty}^{\infty} \varphi_t(x)\delta_K(x)E\big(\sigma_t^2/S_t = x\big)dxdt$$

where φ_t is the risk-neutral density of the spot distribution at time t, and finally:

$$TV = \frac{1}{2}\int_0^T \varphi_t(K)E\big(\sigma_t^2/S_t = K\big)dt.$$

This shows that the time value of an option only depends on

1. The risk-neutral density at the strike.
2. The expected volatility experienced at the strike.

The reader can check, as an exercise, that in the standard Black-Scholes context, when $E\big(\sigma_t^2/S_t = K\big) = c^2K^2$, the time value is given by the Black-Scholes formula.

Furthermore, if we assume a liquid European option market, then the risk neutral density is given by:

$$\varphi_t(K) = \frac{\partial^2 C}{\partial K^2}(t, K).$$

This is discussed, for instance, by Dupire. Proof is straightforward using the Distribution theory:

$$\frac{\partial^2}{\partial K^2}(S_t - K)^+ = \delta_K(S_t),$$

therefore, using the expectation operator over this equality:

$$\frac{\partial^2 C}{\partial K^2}(t, K) = \varphi_t(K).$$

The standard deviation of the local time is significantly non zero, on the contrary of the theoretical standard deviation of the mis-hedge given by delta-replication in a Black-Scholes type model. Empirical results confirm that delta hedge performs way better than stop loss strategies (see, for example, *Hull*, pp. 296).

B) Illustration 2: derivation of the Focker-Planck formula

Let us consider a diffusion X defined by the Stochastic Differential Equation (SDE) :

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

Let us apply Tanaka's formula to the Dirac mass applied to X we get:

$$d\delta_K(X_t) = \left[\delta'_K(X_t)\mu(t, X_t) + \frac{1}{2}\delta_K"(X_t)\sigma_t^2\right]dt + \delta'_K(X_t)_\sigma(t, X_t)dW_t$$

Using the expectation operator over this equality and the Dirac mass derivation rules:

$$d\varphi_t(K) = \left\{ -\frac{\delta}{\delta K}[\varphi_t(K)\mu(t, K)] + \frac{1}{2}\frac{\delta^2}{\delta K^2}[\varphi_t(K)\sigma^2(t, K)]\right\}dt$$

Therefore

$$\frac{\delta}{\delta t}\varphi_t(K) = -\frac{\delta}{\delta K}[\varphi_t(K)\mu(t, K)] + \frac{1}{2}\frac{\delta^2}{\delta K^2}[\varphi_t(K)\sigma^2(t, K)].$$

This equation is known as the Focker-Planck equation, or the Kolmogorov-Forward equation. It puts into light the partial differential equation (PDE) satisfied by the density coming from a diffusion. Since the knowledge of risk-neutral densities is equivalent to that of all European option prices (see previous paragraph), the Focker-Planck equation provides a noticeable numerical benefit: it allows to calculate all European option prices in one single run of a one dimensional PDE grid. The use of it allows to considerably speed-up calibration processes in the equity world (when calibrating a local volatility model to the market smile for instance), as well as in a fixed income setting (when calibrating Markov interest rate models).

The standard Focker-Planck equation as presented above is a PDE on probability densities. In the presence of (possibly stochastic) interest rates, densities are not as interesting as Arrow-Debreu securities prices. Denote $\bar{\varphi}_t(x)$ the density and $\varphi_t(x)$ the corresponding Arrow-Debreu security price.

$$\bar{\varphi}_t(x) = E[\delta_{(x)}(X_t)], \qquad \varphi_t(x) = E\left[e^{-1\int_0^t r-sds}\delta_{(x)}(X_t)\right]$$

and φ satisfies the following PDE:

$$\frac{\delta}{\delta t}\varphi_t(K) = -\frac{\delta}{\delta K}\left[\varphi_t(K)\mu(t, K)\right] + \frac{1}{2}\frac{\delta^2}{\delta K^2}\left[\varphi_t(K)\sigma^2(t, K)\right] - r\varphi_t(K)$$

C) Dupire's smile theory

The Smile Theory, sometimes referred to as Dupire's Implied Diffusion Theory, puts into light the forward volatility assumptions implied by market prices of European options. It allows to price exotics consistently with all European prices, and hedge them accordingly. All major derivatives houses have Dupire's model implemented.

The theory is fully described in Dupire's paper. Our purpose here is to show that his main result can be proven in two lines, and easily extended, using Tanaka's formula.

Let us use the assumptions of the previous paragraph (no interest rates, non dividend bearing stock) plus the extra assumption that volatility directly depends on stock price: $\sigma_t = h(t, S_t)$, and apply Tanaka's formula, once again, to $(S_t - K)^+$:

$$d(S_t - K)^+ = 1_{\{S_t > K\}} dS_t + \frac{1}{2}\delta_K(S_t) h(t, S_t)^2 dt$$

Using the expectation operator over this equality we get:

$$dE\big[(S_t - K)^+\big] = 0 + \frac{1}{2}\varphi_t(K) h(t, K)^2 dt.$$

Since there is no interest rates, repo or dividends, $E[dS_t] = 0$. $E[(S_t - K)^+]$ is today's value of a call of maturity t and strike K, denoted $C(t, K)$, and

$$\frac{\partial C}{\partial t}(t, K) = \frac{1}{2}\varphi_t(K) h(t, K)^2.$$

Since $\varphi_t(K) = \frac{\partial^2 C}{\partial K^2}(t, K)$ (previous paragraph), we finally have:

$$\frac{\partial C}{\partial t}(t, K) = \frac{1}{2}\frac{\partial^2 C}{\partial K^2}(t, K) h(t, K)^2.$$

This formula, known as Dupire's formula, shows that in this simplified context (no interest, repo or dividends, volatility directly dependent on spot level), the necessary and sufficient condition on local volatility to hit all European option prices is:

$$h(t, K) = \sqrt{2\frac{\frac{\partial C}{\partial t}(t, K)}{\frac{\partial^2 C}{\partial K^2}(t, K)}}$$

2.- Beyond the smile theory

Dupire's formula allows to "calibrate" a diffusion model, so as to force it to hit the European option market, before using it to price more complex derivatives. Nonetheless, this formula has been derived in a simplified context. This section looks into its extensions under richer models considering a more realistic behavior of the market.

A) Interest rates, repo and dividends

In the real world, a little extra difficulty is introduced by the presence of interest rates, repo and (possibly absolute) dividends. In this section, we assume these market parameters are not stochastic. First of all, the presence of interest rates makes it necessary to discount call option prices:

$$C(y, K) = DF(0,t)E[(S_t - K)^+]$$

where $DF(s,t)$ is the price at time s of the zero coupon maturing at time t. It follows that

$$\frac{\partial C}{\partial K}(t, K) = -DF(0,t)E(1_{\{S_t > K\}})$$

and

$$\frac{\partial^2 C}{\partial K^2}(t, K) = DF(0,t)\varphi_t(K)$$

Secondly, the presence of interest rates, repo and dividends alter the risk-neutral SDE followed by the stock price. Let us assume the general form:

$$dS_t = \mu(t, S_t)dt + h(t, S_t)dW_t$$

with the drift being such that

$$E[dS_t] = [\alpha(t)S_t - \beta(t)]dt$$

where α and β are deterministic functions of time, possibly being a Dirac mass at absolute ex-dividend dates.

Once again, using the expectation operator over Tanaka's formula, we get:

$$dE[(S_t - K)^+] = E[1_{\{S_t > K\}}dS_t] + \frac{1}{2}E[\delta_K(S_t)h(t, S_t)^2]dt$$

multiplying both sides by $DF(0,t)$:

$$dC(t, K) = DF(0,t)E[1_{\{S_t > K\}}dS_t] + \frac{1}{2}DF(0,t)E[\delta_K(S_t)h(t, S_t)^2]dt$$

developing the last term as in section 4:

$$\frac{\partial C}{\partial t}(t, K) = \frac{DF(0,t)E[1_{\{S_t > K\}}dS_t]}{dt} + \frac{1}{2}\frac{\partial^2 C}{\partial K^2}(t, K)h(t, K)^2$$

and finally[f]:

$$\frac{\partial C}{\partial t}(t,K) = \alpha(t)C(t,K) + (\beta(t) - \alpha(t)K)\frac{\partial C}{\partial K}(t,K) + \frac{1}{2}\frac{\partial^2 C}{\partial K^2}(t,K)h(t,K)^2$$

Thus Dupire's formula extends to:

$$h(t,K) = \sqrt{2\frac{\frac{\partial C}{\partial t}(t,K) - [\alpha(t)C(t,K) + (\beta(t) - \alpha(t)K)\frac{\partial C}{\partial K}(t,K)]}{\frac{\partial^2 C}{\partial K^2}(t,K)}}$$

In the case of an absolute dividend D falling at time d, $\beta(t) = D\delta_d(t)$, and also, by arbitrage, $\frac{\partial C}{\partial t} = \frac{\partial C}{\partial K}D\delta_d(t)$, therefore the Dirac terms in the formula above disappear and it remains calculable.

B) Stochastic volatility

Now we are about to prove one of the most surprising results in the volatility theory, first discovered by Dupire and referred to as Unified Theory of Volatility (UTV).

Let us suppose zero interest rates, repo and dividends again in order to ease the notation, and let us assume a **generic** process $dS_t = \sigma_t dW_t$ on the stock price. By generic, we mean no restriction on the volatility term: it can depend on time or stock price level or even another hidden variable, it can follow its own process, etc. We actually cover all continuous models, the discontinuous ones being taken care of in the next section.

The result we present is as follows: Whatever the model, the necessary and sufficient condition in order to hit all European option prices today is:

$$E[\sigma_t^2/S_t = K] = 2\frac{\frac{\partial C}{\partial t}(t,K)}{\frac{\partial^2 C}{\partial K^2}(t,K)}$$

In the deterministic case we are back to the Dupire formula in the previous section.

As a proof, let us apply the expectation operator on Tanaka's expansion of $(S_t - K)^+$, as we usually do:

$$\frac{dE[S_t - K)^+]}{dt} = E[\delta_K(S_t)\sigma_t^2] = E[\delta_K(S_t)E(\sigma_t^2/S_t)] = \varphi_t(K)E(\sigma_t^2/S_t = K).$$

---

[f] Given that $\frac{DF(0,t)E[1_{\{S_t>K\}}dS_t]}{dt} = DF(0,t)E[1_{\{S_t>K\}}(\alpha(t)S_t - \beta(t))]$, which extends into $\alpha(t)DF(0,t)E[1_{\{S_t>K\}}(S_t - K)] - DF(0,t)E[1_{\{S_t>K\}}(\beta(t) - \alpha(t)K]$ and $\alpha(t)C(t,K) + (\beta(t) - \alpha(t)K)\frac{\partial C}{\partial K}(t,K)$.

To complete the proof, we just notice once again that $C(t,K) = E[(S_t - K)^+]$ and $\varphi_t(K) = \frac{\partial^2 C}{\partial K^2}(t,K)$.

C) Jumps

Let us complete the section by a word on discontinuities. Jump diffusion processes are fully treated in the excellent paper published by Andersen & Andreasen. We only show here how we can extend the previous results in the case of discontinuities. We consider an economy with no financing costs, dividends or repo to ease notations, and limit ourselves to a time and state dependent volatility.

In this case the risk-neutral process becomes

$$dS_t = -\lambda(t)m(t)dt + h(t,S_t)dW_t + J_t dN_t$$

where N is a Poisson process with jump intensity λ and J is a sequence of independent Gaussian variables G_t with mean $m(t)$ and standard deviation $s(t)$. N is the jump process, λ is its jump frequency, and the jump sizes at different times are modeled as random variables J. Since S has to remain a martingale, the drift term is necessary in order to compensate for the non-centered behavior of the jumps. We assume that W, N and J are independent.

An extension of Tanaka's formula allows to find the following result, determining the only function h that makes the model hit all European option prices today. It can be seen as a variation on Dupire's formula accounting for jumps. In other terms, jumps explain by themselves a considerable part of the smile, but not everything. The following PIDE (Partial Integrated Differential Equation) shows what h has to be in order to account for the difference.

Applying Tanaka's formula on $(S_t - K)^+$ in this context, we get:

$$d(S_t - K)^+ = 1_{\{S_t > K\}}[dS_t - J_t dN_t] + \frac{1}{2}\delta_K(S_t)h(t,S_t)^2 dt$$

$$+[(S_t + J_t - K)^+ - (S_t - K)^+]dN_t$$

and applying the expectation operator:

$$\frac{dE(S_t - K)^+}{dt} = -E(1_{\{S_t > K\}})\lambda(t)m(t) + \frac{1}{2}\varphi_t(K)h(t,K)^2$$

$$+\lambda(t)E[(S_t + J_t - K)^+ - (S_t - K)^+]$$

In other terms[g]:

$$\frac{\partial C}{\partial t}(t,K) = \frac{\partial C}{\partial K}(t,K)\lambda(t)m(t) + \frac{1}{2}\frac{\partial^2 C}{\partial K^2}(t,K)h(t,K)^2 - \lambda(t)C(t,K)$$
$$+\lambda(t)\int\int(x+y-K)^+\frac{\partial^2 C}{\partial x^2}(t,x)G_t(y)dxdy$$

Andersen & Andreasen give the means to efficiently solve the PIDE, with a particular focus on the non-standard term $\int\int(x+y-K)^+\varphi_t(x)G_t(y)dxdy$.

3.- The Interest Rate Smile theory

A) The model

While the theory presented above easily fits in an equity derivatives setting, it does not extend easily to term structure modeling. In this last section, we derive similar results in the context of a Markov model on the short rate, actually a Vasicek model extended to fit the market smile.

The celebrated Vasicek model has been widely described, for example by Jamshidian. It is based on a Gaussian mean-reverting risk-neutral dynamics for the short rate:

$$dr_t = \left\{\frac{\partial f(0,t)}{\partial t} + \psi_t - \lambda[r_t - f(0,t)]\right\}dt + \sigma(t)dW_t$$

where r is the short rate, λ is the mean-reversion, $f(t,T)$ is the instantaneous forward rate as defined by Heath-Jarrow-Morton (HJM), σ is the time-dependent normal volatility of the short rate, and ψ is the convexity bias and the solution of the following ordinary differential equation:

$$\psi_0 = 0, \qquad d\psi_t = [\sigma^2(t) - 2\lambda\psi_t]dt$$

The benefits of this model, also referred to as Linear Gauss Markov (LGM) lie in the following theorem (see Jamshidian for a proof):

1. The model dynamics is arbitrage-free in the sense of HJM[h]

[g] The previous formula extends to

$$-E(1_{\{S_t>K\}})\lambda(t)m(t) + \frac{1}{2}\varphi_t(K)h(t,K)^2$$
$$+\lambda(t)\int\int(x+y-K)^+\varphi_t(x)G_t(y)dxdy - \lambda(t)E[(S_t-K)^+]$$

[h] If $\mu_f(t,T)$ and $\sigma_f(t,T)$ denote the risk neutral drift and volatility of $f(t,T)$, the dynamics is arbitrage-free in the sense of HJM if $\mu_f(t,T) = \sigma_f(t,T)\int_0^T \sigma_f(t,u)du$, which is equivalent to r is the drift of all discount factors.

2. The dynamics is Markov, therefore all discount factors can be deduced from the short rate, more precisely[i]:

$$DF(t,T) = \frac{DF(0,T)}{DF(0,t)} exp\Big\{ -\frac{1-e^{-\lambda(T-t)}}{\lambda}[r_t - f(0,t)] - \frac{1}{2}\Big(\frac{1-e^{-\lambda(T-t)}}{\lambda}\Big)^2 \psi_t \Big\}$$

3. European caps and swaptions can be valued with a closed-form, see for example *El Karoui & Rochet.*

This last property makes the model particularly convenient to calibrate to the market prices of European instruments. However, the model is Gaussian, thus unable to generate anything but a Normal smile. It cannot be made to hit a given market volatility smile. Therefore *Ritchken & Sankarasubramanian* (among others) proposed an extension similar to Dupire's into the following model:

$$dr_t = \Big\{ \frac{\partial f(0,t)}{\partial t} + \psi_t - \lambda[r_t - f(0,t)] \Big\} dt + \sigma(t, r_t) dW_t$$

and proved that properties 1 and 2 hold, making the model tractable. An additional difficulty, however, comes from what ψ is not deterministic any more and captures a path-dependence:

$$\psi_0 = 0, \qquad d\psi_t = [\sigma^2(t, r_t) - 2\lambda\psi_t]dt$$

Therefore the system is bi-dimensional and degenerated in the direction of ψ (no diffusion coefficient), making it more delicate, though possible, to implement efficiently in a grid.

The main difference with the Gaussian model comes from property 3. There is no closed-form formula for European option pricing, and calibration has to be performed through numerical methods. An obvious one would be to use a grid to price each European option at each iteration of the procedure, but this is innefficient and requires a prohibitive computation time. A first order tweak would consist in pricing all European options with a single grid with one layer per calibration instrument. Unfortunately, a n layer grid has the same complexity as n single layer grids, the amount of computations that can be factorised is negligible and the increase in performance proves to be marginal.

The next paragraph is dedicated to an application of the Focker-Planck formula in order to price an arbitrary number of European instrument with only one run through a single layer grid, making the calibration of the model almost instantaneous.

[i]This property allows the model to be easily and efficiently implemented in a one-factor grid, for pricing any kind of contingent claims. ψ is deterministic.

B) Fast calibration

Let us call $\overline{\varphi}_t(x,y)$ the joint risk-neutral density of (r_t, ψ_t) and $\varphi_t(x,y)$ the corresponding Arrow-Debreu security price, defined by:

$$\overline{\varphi}_t(x,y) = E[\delta_{(x,y)}(r_t,\psi_t)], \qquad \varphi_t(x,y) = E\left[e^{-\int_0^t r_s ds}\delta_{(x,y)}(r_t,\psi_t)\right]$$

We shall prove the following theorem.

1. The profile at maturity of any interest rate contingent European claim fixing at time t can be written as a function $F(r_t, \psi_t)$.
2. The price of this claim is $\int_0^\infty \int_0^\infty F(x,y)\varphi_t(x,y)dxdy$.
3. The Arrow-Debreu security prices follow the following PDE:

$$\frac{\delta\varphi_t(x,y)}{\delta t} + \frac{\delta\left\{[\frac{\partial f(0,t)}{\partial t} + y - \lambda x]\varphi_t(x,y)\right\}}{\delta x} + \frac{\delta[(\sigma^2(t,x) - 2\lambda y)\varphi_t(x,y)]}{\delta y}$$

$$-\frac{1}{2}\frac{\delta^2[\sigma^2(t,x)\varphi_t(x,y)]}{\delta x^2} + x\varphi_t(x,y) = 0$$

subject to the initial condition $\varphi_0(x,y) = \delta_{(r_0,0)}(x,y)$.

1 is due to the reconstruction formula (property 2 of the model), which allows to deduce any interest rate, thus any contingent claim payoff, from r and ψ. 2 is immediately deduced from the property of Dirac masses. 3 is a direct application of the Focker-Planck formula presented in the section 1 of the present paper. These properties allow to value an arbitrary number of European options of different strikes and maturities with a single grid run. First, run the PDE 3 using a finite-difference method (FDM), keeping track of the Arrow-Debreu security values. Then, perform a bi-dimensional numerical integral to value European claims using property 2.

This methodology can be linked towards a global calibration procedure (value all calibration instruments at each iteration, minimize quadratic average error) or a local one (successively calibrate to instruments of increasing maturity using a partial grid between two maturities).

One efficient way to use this methodology is to link it to a parametric form for the local volatility. For instance, the Constant Elasticity Volatility (CEV) form is particularly popular among fixed income derivatives professionals:

$$\sigma(t, r_t) = \sigma(t)r_t^\beta$$

σ can be calibrated to the at-the-money (ATM) market implied volatility matrix, while β is set so as to hit the market skew. More complex parametric forms can be used to hit higher order market information such as kurtosis.

C) The caplet smile formula

Though the previous paragraph gives the means to quickly and efficiently calibrate an interest rate model consistent with the smile, it does not propose a direct mapping from market prices to local volatility, the way Dupire's formula does it in an equity framework. This is what the following is dedicated to.

We assume caplet prices $C(T,K)$ of all strikes K and maturities T are quoted in a liquid market[j] The link between caplet prices and our model is as follows.

$$C(T,K) = E[M_T(r_T - K)^+], \qquad M_T = e^{-\int_0^T r_t dt}$$

We can also define the price of digital options and Arrow-Debreu securities:

$$D(T,K) = E[M_T 1_{\{r_T > K\}}] = -\frac{\partial C(T,K)}{\partial K},$$

$$AD(T,K) = E[M_T \delta_K(r_t)] = \frac{\partial^2 C(T,K)}{\partial K^2}$$

Eventually, we will manipulate the price of parabolic contracts, delivering a parabolic profile at time T.

$$Par(T,K) = \frac{1}{2} E\{[M_T(r_t - K)^+]^2\} = \int_L^\infty C(T,y)dy$$

We shall prove the following result (caplet smile formula):

In order to hit caplet prices, the local volatility in the model has to be set according to the formula:

$$\frac{\sigma^2(t,K)}{2} = \frac{g(t,K) - \zeta(t,K)}{AD(t,K)}$$

where

$$g(t,K) = 2Par(t,K) + (\lambda + K)C(t,K) + \frac{\partial C(t,K)}{\partial t}$$
$$-\left\{\frac{\partial f(0,t)}{\partial t} - \lambda[K - f(0,t)]\right\} D(t,K)$$

is a function of market option prices, and

$$\zeta(t,K) = \int_0^\infty dy \int_y^\infty dx [y \varphi_t(x,y)]$$

[j] We work with "instantaneous" caplets delivering $(r_T - K)^+$ at time T.

where φ represents Arrow-Debreu security prices and can be valued using Focker-Planck's equation (see paragraph B).

Please note that this formula contains Dupire's formula

$$\sigma^2(t,K) = 2\frac{\frac{\partial C(t,K)}{\partial t}}{AD(t,K)}$$

as well as the following additional terms:

1. A term $2\frac{\lambda C(t,K)-\left\{\frac{\partial f(0,t)}{\partial t}-\lambda[K-f(0,t)]\right\}D(t,K)}{AD(t,K)}$ coming from the drift of r, exactly similar to the Dupire's formula with a drift (see section 2).

2. A term $2\frac{2Par(t,K)+KC(t,K)}{AD(t,K)}$ coming from the discounting of the options at the short rate. Since the short rate is at the same time the index and the discounting factor for the options, the development of a calendar spread will contain a squared r term, which can be locked through a parabolic profile.

3. A non-standard term $-2\frac{\zeta(t,K)}{AD(t,K)}$ coming from the presence of ψ in the drift of r. Since ψ captures a path-dependence, this term cannot be directly linked to option prices. However, it can be valued numerically.

Proof:

From the state variable dynamics:

$$\frac{dM_t}{M_t} = -r_t dt$$

$$d(r_t-K)^+ = \left[1_{\{r_t>K\}}\left\{\frac{\delta f(0,t)}{\delta t}+\psi_t-\lambda[r_t-f(0,t)]\right\}+\frac{1}{2}\delta_K(r_t)\sigma^2(t,r_t)\right]dt$$
$$+1_{\{r_t>K\}}\sigma(t,r_t)dW_t$$

Applying Tanaka's formula to $M_t(r_t-K)^+$ we get:

$$d[M_t(r_t-K)^+] = \left[M_t 1_{\{r_t>K\}}\left\{\frac{\delta f(0,t)}{\delta t}+\psi_t-\lambda[r_t-f(0,t)]\right\}\right.$$

$$\left.+\frac{1}{2}M_t\delta_K(r_t)\sigma^2(t,r_t)-M_t r_t(r_t-K)^+\right]dt+M_t 1_{\{r_t>K\}}\sigma(t,r_t)dW_t$$

and after applying the expectation operator:

$$\frac{dC(t,K)}{dt} = \begin{cases} (1) & E[M_t 1_{r_t>K}]\frac{\partial f(0,t)}{\partial t} \\ (2) & -\lambda E[M_t r_t 1_{\{r_t>K\}}] \\ (3) & +\lambda f(0,t)E[M_t 1_{\{r_t>K\}}] \\ (4) & +\frac{1}{2}E[M_t\delta_K(r_t)\sigma^2(t,r_t)] \\ (5) & -E[M_t r_t(r_t-K)^+]+E[M_t\psi_t 1_{\{r_t>K\}}] \end{cases} \tag{3}$$

(1) $E[M_t 1_{r_t>K}] = D(t,K)$ so (1) $= \frac{\partial f(0,t)}{\partial t} D(t,K)$

(2) $E[M_t r_t 1_{\{r_t>K\}}] = E[M_t(r_t-K)1_{\{r_t>K\}}]+KE[M_t 1_{\{r_t>K\}}] = C(t,K)+KD(t,K)$ so (2) $= -\lambda[C(t,K)+KD(t,K)]$

(3)$= \lambda f(0,t)D(t,K)$

(4)$= \frac{\sigma^2(t,K)}{2} AD(t,K)$

(5)$= -E\{M_t 1_{\{r_t>K\}}[(r_t-K)^2 + K(r_t-K)]\} = -2Par(t,K) - KC(t,K)$

(6) cannot be valued closed-form - let us denote it $\zeta(t,K)$.

And finally:

$$\frac{\sigma^2(t,K)}{2} = \frac{1}{AD(t,K)}\Big\{2Par(t,K) + (\lambda+K)C(t,K) + \frac{\partial C(t,K)}{\partial t}$$
$$-\{\frac{\partial f(0,t)}{\partial t} - \lambda[K - f(0,t)]\}D(t,K) - \zeta(t,K)\Big\}$$

or

$$\frac{\sigma^2(t,K)}{2} = \frac{g(t,K) - \zeta(t,K)}{AD(t,K)}$$

where

$$g(t,K) = 2Par(t,K) + (\lambda+K)C(t,K) + \frac{\partial C(t,K)}{\partial t}$$
$$-\{\frac{\partial f(0,t)}{\partial t} - \lambda[K - f(0,t)]\}D(t,K)$$

and

$$\zeta(t,K) = E[M_t \psi_t 1_{\{r_t>K\}}] = \int_0^\infty dy \int_K^\infty dx[y\varphi_t(x,y)]$$

References

1. L. Andersen, J. Andreasen, Jump Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Pricing, 1999.
2. B. Dupire, Pricing and Hedging with Smile, 1993.
3. B. Dupire, A Unified Theory of volatility, 1995.
4. N. El Karoui, N., H. Geman, J.-C. Rochet, Changes of Numeraire, Changes of Probability Measure and Option Pricing, 1995.
5. N. El Karoui, J.-C. Rochet, A Pricing Formula for Options on Coupon Bonds, 1989.

6. D. Heath, R. Jarrow, A. Morton, Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation, 1992.
7. L. Hormander, The analysis of Linear Partial Differential Operators, Springer-Verlag, 1983.
8. J. Hull, Options, Futures, and other Derivatives Securities Prentice Hall, 1993, Second Edition.
9. F. Jamshidian, Bond and Option Evaluation in the Gaussian Interest Rate Model, 1991.
10. I. Karatzas, S. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag.
11. D. Revuz, M. Yor, Continuous Martingales and Brownian Motion, Springer-Verlag.
12. P. Ritchken, L. Sankarasubramanian, Volatility Structures of Forward Rates and the Dynamics of the Term Structure, 1995.
13. L. Schwartz, Theorie des Distributions Hermann.

Discrete Time Markets with Transaction Costs

Lukasz Stettner
Institute of Mathematics, Polish Academy of Sciences
Sniadeckich 8, 00-950 Warsaw, Poland
Warsaw School of Management and Marketing [a]

Abstract

In the paper various aspects of arbitrage and pricing with concave and proportional transaction costs are studied.

1 Introduction

Assume that we are two accounts: bank and stock account. Transaction costs corresponding to transfers from the bank to the stock account and vice versa are described by the pair of functions τ_1 and τ_2, which transform R^+ into R^+, $\tau_1(0) = \tau_2(0) = 0$, τ_1 is concave and τ_2 is convex. Moreover, we assume that both functions are differentiable with $\dot{\tau}_1(0) = 1 + \lambda$, $\dot{\tau}_2(0) = 1 - \mu$ for positive λ and μ, $\dot{\tau}_1$ is decreasing greater than 1, while $\dot{\tau}_2$ is increasing less than 1. When we transfer ammount $\tau_1(x)$, $x \geq 0$ from the bank to the stock account (buy assets), the last account is increased by x. In the case, when we transfer ammount x from the stock to the bank account (sell assets), the bank account is increased by $\tau_2(x)$. By the form of the functions τ_1 and τ_2 we see that when they are not linear (i.e. $\tau_1(x) \not\equiv (1 + \lambda)x$ and $\tau_2(x) \not\equiv (1 - \mu)x$), buying or selling small ammounts of assets we pay almost proportional transaction costs, while for large transactions the transaction costs are relatively (per unit of the transaction) smaller. Such situation is significantly different from the one with proportional transaction costs, since the transaction costs are not additive (buying or selling assets for $x_1 + x_2$ is less expensive than buying or selling assets first for x_1 and then for x_2).

In this paper we mainly concentrate on strictly concave transaction costs (the situation when τ_1 and τ_2 are nonlinear), trying to treat the proportional transaction costs as a particular case of the concave transaction costs. For simplicity of the presentation we restrict ourselves to the case with one asset or one stock account. All results except of those concerning Cox Ross Rubinstein model can be easily extended to finite number of assets. Our market position is usually characterized by the vector $\binom{x}{y}$, where x corresponds to the ammount on the bank account, while y to the ammount on the stock account. We admit negative values for x or y, which correspond to debits. Moreover for simplicity we assume that rate of interests on the banking account is equal to 0. Given a

[a]The paper supported by grant no. PBZ KBN 016 / P03 / 99

probability space (Ω, F, P) endowed with an increasing family of σ - fields (F_t), we consider the prices S_t of the asset at time $t = 0, \ldots, T$ as F_t measurable random variables which are strictly positive. We assume for simplicity that σ - field F_0 is trivial. In what follows we denote by $L^0(G, F_T)$ the set of all F_T measurable random variables taking values in a set G, which may be also random. Moreover when $\tau_1(x) \equiv (1+\lambda)x$ and $\tau_2(x) \equiv (1-\mu)x$ we say that we have proportional transaction costs, while in the case when $\tau_1(x) \not\equiv (1+\lambda)x$ and $\tau_2(x) \not\equiv (1-\mu)x$ we say about nonlinear concave transaction costs.

2 Arbitrage

Let

$$\begin{aligned} M = \Big\{ \tbinom{x}{y} : y \geq -(\tau_1)^{-1}(x), \quad &for \quad x \geq 0 \\ and \quad y \geq -(\tau_2)^{-1}(-x), \quad &for \quad x \leq 0 \Big\}, \end{aligned} \tag{1}$$

where $(\tau_1)^{-1}$, $(\tau_2)^{-1}$ denote the inverse functions to τ_1, τ_2 respectively. One can notice that M is the set of all nonnegative market positions (portfolios), i.e. positions from which we can enter $\binom{0}{0}$. The set $-M$ is of the form

$$\begin{aligned} -M = \Big\{ \tbinom{x}{y} : y \leq (\tau_1)^{-1}(-x), \quad &for \quad x \leq 0 \\ and \quad y \leq -(\tau_2)^{-1}(x), \quad &for \quad x \geq 0 \Big\} \end{aligned} \tag{2}$$

and is set of all positions which can be achieved starting from $\binom{0}{0}$.
We say that we have a *weak arbitrage at time T* (WA) if starting from the position $\binom{0}{0}$ at time 0 we enter the set M a.s. and with a positive probability the set $M \setminus \binom{0}{0}$ at time T. If starting from the position $\binom{0}{0}$ we enter the set M a.s. and $intM$ with a positive probability at time T, we say that we have a *strict arbitrage opportunity at time T* (SA). Similarly we can characterize the absense of arbitrage. Denote by A_T the set of all portfolios which can be achieved at time T starting at time 0 from $\binom{0}{0}$. If there are no weak arbitrage at time T we have a *strict absense of arbitrage* (SAA), which means that

$$A_T \cap L^0(M, F_T) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}. \tag{3}$$

If there are no strict arbitrage at time T, which is the case when

$$A_T \cap L^0(M, F_T) \subset L^0(\partial M, F_T) \tag{4}$$

we have a *weak absense of arbitrage* (WAA).

Remark 1 *By the very definitions we immediately have:*

$$(SAA) \Rightarrow (WAA) \quad \textit{and} \quad (SA) \Rightarrow (WA).$$

Let $\binom{x_t}{y_t}$ be our market position at time t before possible transactions. It can be completely characterized by initial position $\binom{x_0}{y_0}$ and the pair of (F_t) adapted nonnegative processes l_t, m_t by the following formula

$$\begin{aligned} x_{t+1} &= x_t - \tau_1(l_t) + \tau_2(m_t) \\ y_{t+1} &= \tfrac{S_{t+1}}{S_t}\,(y_t + l_t - m_t)\,. \end{aligned} \tag{5}$$

Define for $t = 0, 1, \ldots, T$ sequences h_t, (H_t) of random transformations of R and R^2 respectively

$$h_t(y) := \frac{y}{S_t}, \qquad H_t\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} x \\ h_t(y) \end{pmatrix}. \tag{6}$$

Clearly

$$\begin{aligned} H_t M = \Big\{ \tbinom{x}{y} : y \geq \tfrac{-(\tau_1)^{-1}(x)}{S_t}, \quad &\textit{for} \quad x \geq 0 \\ \textit{and} \quad y \geq \tfrac{-(\tau_2)^{-1}(-x)}{S_t}, \quad &\textit{for} \quad x \leq 0 \Big\}, \end{aligned} \tag{7}$$

Let $\hat{y}_t = h_t(y_t)$, $\hat{l}_t = h_t(l_t)$, $\hat{m}_t = h_t(m_t)$ and $\hat{A}_T = H_T A_T$. We have

$$\hat{y}_{t+1} = \hat{y}_t + \hat{l}_t - \hat{m}_t, \tag{8}$$

where $\hat{y}_t$ is the number of assets in our portfolio before possible transactions at time t. Moreover from (8) we obtain

$$\hat{A}_T = \sum_{s=0}^{T} N_s, \tag{9}$$

where $N_s = -L^0(H_s M, F_s)$. Furthermore, (3) and (4) are equivalent respectively to

$$\hat{A}_T \cap L^0(H_T M, F_T) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \tag{10}$$

and

$$\hat{A}_T \cap L^0(H_T M, F_T) \subset L^0(\partial H_T M, F_T). \tag{11}$$

From the above consideration taking into account Remark 1 and (10), (11) we immediately have

Lemma 1 *(WA) with proportional transaction costs ($\tau_1(x) \equiv (1+\lambda)x$ and $\tau_2(x) \equiv (1-\mu)x$) implies (SA) with concave transaction costs. (WA) with concave transaction costs implies arbitrage without transaction costs. Conversely, absense of arbitrage without transaction costs implies (SAA) with concave transaction costs, while (WAA) with concave transaction costs implies (SAA) with proportional transaction costs.*

The following two examples explain the last lemma

Examples. Assume $T = 1$, and $S_1 = S_0(1+\xi)$, where the rate of return ξ is such that $P\{\xi = 0\} = p > 0$, $P\{\xi = r\} = 1-p > 0$ for $r > 0$. We clearly see that we have an arbitrage without transaction costs. In the case with concave transaction costs starting from $\binom{x}{(\tau_1)^{-1}(-x)}$, $x < 0$ we would like to reach (or cross) the position $\binom{x}{(\tau_2)^{-1}(-x)}$ which is impossible, so that even (WA) with concave trasaction costs does not hold. Consider now deterministic rate of return $\xi = r > 0$. When $\frac{1+r}{1+\lambda} > \frac{1}{1-\mu}$ we have (SA) with proportional transaction costs. If

$$\lim_{x\to\infty} \dot{\tau}_1(x) = 1 = \lim_{x\to\infty} \dot{\tau}_2(x)$$

and $r > \lambda$, we have (SA) with concave transaction costs (we choose x so small that $(1+r)(\tau_1)^{-1}(-x) > (\tau_2)^{-1}(-x)$). On the other hand, when $\lambda < r < \frac{\lambda+\mu}{1-\mu}$, we have (WAA) with proportional transaction costs.

We have the following

Proposition 1 *Under nonlinear concave transaction costs (WAA) is equivalent to (SAA).*

Proof. By Remark 1 we have $(SAA) \Rightarrow (WAA)$. Assume we have (WAA) which means in its equivalent form that (11) holds. By the form of the set M if $\binom{x}{y} \in M$, then $k\binom{x}{y} \in intM$ for $k > 1$. The above property is preserved under the transformation H_s, so that whenever $n_s \in N_s$ then $kn_s \in L^0(intH_sM \cup \binom{0}{0}, F_s)$ for $k > 1$. If (10) is not satisfied (i.e. (SAA) does not hold), then there is a $v \in A_T \cap L^0(H_TM, F_T)$ such that $v \not\equiv \binom{0}{0}$. Consequently $kv \in A_T \cap L^0(H_TM, F_T)$ and $kv \notin L^0(\partial H_TM, F_T)$ for $k > 1$, which contradicts (WAA).

□

Remark 2 *Notice that for proportional transaction costs (WAA) does not imply (SAA) (see theorem 1 below). Moreover arbitrage without transaction costs at time T implies the existence of a one step arbitrage at certain deterministic time $t \le T$ (see*[1] *theorem 3.3). This is no longer true in the case of concave transaction costs.*

Below we formulate a sufficient condition for arbitrage with concave transaction costs

Lemma 2 *If* $\lim_{x\to\infty}\dot{\tau}_1(x)=\lim_{x\to\infty}\dot{\tau}_2(x)$ *and there is* $\epsilon>0$ *and* $0\le t\le T$ *such that* $\frac{S_t}{S_{t+1}}\le 1-\epsilon$, *P a.e., then we have a strong arbitrage at time* T *with concave transaction costs.*
Proof. Let $K>0$ be such that

$$\frac{(\tau_1)^{-1}(K)}{(\tau_2)^{-1}(K)}>1-\epsilon.$$

Choosing the position $\binom{-K}{(\tau_1)^{-1}(K)}$ at time t we obtain

$$(\tau_1)^{-1}(K)>(\tau_2)^{-1}(K)\frac{S_t}{S_{t+1}},$$

from which the arbitrage follows.

□

Another charaterizations of arbitrage use an analog of "positive dual set" $(H_tM)^*$ to the set H_tM defined as follows

$$(H_tM)^*:=\left\{\begin{pmatrix}x\\y\end{pmatrix}:\tau_2(S_tx)\le y\le\tau_1(S_tx)\ \ and\ \ x\ge 0\right\}. \tag{12}$$

The following theorem characterizes arbitrage under proportional transaction costs

Theorem 1 *In the case of proportional transaction costs (SAA) is equivalent to the existence of an equivalent measure* Q *and an* F_t *adapted process* ρ_t *such that* $1-\mu<\rho_t<1+\lambda$, *for* $t=0,1\ldots,T$ *and* (ρ_tS_t) *is a* Q *martingale.*
If the market is finite (the prices S_t *of the asset admit a finite number of values only), in the case of proportional transaction costs, (WAA) is equivalent to the existence of an equivalent measure* Q *and an* F_t *adapted process* ρ_t *such that* $1-\mu\le\rho_t\le 1+\lambda$, *P a.e. for* $t=0,1\ldots,T$ *and* (ρ_tS_t) *is a* Q *martingale.*
Proof. Under (SAA) by theorem 2.1 of[3] there exists a bounded martingale $\left(\binom{z_1(t)}{z_2(t)}\right)$ with strictly positive components such that

$$\begin{pmatrix}z_1(t)\\z_2(t)\end{pmatrix}\in L^0(int(H_tM)^*,F_t).$$

Let $dQ=\frac{z_1(T)}{z_1(0)}dP$ and $\rho_t:=\frac{z_2(t)}{z_1(t)S_t}$. By Lemma 2 of[9] (ρ_tS_t) is a Q martingale. Moreover by the form of the set $(H_tM)^*$ we have that $1-\mu\le\rho_t\le 1+\lambda$, P a.e.. On the other hand the process $\left(\binom{E[\frac{dQ}{dP}|F_t]}{\rho_tS_tE[\frac{dQ}{dP}|F_t]}\right)$ is a P martingale taking values in $int(H_tM)^*$. Consequently for each $t\le T$ there is a martingale (Z_s^t) such that Z_s^t takes values in $(H_sM)^*$, while Z_t^t in $int(H_tM)^*$. Adapting the proof of Lemma 3.6 of[3] we obtain (10), which completes the proof of the first equivalence. The second equivalence follows from theorem 3.2 of[4]. Namely, under (WAA) there is a martingale with strictly positive components $(\binom{z_1(t)}{z_2(t)})$ (see lemma 3.1 of[4]) taking values respectively in the set $(H_tM)^*$. As above we construct the process (ρ_t) and the measure Q. Conversely, if there exists an equivalent measure Q and an F_t adapted process ρ_t such that $1-\mu \leq \rho_t \leq 1+\lambda$, P a.e. for $t = 0, 1 \ldots, T$ such that $(\rho_t S_t)$ is a Q martingale, the process $(\binom{1}{\rho_t S_t})$ is a P martingale taking values in $(H_tM)^*$ and by theorem 3.2 of[4] (WAA) follows.

□

Remark 3 *By the proof of theorem 1 one can notice that under proportional transaction costs, (SAA) is also equivalent to the existence of an equivalent measure Q such that $\frac{dQ}{dP}$ is bounded and an F_t adapted process ρ_t such that $1 - \mu < \rho_t < 1 + \lambda$, for $t = 0, 1 \ldots, T$ and $(\rho_t S_t)$ is a Q martingale.*

In an analogy to theorem 1 we show below a sufficient condition for absense of arbitrage under nonlinear concave transaction costs. We denote by $(\cdot, \cdot)$ and $|\cdot|$ respectively the scalar product and Eulidean norm in R^2.

Proposition 2 *Assume that the market is finite and for each $r > 0$ there is a supermartingale $Z(r)$ such that $|Z_s(r)| \geq r$, P a.e. for $s = 0, 1 \ldots, T$ and*

1) $(Z_s(r), \xi) \leq 0$ for any $\xi \in N_s$ such that $|\xi| \leq |Z_s(r)|$,

2) if $(Z_T(r), \xi) = 0$ for $\xi \in N_T$ such that $|\xi| \leq |Z_T(r)|$, then $\xi = 0$,

Then we have (SAA).

Proof. If (10) does not hold, there exist $\xi_s \in N_s$ and $0 \neq \xi'_T \in N_T$ such that $\sum_{s=0}^{T} \xi_s = -\xi'_T$. Therefore $\sum_{s=0}^{T-1} \xi_s = -\xi_T - \xi'_T \in N_T$. Choose

$$r \geq \max\left\{\max_s |\xi_s|, |\xi_T + \xi'_T|\right\}.$$

Then

$$(Z_T(r), \sum_{s=0}^{T-1} \xi_s) = -(Z_T(r), (\xi_T + \xi'_T)) \geq 0$$

and by the supermartingale property

$$E\left\{(Z_T(r), \sum_{s=0}^{T-1} \xi_s)\right\} \leq \sum_{s=0}^{T-1} E\left\{(Z_s(r), \xi_s)\right\} \leq 0.$$

Therefore $(Z_T(r), \xi_T + \xi'_T) = 0$ and by (2) $\xi_T + \xi'_T = 0$ from which it follows that $\xi_T = \xi'_T = 0$, a contradiction.

□

3 Set of hedging endowments

Let $C = \binom{C_1}{C_2}$ be a pair of integrable F_T measurable random variables called a *contingent claim*. If for an initial portfolio $\binom{x}{y}$ there is an investment strategy (l_t, m_t) such that $\binom{x_T}{y_T} \in \binom{C_1}{C_2} + M$, we say that starting from $\binom{x}{y}$ we *hedge* the contingent claim C. Our purpose now is to characterize the set of hedging endowments Γ described as follows

$$\Gamma = \left\{ v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} : C \in v + A_T \right\}. \tag{13}$$

Let $\mathcal{P}$ be a class of pairs $((\rho_t), Q)$ consisting of (F_t) adapted processes (ρ_t) taking values in the interval $[1-\mu, 1+\lambda]$ and an equivalent probability measure Q such that $\frac{dQ}{dP}$ is bounded and $(\rho_t S_t)$ is a Q martingale.
Following[2] and[3] (see also[9]) we have
Theorem 2 *Under (SAA) for proportional transaction costs we have*

$$\Gamma = \bigcap_{((\rho_t),Q)\in\mathcal{P}} \left\{ \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} : v_1 + \rho_0 v_2 \geq E^Q\{C_1\} + E^Q\{\rho_T C_2\} \right\}. \tag{14}$$

Proof. By theorem 4.2[3]

$$\Gamma = \bigcap_{(Z(t))\in\mathcal{Z}} \left\{ \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} : E\{z_1(T)C_1 + h_T(z_2(T))C_2\} \leq z_1(0)v_1 + h_0(z_2(0))v_2 \right\} \tag{15}$$

where $\mathcal{Z}$ is the set of bounded nonnegative martingales $(\binom{z_1(t)}{z_2(t)})$ such that

$$(1-\mu)z_1(t) \leq h_T(z_2(t)) \leq (1+\lambda)z_1(t), \quad P \ a.e.$$

for $t = 0, 1 \ldots, T$. Notice that we can restrict the set $\mathcal{Z}$ to strictly positive martingales. Let $(Z(t)) \in \mathcal{Z}$ be a strictly positive martingale. Define $dQ = \frac{z_1(T)}{z_1(0)} dP$ and $\rho_t := \frac{z_2(t)}{z_1(t)S_t}$. Clearly $((\rho_t), Q) \in \mathcal{P}$. Therefore the set defined by the right hand side of (14) is contained in the set Γ given in (15). On the other hand if $v \in \Gamma$ defined in (13) then

$$\begin{pmatrix} C_1 \\ \frac{C_2}{S_T} \end{pmatrix} = \begin{pmatrix} v_1 \\ \frac{v_2}{S_0} \end{pmatrix} + \sum_{s=0}^{T} \xi_s,$$

with $\xi_s \in N_s$. Let $((\rho_t), Q) \in \mathcal{P}$. Since

$$[1 \quad \rho_T S_T] \sum_{a=0}^{T} \xi_s = C_1 + \rho_T C_2 - v_1 - v_2 \frac{S_T}{S_0} \rho_T$$

is Q integrable, by Lemma 3.5 [3] we obtain

$$v_1 + v_2 E^Q \left\{ \frac{\rho_T S_T}{S_0} \right\} \geq E^Q \{C_1\} + E^Q \{\rho_T C_2\}.$$

Therefore v is a element of the set defined in (14). Thus the sets defined in (13), (14) and (15) coincide, which completes the proof.

□

Remark 4 *There is a problem to characterize the set Γ in the case of nonlinear concave transaction costs even when we have a finite financial market. We conjecture that*

$$\begin{aligned}\Gamma = \Big\{ v = \tbinom{v_1}{v_2} : E\{z_1(T)C_1 + h_T(z_2(T))C_2\} \leq \\ z_1(0)v_1 + h_0(z_2(0))v_2 \quad \textit{for each martingale} \\ Z(t) = \tbinom{z_1(t)}{z_2(t)} \in L^0((H_t M)^*, F_t) \ \textit{such that} \\ |Z(t)| \geq |\xi_t|, \quad \textit{where} \ \ \xi_t \in N_t \ \ \textit{for} \ \ t = 0, 1, \ldots, T \\ \textit{and} \ \ H_T C - H_0 v = \textstyle\sum_{s=0}^T \xi_s, \Big\}. \end{aligned} \tag{16}$$

It is not clear that for each $r > 0$ there is a martingale $Z(t)$ taking values in $(H_t M)^*$ such that $|Z(t)| \geq r$, which we impose for a sufficient condition of (SAA) in Proposition 2.

The remaining part of this section we devote to the study of the form of the set Γ for a so called *Cox Ross Rubinstein model* (CRR) with concave transaction costs, which is a finite market model with i.i.d. rates of return $\xi_t := \frac{S_{t+1} - S_t}{S_t}$ such that $0 < P\{\xi_t = a\} = 1 - P\{\xi_t = b\} < 1$, for $-1 < a < 0 < b$. We assume now that $C_1 = c_1(S_T)$ and $C_2 = S_T c_2(S_T)$ for Borel functions c_1 and c_2.

Define

$$\tau(z) := \begin{cases} \tau_1(z) \text{ if } \ z \geq 0 \\ -\tau_2(-z) \text{ if } \ z \leq 0 \end{cases} \tag{17}$$

which is the ammount of money we spend (or we get) to achieve the position z on the stock account. Denote by (η_t, θ_t) the market position consisting of the ammount of money η_t on the bank account and ammount of assets θ_t held in the portfolio at time t after possible transactions. If $\eta_T = c_1(S_T)$ and $\theta_T = c_2(S_T)$, we say we *replicate* the contingent claim C. In the case of CRR model we have replication when

$$\eta_{T-1} - c_1(S_{T-1}(1+e)) = \tau((c_2(S_{T-1}(1+e)) - \theta_{T-1})S_{T-1}(1+e)) \tag{18}$$

holds for $e = a, b$. Let

$$\Phi(c_1, c_2)(\theta, s) = c_1((1+a)s) - c_1((1+b)s) + \tau((c_2((1+a)s) - \theta)s((1+a)) - \tau((c_2((1+b)s) - \theta)s(1+b)). \tag{19}$$

One can notice (see proof of theorem 5 of[9]) that under

$$\frac{1+b}{1+a} > \frac{1+\lambda}{1-\mu} \tag{20}$$

the mapping $\theta \mapsto \Phi(c_1, c_2)(\theta, s)$ is strictly increasing with the range equal to R, so that $[\Phi(c_1, c_2)(\cdot, s)]^{-1}(0)$ is uniquely defined. Let

$$\Psi_2(c_1, c_2)(s) := [\Phi(c_1, c_2)(\cdot, s)]^{-1}(0),$$

$$\Psi_1(c_1, c_2)(s) := c_1((1+a)s) + \tau((c_2(1+a)s) - \Psi_2(c_1, c_2)(s)s(1+a))$$

and $\Psi(c_1, c_2)(s) := (\Psi_1(c_1, c_2)(s), \Psi_2(c_1, c_2)(s))$. Combining theorem 5 of[9] and[7] we obtain

Theorem 3 *Under (20) there is a unique replicating strategy (η_t, θ_t) for CRR model and is in the form of iterations of the operator Ψ as follows*

$$(\eta_t, \theta_t) = \Psi^{T-t}(c_1, c_2)(s_t) \tag{21}$$

If additionally

$$c_2((1+a)s) \le \Psi_2(c_1, c_2)(s) \le c_2((1+b)s) \tag{22}$$

for $s \ge 0$, then

$$\Psi_2(c_1, c_2)((1+a)s) \le \Psi_2^2(c_1, c_2)(s) \le \Psi_2(c_1, c_2)((1+b)s). \tag{23}$$

If $b > \lambda$ and $a + \mu > 0$ and (22) is satisfied then

$$\Gamma = \begin{pmatrix} \eta_0 \\ \theta_0 S_0 \end{pmatrix} + M. \tag{24}$$

Remark 5 *Under (22), when $S_{T-1} = (1+a)S_{T-2}$ we have*

$$\theta_{T-1} = \Psi_2(c_1, c_2)((1+a)S_{T-2}) \le \theta_{T-2}.$$

Moreover if $S_{T-1} = (1+b)S_{T-2}$ we have

$$\theta_{T-1} = \Psi_2(c_1, c_2)((1+b)S_{T-2}) \ge \theta_{T-2}.$$

By induction we easily obtain the following property of the replicating strategy: when the price of the asset increases we buy assets, while when it decreases we sell assets. The form (24) of the set Γ is true in a number of cases (even when (20) does not hold) for proportional transaction costs (see[8] or[9] for details).

4 Pricing

Given a characterization of the set of initial endowments Γ we are interested to know the *seller price* $p_s(C)$ at time $t = 0$ of the contingent claim C. It is the minimal ammount of money on our bank account from which we can enter the set Γ.

Theorem 4 *Under (SAA) for proportional transaction costs we have*

$$p_s(C) = \sup_{(\rho_t),Q)\in\mathcal{P}} E^Q\left\{C_1 + \rho_T C_2\right\}. \tag{25}$$

Proof. Notice first that by theorem 2, $p_s(C) \geq E^Q\left\{C_1 + \rho_T C_2\right\}$. To show the inverse inequality it remains to follow the proof of Theorem 4 in [9] which is based on the arguments from the proof of theorem 1 in [6].

□

In the remaining part of this section we shall consider a contingent claim C of the form $C_1 = c_1(S_T)$ and $C_2 = S_T c_2(S_T)$. Let Θ denote the class of (F_t) adapted processes (θ_t) such that there exists an adapted process (η_t) for which we have

$$\eta_{t+1} - \eta_t + \tau((\theta_{t+1} - \theta_t)S_{t+1}) = 0 \tag{26}$$

for $t = 0, \ldots, T-1$, and

$$c_1(S_T) - \eta_{T-1} + \tau((c_2(S_T) - \theta_{T-1})S_T) \leq 0. \tag{27}$$

Notice that (26) means that strategy (η_t, θ_t) is *selffinacing* i.e. there are not exogenous infusion or withdrawal of money, while (27) says that strategy (η_t, θ_t) hedges the contingent claim C.

In what follows we extend functions τ_1 and τ_2 letting $\tau_1(x) = -\tau_1(-x)$ and $\tau_2(x) = -\tau_2(-x)$ for $x \leq 0$.

Given $(\theta_t) \in \Theta$ define the class $\Lambda((\theta_t))$ consisting of the pairs $((\rho_t), Q)$, where Q is an equivalent measure such that $\frac{dQ}{dP}$ is bounded, and ρ_t is an (F_t) adapted function valued process for which

$$\min\left\{\tau_1(x), \tau_2(x)\right\} \leq \rho_t(x) \leq \max\left\{\tau_1(x), \tau_2(x)\right\} \tag{28}$$

for $x \in R$,

$$E^Q[\rho_t(\theta_t S_t) - \rho_t((\theta_t - \theta_{t-1})S_t)|F_{t-1}] \leq \rho_{t-1}(\theta_{t-1}S_{t-1}) \tag{29}$$

Q a.e. for $t = 1, \ldots, T-1$ and

$$E^Q[\rho_T(c_2(S_T)S_T) - \rho_T((c_2(S_T) - \theta_{T-1})S_T)|F_{T-1}] \leq \rho_{T-1}(\theta_{T-1}S_{T-1}) \tag{30}$$

Q a.e..

We have

Proposition 3 *If there is $(\theta_t) \in \Theta$ such that $\Lambda((\theta_t))$ is nonempty, then*

$$p_s(C) \geq \inf_{(\theta_t)\in\Theta} \sup_{(\rho_t),Q)\in\Lambda((\theta_t))} E^Q\left\{c_1(S_T) + \rho_T(c_2(S_T)S_T)\right\}. \tag{31}$$

Proof. Let (η_t, θ_t) be a strategy which hedges the contingent claim C and $((\rho_t), Q) \in \Lambda((\theta_t))$. By (27) and (30) taking into account that $\tau \geq \rho_T$ we obtain

$$\begin{aligned}
E^Q[c_1(S_T) + \rho_T(S_T c_2(S_T))|F_{T-1}] \leq \\
E^Q[\eta_{T-1} + \rho_T(S_T c_2(S_T)) - \tau((c_2(S_T) - \theta_{T-1})S_T)|F_{T-1}] \\
\leq \eta_{T-1} + E^Q[\rho_T(S_T c_2(S_T)) - \rho_T((c_2(S_T) - \theta_{T-1})S_T)|F_{T-1}] \\
\leq \eta_{T-1} + \rho_{T-1}(\theta_{T-1}S_{T-1})
\end{aligned} \tag{32}$$

Q a.e.. Moreover, the process $(\eta_t + \rho_t(\theta_t S_t))$ is for $t \leq T-1$ a Q supermartingale. In fact, by (26), inequality $\tau \geq \rho_t$ and (29) we obtain that

$$E^Q[\eta_t + \rho_t(\theta_t S_t)|F_{t-1}] \leq \eta_{t-1} + \rho_{t-1}(\theta_{t-1}S_{t-1}),$$

Q a.e. for $t \leq T-1$. Therefore by (32) and inequality $\tau \geq \rho_0$

$$\sup_{(\rho_t),Q)\in\Lambda((\theta_t))} E^Q\left\{c_1(S_T) + \rho_T(c_2(S_T)S_T)\right\} \leq \eta_0 + \tau(\theta_0 S_0)$$

from which (31) follows.

□

Notice that when ρ_t is a linear function with linear coefficient denoted with an ambiguity by ρ_t, such that $(\rho_t S_T)$ is a Q martingale, the inequalities (29) and (30) are in fact equalities and comparing (31) with (25) we obtain

Corollary 1 *In the case of proportional transaction costs under (SAA) we have an equality in the formula (31).*

To study the formula (31) for CRR model with nonlinear concave transaction costs we have to introduce a new notation. Let

$$\tau'(\theta) := \begin{cases} \tau_1 \text{ if } \theta \geq 0 \\ -\tau_2 \text{ if } \theta \leq 0 \end{cases} \tag{33}$$

In the case of CRR model we can consider $\Omega = \{a, b\}^T$ and $\omega = (\omega_1, \ldots, \omega_T)$. For an F_{t+1} adapted random variable ξ_{t+1} define F_t measurable random variables ξ_t^a and ξ_t^b as follows: we let for $\xi_t^a(\omega)$ and $\xi_t^b(\omega)$ the values of $\xi_{t+1}(\omega')$, where ω' differs from ω only at ω_{t+1} which is equal respectively to a or b.

Under (20) by theorem 3 there is a unique replicating strategy (η_t, θ_t) given by the formula (21). Let

$$q_t^b := \frac{n_t^b}{d_t^b}, \tag{34}$$

where

$$n_t^b := \tau((\theta_t^a - \theta_t)(1+a)S_t) - \tau'(\theta_t^a - \theta_t)(\theta_t^a S_t(1+a)) + \tau'(\theta_t - \theta_{t-1})(\theta_t S_t)$$

and

$$d_t^b := \tau((\theta_t^a - \theta_t)(1+a)S_t) - \tau'(\theta_t^a - \theta_t)(\theta_t^a S_t(1+a)) +$$
$$\tau'(\theta_t^b - \theta_t)(\theta_t^b S_t(1+b)) - \tau(\theta_t^b - \theta)(1+b)S_t).$$

We have

Lemma 3 *Under (20), (22), when $a + \mu > 0$ we have*

$$0 < q_t^b < 1. \tag{35}$$

Proof. Technical based on the properties of the functions τ_1 and τ_2 consists in consideration of all possible positions of $\theta_t^a \leq \theta_t \leq \theta_t^b$ (inequality holds by (23)) and 0.

□

We can now generalize Proposition 3.5 of[5]

Proposition 4 *Under (20), (22), assuming furthermore that $b > \lambda$ and $a + \mu > 0$ for the replicating strategy (η_t, θ_t) in CRR model we have*

$$E^Q \{c_1(S_T) + \rho_T(c_2(S_T)S_T)\} = \eta_0 + \tau(\theta_0 S_0) \tag{36}$$

where $Q\{\xi_t = a\} = 1 - q_t^b$, $\rho_T(x) = \tau'(\theta_T - \theta_{T-1})(x)$ and $\rho_t(x) = \tau'(\theta_t - \theta_{t-1})(x)$, for $0 \leq t \leq T-1$ with $\theta_{-1} := 0$.

Consequently we have the equality in (31).

Proof. Notice that the process $(\eta_t + \rho_t(\theta_t S_t))$ is a Q martingale. Since $\rho_0(\theta_0 S_0) = \tau(\theta_0 S_0)$ we therefore have (36). Combining (36) with (31) and (24) we obtain the equality in (31).

□

Remark 6 *We have showed equality in (31) for proportional transaction costs and for CCR model with nonlinear concave transaction costs. One can expect the equality (31) to be true as a general rule.*

References:

1. Dalang R. C., A. Morton and W. Willinger (1990): "Equivalent Martingale Measures and No-Arbitrage Securities Market Models", *Stochastics Stoch. Rep.* 29, 185-201,
2. Delbaen F., Yu.M. Kabanov and E.Valkeila (1998): "Hedging under transaction costs in currency markets: a discrete - time model", preprint
3. Kabanov Yu. M., M. Rásonyi, Ch. Stricker (2001): "No-arbitrage criteria for financial markets with efficient friction", preprint
4. Kabanov Yu. M., Ch. Stricker (2000): "The Harrison - Pliska arbitrage theorem under transaction costs", *J. Math. Econ.*, to appear
5. Koehl P.F., H. Pham and N. Touzi (1997a): "Hedging in discrete - time under transaction costs and continuous time limit", preprint
6. Koehl P.F., H. Pham and N. Touzi (1997b): "On super - replication in discrete time under transaction costs", preprint
7. Kocinski M., (2001): "Hedging in the CRR model under concave transaction costs", *Demonstratio Math.* 34, 497-512,7
8. Stettner L., (1997): "Option pricing in the CRR model with proportional transaction costs: a cone transformation approach", *Applicationes Mathematicae* 24,4, 475-514,
9. Stettner L., (2000): "Option Pricing in Discrete - Time Incomplete Market Models", *Math. Fin.* 10/2, 305-321,

The Necessity of No Asymptotic Arbitrage in APT Pricing

Xiaoai Lin[a], Xia Liu[†] and Yeneng Sun[bc d]

Abstract: A typical APT type formula states that the square of the deviations of the individual rates of return from a factor-pricing formula sum to a finite number. The assumption of no asymptotic arbitrage is, in general, sufficient but not necessary for such an APT type formula to hold. Under certain additional assumptions on the residual risks, the desired necessity result can be obtained for a market with a countably infinite or an uncountably infinite number of assets.

1 Introduction

In the arbitrage pricing theory (APT) of Ross, the asset returns in a financial market with a countably infinite number of assets are assumed to conform to a strict factor structure, where a given finite number of factors is used as a formalization of *systematic* risks in the market, and the residual components in the asset returns, called *unsystematic* or *idiosyncratic* risks, are orthogonal to the factors and to each other. If there is an absence of asymptotic arbitrage in the sense that a sequence of asymptotically costless and riskless finite portfolios does not asymptotically yield a positive return, then one can derive an APT pricing formula stating that the total square deviations of the individual rates of return from a factor-pricing formula is finite. Thus, the expected returns for all but a finite number of assets are approximately linearly related to its factor loadings (see [10] and [11]).

It is argued in [3] that the definition of a strict factor structure is sufficiently stringent and it is unlikely that any large asset market has a usefully small number of factors. The concept of an approximate factor structure is then introduced for a market with countably many assets. It is shown that the result of Ross can be generalized to the case of an approximate factor structure. Asset returns in [2] and [3] are viewed as elements of a suitable Hilbert space. The

[a]DBS Bank, 6 Shenton Way, Singapore 068809
[b]Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543
[c]Institute for Mathematical Sciences, National University of Singapore, 3 Prince George's Park, Singapore 118402
[d]Centre for Financial Engineering, National University of Singapore, 12 Prince George's Park, Singapore 118411

no asymptotic arbitrage condition is shown to imply the continuity of the cost functional on portfolio returns generated by the assets.

Based on the methods developed in[2],[3],[4] and[9], it is shown in[6] that an interesting APT theory can be developed in a setting with an *arbitrary* (countably or uncountably) infinite number of assets, and with or without correlations among their idiosyncratic risks. In comparison to the earlier APT result, exact factor pricing is obtained by moving from a countable to an uncountable domain of assets and by neglecting a countable rather than finite number of them. It is also noted in[6] that the no asymptotic arbitrage assumption, "while sufficient, is not necessary for the validity of the usual APT pricing formula." In other words, for a general infinite asset market with a factor structure, no asymptotic arbitrage is strictly stronger than claiming the validity of the usual APT type formula in the literature.

The exact law of large numbers for a continuum of independent (or uncorrelated) random variables is needed in a large literature in economic theory. Various versions of such exact law have been shown recently in [12], which allows complete elimination of the idiosyncratic risks in well-diversified portfolios. Based on that, a new concept of *no exact arbitrage* (as opposed to *no asymptotic arbitrage*) is introduced in[5] and[7]. This notion of no exact arbitrage is shown to be necessary *and* sufficient for an exact APT pricing formula to hold.

A natural question arises. Can the condition of no asymptotic arbitrage be necessary for an APT type formula to hold under some additional assumptions? Since the APT model is one of the main asset pricing models (see, for examples, [8]), an answer to this question should be of sufficient interest. The purpose of this paper is to prove a general "converse" result for the APT in terms of no asymptotic arbitrage. We show that if the idiosyncratic risks are not too small in some precise sense to be defined below, then the validity of an APT formula implies no asymptotic arbitrage. The rest of the paper is organized as follows. Some general results on the APT in[6] are collected in Section 2 for easy references. Section 3 contains some examples. The main necessity result is then presented in Section 4.

2 Some results on the APT pricing

We shall follow exactly the same notation as stated in Section 2 of[6]. Let the financial market consist of assets indexed by $i \in I$, where the index set I is an infinite set. Thus, we work with a countably infinite or an uncountably infinite number of assets. For each $i \in I$, let x_i be the random one-period rate of return to a dollar invested in the asset i, and as such, each asset has a unit cost. Each x_i is assumed to have a finite second moment; and its mean and variance are denoted by μ_i and $V(x_i)$. For simplicity, we also assume that there is a riskless asset s with a positive return ρ, and we let s be one of the random variables x_i. For any two random variables ϕ and ψ, let $cov(\phi, \psi)$ be the covariance between ϕ and ψ.

We shall assume that for each $i \in I$,

$$x_i = \mu_i + \beta_{i1} f_1 + \cdots + \beta_{iK} f_K + e_i,$$

where the factors $f_1, \cdots, f_K$ are orthogonal to each other, to all the e_i, and have zero mean, unit variance. The idiosyncratic disturbances e_i are assumed to have zero means, and we shall make additional assumptions on them as they are needed in the sequel.

A finite portfolio p assigns the share α_j to asset i_j, $1 \leq j \leq n$ for some $n \geq 1$. In this case, the cost $C(p)$ and mean $E(p)$ of the portfolio p are given respectively by $\sum_{j=1}^{n} \alpha_j$ and $\sum_{j=1}^{n} \alpha_j \mu_{i_j}$. The random return $R(p)$ of the portfolio p is given by $\sum_{j=1}^{n} \alpha_j x_{i_j}$, and its variance by $V(p)$.

We can also view a finite portfolio p as a function $\alpha : I \longrightarrow \mathbb{R}$ for which $\alpha(i) \neq 0$ for only finitely many i. In this case, one can simply write $C(p) = \sum_{i \in I} \alpha_i$, $E(p) = \sum_{i \in I} \alpha_i \mu_i$, $R(p) = \sum_{i \in I} \alpha_i x_i$ and

$$V(p) = V\left(\sum_{k=1}^{K} \sum_{i \in I} \alpha_i \beta_{ik} f_k + \sum_{i \in I} \alpha_i e_i\right) = \sum_{k=1}^{K} \left(\sum_{i \in I} \alpha_i \beta_{ik}\right)^2 + V(\sum_{i \in I} \alpha_i e_i).$$

Note that the finiteness assumption means that in all the sums $\sum_{i \in I}$ involving α_i, only finitely many terms are non-zero, and thus they are all well-defined.

We begin with a formal definition of the concept *no asymptotic arbitrage.*

Definition 1 *A financial market is said to have no asymptotic arbitrage if for any sequence of finite portfolios* $\{p_n\}_{n=1}^{\infty}$,

$$\lim_{n \to \infty} V(p_n) = 0 \text{ and } \lim_{n \to \infty} C(p_n) = 0 \Longrightarrow \lim_{n \to \infty} E(p_n) = 0.$$

The following definition generalizes the concept of an approximate factor structure in [3] for a market with countably many assets to the general case.

Definition 2 *A financial market is said to have an approximate factor structure if there exists a positive real number M such that for any finite subset I_F of I, the* **maximal** *eigenvalue of the covariance matrix $\sum_{I_F}$ of $\{e_i\}_{i \in I_F}$ is less than M.*

The result in following theorem was presented as Theorem 2 in [6]. It is a generalization of the main result in [3] to the setting of a financial market with a general index set.

Theorem 1 *If there is no asymptotic arbitrage in a market with an approximate factor structure, then there exist real numbers $\tau_1, \tau_2, \cdots, \tau_K$ such that*

$$\sum_{i \in I} (\mu_i - \rho - \tau_1 \beta_{i1} - \cdots - \tau_K \beta_{iK})^2 < \infty.$$

Since the sum $\sum_{i \in I} (\mu_i - \rho - \tau_1 \beta_{i1} - \cdots - \tau_K \beta_{iK})^2$ is finite, it is obvious that all except a countable number of these pricing errors are zero. Hence, the absence of asymptotic arbitrage in a financial market with uncountably many assets implies that all but a countable number of them can be priced-out exactly in terms of factors.

The following definition generalizes the concept of an exact factor structure as used in [1] and [10] to the general case.

Definition 3 *A financial market is said to have an exact factor structure if for each $i, j \in I$ with $i \neq j$, $cov(e_i, e_j) = 0$, and that there exists $0 \leq \zeta < \infty$ such that $V(e_i) \leq \zeta$ for all $i \in I$.*

If the market has an exact factor structure, then for any finite subset I_F of I, the maximal eigenvalue of the covariance matrix $\sum_{I_F}$ of $\{e_i\}_{i \in I_F}$ is certainly less than ζ. The following corollary, which generalizes the result of Ross in [10], is thus obvious. It is Theorem 1 in [6].

Corollary 1 *If there is no asymptotic arbitrage for a market with an exact factor structure, then there exist real numbers $\tau_1, \tau_2, \cdots, \tau_K$ such that*

$$\sum_{i \in I} (\mu_i - \rho - \tau_1 \beta_{i1} - \cdots - \tau_K \beta_{iK})^2 < \infty.$$

3 Examples

In this section, we construct two examples showing that the assumption of no asymptotic arbitrage is not necessary for the APT type formula to hold. We first consider a simple example without factors. This is Example 1 in [6].

Example 1 *Let $A = \{j_l : l = 0, 1, 2, \cdots\} \subseteq I$. Let e_i, $i \in I$ be mutually orthogonal random variables with mean zero. For each $l = 1, 2, \cdots$, the variance of e_{j_l} is $1/l^2$. Let the financial market consist of risky assets $\{x_i\}_{i \in I}$ where*

$$x_i = \begin{cases} \rho & \text{if } i = j_0 \\ \rho + 1/l + e_i & \text{if } i = j_l,\ l = 1, 2, \cdots, \\ \rho + e_i & \text{if } i \neq j_l,\ l = 0, 1, 2, \cdots. \end{cases}$$

Let μ_i be the mean of x_i. Since the market has no factors and $\sum_{i \in I} (\mu_i - \rho)^2 < \infty$, the APT type formula does hold. For each $n \geq 1$, take a portfolio

$p_n = (\alpha_{j_0}, \alpha_{j_1}, \ldots, \alpha_{j_n})$ *with* $\alpha_{j_0} = -\sum_{l=1}^{n} \alpha_{j_l}$ *and* $\alpha_{j_l} = l/n$ *for* $1 \le l \le n$, *where* α_{j_l} *is the share of asset* j_l *in the portfolio. Then, it is obvious that* $C(p_n) = 0$. *It is also easy to obtain that*

$$V(p_n) = \sum_{l=1}^{n} \alpha_{j_l}^2 V(e_{j_l}) = \sum_{l=1}^{n} 1/n^2 = 1/n.$$

Now

$$E(p_n) = \alpha_{j_0}\rho + \sum_{l=1}^{n} \alpha_{j_l}\mu_{j_l} = \sum_{l=1}^{n} \alpha_{j_l}/l = 1.$$

Since $\lim_{n\to\infty} V(p_n) = 0$, *and for all* $n \ge 1$, $C(p_n) = 0$ *and* $E(p_n) = 1$, *the market does permit asymptotic arbitrage.*

The following is an example with K factors.

Example 2 *As in the previous example, let* $A = \{j_l : l = 0, 1, 2, \cdots\} \subseteq I$. *Let* e_i, $i \in I$ *be mutually orthogonal random variables with mean zero. For each* $l = 1, 2, \cdots$, *the variance of* e_{j_l} *is* $1/l^4$. *Let the financial market consist of risky assets* $\{x_i\}_{i\in I}$ *where*

$$x_i = \begin{cases} \rho & \text{if } i = j_0 \\ \rho + \frac{1}{l^2} + \sum_{k=1}^{K} \tau_k \beta_{j_l k} + \sum_{k=1}^{K} \beta_{j_l k} f_k + e_i & \text{if } i = j_l,\ l = 1, 2, \cdots, \\ \rho + e_i & \text{if } i \ne j_l,\ l = 0, 1, 2, \cdots \end{cases}$$

where ρ, τ_k *are real numbers, and* $\beta_{j_l k} = 1/(l^3 k)$. *This market has a strict factor structure. Let* μ_i *be the mean of* x_i, *and then*

$$\mu_i = \begin{cases} \rho & \text{if } i = j_0 \\ \rho + \frac{1}{l^2} + \sum_{k=1}^{K} \tau_k \beta_{j_l k} & \text{if } i = j_l,\ l = 1, 2, \cdots, \\ \rho & \text{if } i \ne j_l,\ l = 0, 1, 2, \cdots \end{cases}$$

When $i \ne j_l$, $l = 1, 2, \cdots$, *let* $\beta_{ik} = 0$ *for all* $1 \le k \le K$. *Then* $\sum_{i\in I}(\mu_i - \rho - \sum_{k=1}^{K} \tau_k \beta_{ik})^2 = \sum_{l=1}^{\infty} \frac{1}{l^4} < \infty$, *i.e., the APT type formula does hold.*

For each $n \ge 1$, *take a portfolio* $p_n = (\alpha_{j_0}, \alpha_{j_1}, \ldots, \alpha_{j_n})$ *with* $\alpha_{j_0} = -\sum_{l=1}^{n} \alpha_{j_l}$ *and* $\alpha_{j_l} = l^3/n^2$ *for* $1 \le l \le n$, *where* α_{j_l} *is the share of asset* j_l *in the portfolio. Then, it is obvious that* $C(p_n) = 0$, *and*

$$R(p_n) = \sum_{l=1}^{n} \alpha_{j_l}\mu_{j_l} + \sum_{k=1}^{K}\left(\sum_{l=1}^{n} \alpha_{j_l}\beta_{j_l k}\right) f_k + \sum_{l=1}^{n} \alpha_{j_l} e_{j_l}.$$

It is also easy to obtain that

$$\begin{aligned}
V(p_n) &= \sum_{k=1}^{K}(\sum_{l=1}^{n}\frac{l^3}{n^2}\cdot\frac{1}{l^3 k})^2+\sum_{l=1}^{n}\frac{l^6}{n^4}\cdot\frac{1}{l^4} \\
&= \sum_{k=1}^{K}(\frac{1}{n^2}\sum_{l=1}^{n}\frac{1}{k})^2+\sum_{l=1}^{n}\frac{l^2}{n^4} \\
&\le \sum_{k=1}^{K}(\frac{1}{n^2}\sum_{l=1}^{n}1)^2+\frac{1}{n^4}\sum_{l=1}^{n}l^2 \\
&= \frac{K}{n^2}+\frac{1}{n^4}\cdot\frac{n(n+1)(2n+1)}{6},
\end{aligned} \tag{1}$$

and the mean is

$$\begin{aligned}
E(p_n) &= \sum_{l=0}^{n}\alpha_{j_l}\mu_{j_l} \\
&= -\sum_{l=1}^{n}\frac{l^3}{n^2}\rho+\sum_{l=1}^{n}\frac{l^3}{n^2}(\rho+\sum_{k=1}^{K}\tau_k\beta_{j_l k}+\frac{1}{l^2}) \\
&= \sum_{k=1}^{K}\tau_k\sum_{l=1}^{n}\frac{l^3}{n^2}\cdot\frac{1}{l^3 k}+\sum_{l=1}^{n}\frac{l^3}{n^2}\cdot\frac{1}{l^2} \\
&= \frac{1}{n^2}\sum_{k=1}^{K}\tau_k\sum_{l=1}^{n}\frac{1}{k}+\sum_{l=1}^{n}\frac{l}{n^2} \\
&= \frac{1}{n}\sum_{k=1}^{K}\frac{\tau_k}{k}+\frac{n(n+1)}{2n^2}.
\end{aligned} \tag{2}$$

Then $\lim_{n\to\infty}V(p_n)=0$, *and* $\lim_{n\to\infty}E(p_n)=1/2$. *But for all* $n\ge 1$, $C(p_n)=0$. *So the market does permit asymptotic arbitrage.*

4 The necessity of no asymptotic arbitrage

The above examples show that the absence of asymptotically arbitrage opportunities alone is not necessary for the validity of the APT pricing formula. However, as mentioned in the introduction, if the idiosyncratic risks are "not too small", then the validity of an APT formula will imply no asymptotic arbitrage. The following is a precise definition for the idiosyncratic risks to be "not too small".

Definition 4 *The idiosyncratic risks* e_i *in the financial market are said to be "not too small" if there exists a positive real number* m *such that for any* *finite subset I_F of I, the* **minimal** *eigenvalue of the covariance matrix $\sum_{I_F}$ of $\{e_i\}_{i\in I_F}$ is greater than m.*

Theorem 2 *Assume that the idiosyncratic risks e_i in the financial market are "not too small". If the APT type formula holds, i.e., there exist numbers $\tau_1, \ldots, \tau_K$ such that*

$$\sum_{i\in I}(\mu_i - \rho - \tau_1\beta_{i1} - \cdots - \tau_K\beta_{iK})^2 < \infty,$$

then the market has no asymptotic arbitrage.

Proof Consider a sequence of finite portfolios $\{p_n\}_{n=1}^{\infty}$. Then p_n defines a function $\alpha_n : I \longrightarrow \mathbb{R}$ for which $\alpha_n(i) \neq 0$ for only finitely many i.. For simplicity, denote $\alpha_n(i)$ by α_{ni}. Then the random return

$$R(p_n) = \sum_{i\in I}\alpha_{ni}\mu_i + \sum_{k=1}^{K}(\sum_{i\in I}^{n}\alpha_{ni}\beta_{ik})f_k + \sum_{l=1}^{n}\alpha_{ni}e_i.$$

Assume that $\lim_{n\to\infty} V(p_n) = 0$ and $\lim_{n\to\infty} C(p_n) = 0$. Thus

$$V(p_n) = \sum_{k=1}^{K}(\sum_{i\in I}\alpha_{ni}\beta_{ik})^2 + V(\sum_{i\in I}\alpha_{ni}e_i) \to 0. \tag{3}$$

Since all the terms in the above equation are non-negative, it is obvious that $V(\sum_{i\in I}\alpha_{ni}e_i) \to 0$. Let I_n be the finite set $\{i \in I, a_{ni} \neq 0\}$. Since the idiosyncratic risks e_i are assumed to be "not too small", we have

$$\begin{aligned} V(\textstyle\sum_{i\in I}\alpha_{ni}e_i) &= \sum_{i,j\in I_n} cov(a_{ni}e_i, a_{nj}e_j) \\ &= \sum_{i,j\in I_n} a_{ni}a_{nj}cov(e_i, e_j) \\ &\geq m\sum_{i\in I}\alpha_{ni}^2 \end{aligned} \tag{4}$$

Since $m > 0$, we have $\sum_{i\in I}\alpha_{ni}^2 \to 0$.

Note that $V(p_n) \to 0$ also implies that

$$\sum_{k=1}^{K}(\sum_{i\in I}\alpha_{ni}\beta_{ik})^2 \to 0.$$

Thus, $\sum_{i\in I}\alpha_{ni}\beta_{ik} \to 0$, for each $k = 1, \ldots, K$.

Let $\gamma_i \equiv \mu_i - \rho - \sum_{k=1}^{K} \tau_k \beta_{ik}$. Then the validity of the APT formula says that $\sum_{i \in I} \gamma_i^2 < \infty$.

Now we can check the mean return of the portfolio

$$\begin{aligned} E(p_n) &= \sum_{i \in I} \alpha_{ni} \mu_i \\ &= \sum_{i \in I} \alpha_{ni} (\gamma_i + \rho + \sum_{k=1}^{K} \tau_k \beta_{ik}) \\ &= \sum_{i \in I} \alpha_{ni} \gamma_i + \rho \sum_{i \in I} \alpha_{ni} + \sum_{i \in I} \alpha_{ni} \sum_{k=1}^{K} \tau_k \beta_{ik} \\ &\leq (\sum_{i \in I} \alpha_{ni}^2 \sum_{i \in I} \gamma_i^2)^{1/2} + \rho C(P_n) + \sum_{k=1}^{K} \tau_k \sum_{i \in I} \alpha_{ni} \beta_{ik}. \end{aligned} \tag{5}$$

The last inequality follows from the Cauchy-Schwartz inequality. Since $\sum_{i \in I} \alpha_{ni}^2 \to 0$, $\sum_{i \in I} \gamma_i^2 < \infty$, $C(p_n) \to 0$, and $\sum_{i \in I} \alpha_{ni} \beta_{ik} \to 0$, we obtain that $E(p_n) \to 0$ as $n \to \infty$. Therefore, the market has no asymptotic arbitrage.

The following corollary gives a "converse APT result" for a market with an exact factor structure. The proof follows from the fact that for any finite subset I_F of I, the minimal eigenvalue of the covariance matrix $\sum_{I_F}$ of $\{e_i\}_{i \in I_F}$ is greater than ε.

Corollary 2 *Assume that the financial market has an exact factor structure and there exists a $\varepsilon > 0$ such that $V(e_i) > \varepsilon$ for all $i \in I$. If the APT type formula holds, then the market has no asymptotic arbitrage.*

Remark 1 *For the case that the index set I is the countable set of positive integers, the results in Theorem 2 and Corollary 2 respectively show that under suitable conditions on the idiosyncratic risks, the no asymptotic arbitrage condition in Ross[10] and Chamberlain and Rothschild[3] are not only sufficient, but also necessary for the APT type formula to hold in their settings. This result is new even in this countable case.*

1. A. Basilevsky, "Statistical Factor Analysis and Related Methods," John Wiley, New York, 1994.
2. G. Chamberlain, Funds, factors and diversification in arbitrage pricing models, *Econometrica* **51** (1983), 1305-1323.
3. G. Chamberlain and M. Rothschild, Arbitrage, factor structure, and mean-variance analysis on large markets, *Econometrica* **51** (1983), 1281-1304.

4. G. Huberman, A simple approach to arbitrage pricing theory, *Journal of Economic Theory* **28** (1982), 183-191.
5. M. Ali Khan and Y. N. Sun, The capital-asset-pricing model and arbitrage pricing theory: a unification, *Proceedings of the National Academy of Sciences of the United States of America* **94** (1997), 4229-4232.
6. M. Ali Khan and Y. N. Sun, Asymptotic arbitrage and the APT with or without measure-theoretic structures, *Journal of Economic Theory* (2001), in press.
7. M. Ali Khan and Y. N. Sun, Exact Arbitrage, well-diversified portfolios and asset pricing in large markets, preprint.
8. R. C. Radcliffe, "Investment: Concepts, Analysis, Strategy," 5th ed., Addison-Wesley, Reading, Mass., 1997.
9. H. Reisman, A general approach to the arbitrage pricing theory (APT), *Econometrica* **56** (1988), 473-476.
10. S. A. Ross, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* **13** (1976), 341-360.
11. S. A. Ross, Return, risk and arbitrage, *in* "Risk and Return in Finance" (I. Friend and J. L. Bicksler, Eds.), Balinger Publishing Co., Cambridge, Mass., 1977.
12. Y. N. Sun, A theory of hyperfinite processes: the complete removal of individual uncertainty via exact LLN, *Journal of Mathematical Economics* **29** (1998), 419-503.

Financial Mean-Variance Problems and Stochastic LQ Problems: Linear Stochastic Hamilton Systems and Backward Stochastic Riccati Equations

Shanjian Tang
Department of Mathematics
and the Laboratory of Mathematics for Nonlinear Sciences,
Fudan University, Shanghai 200433, China
E-mail: sjtangk@online.sh.cn

Financial mean-variance problems, including the mean-variance hedging, the mean-variance portfolio selection and the variance-optimal martingale measure, have obvious importance in modern finance. They are one-dimensional singular non-homogeneous stochastic linear-quadratic control problems (LQ), and can be solved in terms of the associated Riccati equation. However, the solution of the Riccati equation associated with the general stochastic LQ problem with random coefficients presents a new problem, which in fact has been open since Bismut (1978). Recently, the general one-dimensional case—which is the right case in the financial mean-variance problems—has been solved by Kohlmann and the author. More recently, the general regular case has been solved by the author with the theory of stochastic Hamilton system. In this article, the extension of the latter work is described to the singular case, which therefore provides an alternative approach to financial mean-variance problems with the theory of stochastic Hamilton system.

1 Introduction

Financial mean-variance problems—including the mean-variance hedging, the mean-variance portfolio selection and the variance-optimal martingale measure—are one-dimensional singular non-homogeneous stochastic linear-quadratic control problems. The classical stochastic LQ control theory suggests a general scheme for the complete solution of a stochastic LQ problem in terms of the associated Riccati equation. Bismut[1,2] and Peng[24] have provided two different heuristic ways to the general scheme, starting respectively from the stochastic maximum principle (which leads to a linear stochastic Hamilton system) and from the dynamic programming principle (which leads to a backward stochastic Hamilton-Jacobi-Bellman equation). However, the justification of each heuristic way requires that the associated Riccati equation has a solution.

Unfortunately, a rigorous theory is still lacking in the literature for the solution of the Riccati equation associated with a general stochastic LQ problem. In fact, this problem has been open since Bismut[1]. The Riccati equation associated with a general stochastic LQ problem is a symmetric matrix valued nonlinear backward stochastic differential equation (BSDE). The drift is a quadratic form of the martingale term (the second unknown variable) and also involves the inverse of the first unknown variable. This kind of structure goes far beyond the consideration of Pardoux and Peng's fundamental existence and uniqueness result[23] on BSDEs. Due to the appearance of the inverse of the first unknown variable, its one-dimensional version also goes beyond a direct consideration of Kobylanski's result[14] on one-dimensional BSDEs with quadratic growth in the second unknown variable. Last year, some breakthroughs have been made on the difficulty and two methods are developed—one is the approximation technique with uniformly Lipschitz drifts, and the other is the inverse transform technique. See Kohlmann and Tang[18] for more details. However, both methods have limitations and they can only be applied to some special cases: the approximation technique is difficult to be applied to the multi-dimensional case, and the transformation technique essentially requires among others a matching condition between the dimensions of the state, the control and the underlying Brownian motion.

Recently, the author[32] has developed a constructive method in terms of the solution of the associated linear Hamilton system (HS) so as to get around the above said difficulty. This method turns out to be successful to the proof of the existence and uniqueness result for the Riccati equation associated with a *regular* stochastic LQ problem. In this article, it will be shown that this method also permits an easy extension to the *singular* case. Therefore, a Hamilton system theoretic approach to the financial mean-variance problems is established. It is worth pointing out that this new approach also brings some new insight on the financially concerned fact that the variance-optimal measure is an equivalent martingale measure.

Now let us introduce some notations. Let $W := (W^1, \cdots, W^d)$ be a d-dimensional standard Brownian motion defined on some probability space $(\Omega, \mathcal{F}, P)$. Denote by $\{\mathcal{F}_t, 0 \leq t \leq T\}$ the augmented natural filtration of the standard Brownian motion $\{W_t, 0 \leq t \leq T\}$. $\mathcal{L}^2_{\mathcal{F}}(0, T; R^m)$, denotes the set of all R^m-valued square integrable $\{\mathcal{F}_t, 0 \leq t \leq T\}$-adapted processes. $L^2(\Omega, \mathcal{F}_T, P; R^n)$ is the set of all R^n-valued square integrable random variables defined on $(\Omega, \mathcal{F}_T, P)$. $L^\infty(\Omega, \mathcal{F}_T, P; R^n)$ is the subspace of $L^2(\Omega, \mathcal{F}_T, P; R^n)$ which consists of all those essentially bounded random variables. $\mathcal{S}^n$ denotes the set of all $n \times n$ symmetric matrices. The prime denotes the transpose of a vector or a matrix.

The rest of this article is organized as follows. Section 2 contains three mean-variance problems and some relevant comments. In Section 3, a general stochastic LQ problem is formulated and the associated Riccati equation is introduced. In section 4, some works on Riccati equations and recent progresses are sketched and a long-standing problem is introduced. In Section 5, the interrelationship between the stochastic LQ problem, the Riccati equation, and the Hamilton system is stated in a more general context than the author's previous paper [32]. Finally in section 6, some conclusions are given.

2 Financial Mean-variance Problems

Mean-variance models have a feature of simplicity. They capture the two fundamental concepts of return and risk in a financial market. It is not surprised that they have been receiving much considerations from the beginning of modern financial theory and practice.

In the following, three examples are given to illustrate the applications of the mean-variance models in finance: the first two coming from the considerations of pricing a contingent claim, and the third one coming from the investment considerations.

Suppose that there are in the market one bond and m stocks whose price dynamics are governed by the following:

$$\begin{cases} dS_t^0 = r_t S_t^0 \, dt, \\ dS_t = \mathrm{diag}(S_t)(\mu_t \, dt + \sigma_t \, dW_t). \end{cases} \tag{1}$$

Here r, μ, σ are adapted, bounded processes, representing respectively the interest rate, the appreciation rates vector, and the volatility matrix in the market. Moreover, assume that

$$\sigma_t \sigma_t' \geq \alpha I_{d\times d}, \quad \text{for some constant } \alpha > 0. \tag{2}$$

The risk premium process is

$$\lambda_t := \sigma'(\sigma\sigma')^{-1}\widetilde{\mu}_t \quad \text{where } \widetilde{\mu} := \mu - (r, \cdots, r)'. \tag{3}$$

Denote by $\pi(t)$ the vector whose i-th component is the amount of money invested in the i-th stock. Then the wealth equation is given by

$$\begin{cases} dx_t = [r_t x_t + \langle \widetilde{\mu}_t, \pi_t \rangle] \, dt + \pi_t' \sigma_t \, dW_t, \\ x_0 = h, \quad \pi \in \mathcal{L}^2_{\mathcal{F}}(0, T; R^m). \end{cases} \tag{4}$$

The solution corresponding the risky portfolio process π and the initial data $(0, h)$ is denoted by $x^{0,h;\pi}$.

Problem 1. (the mean-variance hedging) $\forall \xi \in L^2(\Omega, \mathcal{F}_T, P; R)$, *consider*

$$\min_{\pi \in \mathcal{L}^2_{\mathcal{F}}(0,T;R^m)} E|x_T^{0,h;\pi} - \xi|^2. \tag{5}$$

This problem was considered, among others by Duffie and Richardson [9], Schweizer [28,29,30], Gourieroux, Laurent and Pham [11], and Laurent and Pham [21].

Problem 2. (the variance-optimal martingale measure)

$$\min_{\theta\in\mathcal{L}^2_{\mathcal{F}}(0,T;R^m)} E|\mathcal{X}_T^{0,1;\theta}|^2 \tag{6}$$

where $\mathcal{X}^{0,x;\theta}$ solves the SDE

$$\begin{cases} d\mathcal{X} = \mathcal{X}[-r\,dt - \langle\lambda,\, dW\rangle] + \langle[I - \sigma'(\sigma\sigma')^{-1}\sigma]\theta,\, dW\rangle, \\ \mathcal{X}_0 = 1, \quad \theta\in\mathcal{L}^2_{\mathcal{F}}(0,T;R^m), \end{cases}$$

specifying the Radon-Nikodym derivative of a "signed" martingale measure—which is not necessarily positive.

This problem is the dual problem of the homogeneous mean-variance hedging problem, and it was considered by Schweizer [30], Delbaen and Schachermayer [8], Gourieroux, Laurent and Pham [11], and Laurent and Pham [21].

Problem 3. (the mean-variance portfolio selection) *Minimize $E|x_T^{0,h;\pi} - Ex_T^{0,h;\pi}|^2$ over $\pi\in\mathcal{L}^2_{\mathcal{F}}(0,T;R^m)$ under the constraint $Ex_T^{0,h;\pi}=\eta$ for a previously given $\eta > h\exp\left(\int_0^T r_t\,dt\right)$ where h is the initial wealth of an individual and is assumed to be positive.*

The above three problems are all stochastic LQ problems. The concerned financial issues include

(1) the variance-optimal portfolio,

(2) the value function, used to determine the price corresponding ξ,

(3) the positivity of the corresponding optimal wealth when $\xi=0$ and $h=1$, which is used as a hedging numeraire (see Gourieroux, Laurent and Pham [11]),

(4) the positivity of the Radon-Nikodym derivative of the variance-optimal martingale measure, which implies that the variance-optimal one among the signed martingale measures is an equivalent martingale measure! see Shweizer [30], and Delbaen and Schachermayer [8].

The last two financially concerned issues are mathematically the same one: for a one-dimensional homogeneous stochastic LQ problem, the optimal state process always keeps the sign of its initial values; more generally, for a multi-dimensional homogeneous stochastic LQ problem, the optimal state process is almost surely flows of homomorphism. This assertion comes from the fact that for a homogeneous stochastic LQ problem, the optimal control is a linear feedback of the state, which will be stated in Section 5. Therefore, we can give a new systematic proof to the above assertions (3) and (4), which is more general than those of Gourieroux, Laurent and Pham[11], Shweizer[30], and Delbaen and Schachermayer[8] in the sense that it applies to the multi-dimensional case.

3 The General Stochastic LQ Problem

The general stochastic LQ problem, denoted hereafter by $\mathcal{P}_{0,h}$, is the following minimization

$$\inf_{u\in\mathcal{L}^2_{\mathcal{F}}(0,T;R^m)} \widetilde{J}(u,h) \tag{7}$$

where

$$\begin{aligned}\widetilde{J}(u,h) := {} & \frac{1}{2}E\langle Mx_T^{0,h;u}, x_T^{0,h;u}\rangle + E\langle m_T, x_T^{0,h;u}\rangle \\ & + \frac{1}{2}E\int_0^T \left(\langle Q_t x_t^{0,h;u}, x_t^{0,h;u}\rangle + \langle N_t u_t, u_t\rangle\right) dt \\ & + E\int_0^T \left(\langle q_t, x_t^{0,h;u}\rangle + \langle \beta_t, u_t\rangle\right) dt\end{aligned}$$

and $x^{0,h;u}$ solves the SDE:

$$\begin{cases} dx = (Ax + Bu + f)\,dt + \sum_{i=1}^{d}(C^i x + D^i u + g^i)\,dW^i, \\ x_0 = h \in R^n, \quad u(t) \in R^m. \end{cases} \tag{8}$$

Here the new notations appearing in the system and the cost functional will be specified in Section 5.

Starting from the stochastic maximum principle (MP) and from the dynamic programming respectively, Bismut[1,2] and Peng[24] derived in an heuristic way the following backward stochastic Riccati differential equation (BSRDE):

$$\begin{cases} dK_t = -G(t, K_t, L_t)\,dt + \sum_{i=1}^{d} L_t^i\,dW_t^i, \\ K_T = M, \quad 0 \le t < T, \end{cases} \tag{9}$$

where

$$
\begin{aligned}
G(t,K,L) := {} & A_t'K + KA_t + \sum_{i=1}^{d}(C_t^i)'KC_t^i + Q_t \\
& + \sum_{i=1}^{d}[(C_t^i)'L^i + L^iC_t^i] + F(t,K,L), \\
F(t,K,L) := {} & -[KB_t + \sum_{i=1}^{d}(C_t^i)'KD_t^i + \sum_{i=1}^{d} L^iD_t^i][N_t + \sum_{i=1}^{d}(D_t^i)'KD_t^i]^{-1} \\
& \times[KB_t + \sum_{i=1}^{d}(C_t^i)'KD_t^i + \sum_{i=1}^{d} L^iD_t^i]', \\
& \forall (t,K,L) \in [0,T] \times \mathcal{S}^n \times (\mathcal{S}^n)^d.
\end{aligned}
$$

However, the rigorous solution of this equation in the general case had been open for a long period of over two decades since Bismut [1,2].

4 Solution of Riccati Equation: a Long-Standing Problem

4.1 Historical Remarks

In 1978, Bismut [2] commented on page 220 of the Springer yellow book LNM 649 that: "*Nous ne pourrons pas démontrer l'existence de solution pour l'équation (2.49) dans le cas général.*" On page 238, he pointed out that the essential difficulty for solution of the general BSRDE lies in the integrand of the martingale term which appears in the drift in a quadratic way. The difficulty asises if the coefficients are correlated to the *control-dependent* system noise! That is, the difficulty is marked by the simultaneous occurrence of the random change of the coefficients and its correlation with the control-dependent noise. Both features find strong motivation in finance. In particular, the factor models (see Heston [12], Hull and White [13], and Stein and Stein [31], for example) in financial theory support the random nature of the coefficients.

Two decades later in 1998, Peng [25] formally included the solution of the general Riccati equation (9) in his list of open problems on BSDEs.

Some comments, similar to Bismut's can also be found in Chen, Li, and Zhou [3], and Chen and Yong [4].

In the literature, Wonham[35] discussed the case of deterministic coefficients where the drift $F(t,K,L)$ does not contain L (it can also be taken as being zero matrix) and the Riccati equation is an ordinary differential equation. Bismut [1,2] and Peng [24] considered with different methods a special case of random coefficients where $F(t,K,L)$ linearly depends on L—thus the above-described difficulty does not exist there.

In the last few years, there are several related works on indefinite stochastic LQ problems, among which are cited Chen, Li and Zhou[3], Chen and Zhou[7], and Chen and Yong[4,5,6]. The book Yong and Zhou[34] contains one chapter's description in this respect. However, these works mainly concentrate on the indefinite feature of the Riccati equation(9) and unfortunately do not concentrate on dealing with the difficulty initially described by Bismut.

4.2 Some Recent developments

Last year (in the year 2000), on the difficulty initially described by Bismut, some breakthroughs[16,17,20] were made, a detailed account of which can be found in Kohlmann and Tang[18].

Recently, a complete solution to the above Bismut-Peng's problem (the Riccati equation (9) in the regular case) is provided by the author[32] by developing a Hamilton system theoretic approach. Actually, the formulas for the unique solution of Riccati equation (9) are given in terms of a set of "base" solutions of the associated HS.

In the following section, the adaptations to a more general case than the author's recent paper[32] are presented so as to cover the non-homogeneous case and the singular case. Due to the limitation of space, only the main results and comments are given here. The adaptations to the jump case will be presented elsewhere.

5 The Regular and Singular Stochastic LQ Problems: a Hamilton System Theoretic Method

In this section, we shall make the following basic assumptions:

(A1) The coefficients $A, B, C = (C^1, \cdots, C^d), D = (D^1, \cdots, D^d), Q, N$ are adapted bounded matrix-valued processes of appropriate dimensions, and $M \in L^\infty(\Omega, \mathcal{F}_T, P; R^{n\times n})$.

(A2) $M \geq 0, Q_t \geq 0, N_t \geq \alpha I_{m\times m}, \alpha > 0$.

(A3) $f, g = (g^1, \cdots, g^d) \in (\mathcal{L}^2_{\mathcal{F}}(0,T;R^n))^d, \beta \in \mathcal{L}^2_{\mathcal{F}}(0,T;R^m), m_T \in L^2(\Omega, \mathcal{F}_T, P; R^n)$.

And assumption (A2) can always be replaced with the following one

(A2)' $M \geq \alpha I_{n\times n}, Q_t \geq 0, N_t \geq 0, \sum_{i=1}^{d}(D_t^i)' D_t^i \geq \alpha I_{m\times m}, \alpha > 0$.

Under the above assumptions, the cost is then a coercive quadratic functional of the control in a Hilbert space, and the existence and uniqueness of the optimal control is a simple functional fact.

Theorem 5.1 *Let assumptions (A1), (A2) (or (A2)'), and (A3) be satisfied. Then there is a unique optimal control.*

The proof is referred to Bismut [1] and J. L. Lions [22].

Introduce the Hamiltonian

$$\widetilde{H}(t,x,u;y,z) := \langle y, A_t x + B_t u + f_t\rangle + \sum_{i=1}^{d} \langle z^i, C_t^i x + D_t^i u + g_t^i\rangle$$

$$+\frac{1}{2}\langle Q_t x, x\rangle + \frac{1}{2}\langle N_t u, u\rangle + \langle \beta_t, u\rangle + \langle q_t, x\rangle,$$

$$\forall (x,y,u) \in R^n \times R^n \times R^m,\ z = (z^1, \cdots, z^d) \in (R^n)^d.$$

Consider the associated linear stochastic Hamilton system:

$$\begin{cases} dx_t = \partial_y \widetilde{H}(t, x_t, u_t; y_t, z_t)\, dt + \sum_{i=1}^{d} \partial_{z^i} \widetilde{H}(t, x_t, u_t; y_t, z_t)\, dW_t^i \\ dy_t = -\partial_x \widetilde{H}(t,x,u;y,z)\, dt + \sum_{i=1}^{d} z_t^i dW_t^i, \\ x_0 = h, \quad y_T = M x_T + m_T, \\ \ 0 = \partial_u \widetilde{H}(t, x_t, u_t; y_t, z_t). \end{cases} \tag{10}$$

Its solution will be denoted hereafter by $(\widetilde{x}, \widetilde{y}, \widetilde{z}, \widetilde{u})$ or $(\widetilde{x}(h), \widetilde{y}(h), \widetilde{z}(h), \widetilde{u}(h))$. Obviously, the solution will also depend on the parameter $\{f; g^1, \cdots, g^d; \beta, m_T)$. Denote by $(x(h), y(h), z(h), u(h))$ the solution as $f = g^1 = \cdots = g^d = 0, \beta_t \equiv 0, m_T = 0$.

The following theorem contains the equivalence between the stochastic LQ problem $\mathcal{P}_{0,h}$ and the HS.

Theorem 5.2 *Let assumptions (A1), (A2) (or (A2)'), and (A3) be satisfied. Then, (i) if Hamilton system (10) has a solution $(\widetilde{x}(h), \widetilde{y}(h), \widetilde{z}(h), \widetilde{u}(h))$, then $\widetilde{u}(h)$ is the optimal control of problem $\mathcal{P}_{0,h}$; (ii) inversely, if $\widetilde{u}(h)$ is the optimal control of problem $\mathcal{P}_{0,h}$, then there is a triple $(\widetilde{x}(h), \widetilde{y}(h), \widetilde{z}(h))$ such that $(\widetilde{x}(h), \widetilde{y}(h), \widetilde{z}(h), \widetilde{u}(h))$ is a solution to Hamilton system (10). Moreover, Hamilton system (10) has a unique solution $(\widetilde{x}(h), \widetilde{y}(h), \widetilde{z}(h), \widetilde{u}(h))$, satisfying*

$$E \max_{0\le t\le T} |x_t(h)|^2 + E \max_{0\le t\le T} |y_t(h)|^2 + E \int_0^T |z_t(h)|^2\, dt$$

$$\leq \varepsilon E\left(|h|^2 + |m_T|^2\right) + \varepsilon E \int_0^T \left(|f_t|^2 + |g_t|^2 + |\beta_t|^2 + |q_t|^2\right) dt.$$

Here ε is a positive constant.

The proof consists in the application of the stochastic MP, a priori estimates of BSDEs, and the "Energy Equality".

Let e_i denote the i-th column of the $n \times n$ unit matrix $I_{n\times n}$ for $i = 1, \ldots, n$. Define

$$\begin{aligned} X &:= (x(e_1), \cdots, x(e_n)), \quad Y := (y(e_1), \cdots, y(e_n)), \\ Z^i &:= (z^i(e_1), \cdots, z^i(e_n)), \quad i = 1, \ldots, d, \\ Z &:= (z^1, \cdots, z^d), \quad U := (u(e_1), \cdots, u(e_n)). \end{aligned} \tag{11}$$

Then, (X, Y, Z, U) is the fundamental solution matrix of HS (10).

The following theorem contains the existence and uniqueness result for the solution of a general Riccati equation (9), and thus solves the long standing problem mentioned in the last section. Moreover, it provides the formulas in terms of the solutions of the associated HS. This theorem in the regular case constitutes the main result of the author's recent work [32].

Theorem 5.3. *Let assumptions (A1) and (A2) (or (A2)') be satisfied. Then, X_t a.s. has an inverse, and the stochastic Riccati equation (9) has the unique adapted solution (K, L) with*

$$\begin{aligned} K_t &= Y_t X_t^{-1}, \quad L := (L^1, \cdots, L^d), \\ L_t^i &= Z_t^i X_t^{-1} - Y_t X_t^{-1}(C_t^i + D_t^i U_t X_t^{-1}), \quad i = 1, \ldots, d. \end{aligned} \tag{12}$$

Moreover, $K_t \geq 0, \quad E \int_0^T |L_t|^2\, dt < \infty.$

It is noted that, the fact that X_t^{-1} exists for every $t \in [0, T]$ implies that $\{x_t(h) : 0 \leq t \leq T, h \in R^n\}$ is a flow of homomorphism.

Theorem 5.4. *Let assumptions (A1), (A2) (or (A2)'), and (A3) be satisfied. Let (K, L) be the unique adapted solution of Riccati equation (9), and $(\widetilde{x}, \widetilde{y}, \widetilde{z}, \widetilde{u})$ be the unique solution of Hamilton system (10). Define*

$$\begin{aligned} \psi_t &= \widetilde{y}_t - K_t \widetilde{x}_t, \\ \phi_t^i &= \widetilde{z}_t^i - [L_t^i \widetilde{x}_t + K_t(C_t^i \widetilde{x}_t + D_t^i \widetilde{u}_t + g_t^i)], \quad i = 1, \ldots, d. \end{aligned}$$

Then, (ψ, ϕ) with $\phi := (\phi^1, \cdots, \phi^d)$ is the unique adapted solution of the following BSDE:

$$\begin{cases} d\psi_t = -\{\widehat{A}_t'\psi_t + \sum_{i=1}^{d}(\widehat{C}_t^i)'(\phi_t^i + K_t g_t^i) + q_t + K_t f_t \\ \qquad + \sum_{i=1}^{d} L_t^i g_t^i + \Gamma(t, K_t, L_t)'\beta_t\}\, dt + \sum_{i=1}^{d} \phi_t^i \, dW_t^i, \\ \psi_T = m_T \end{cases} \tag{13}$$

with $E\int_0^T |\psi|^2\, ds < \infty$ and $\int_0^T |\phi|^2\, ds < \infty$ a.s., where

$$\begin{aligned} \Gamma(t, K, L) &:= -[N_t + \sum_{i=1}^{d}(D_t^i)'KD_t^i]^{-1}[KB_t + \sum_{i=1}^{d}(C_t^i)'KD_t^i + \sum_{i=1}^{d} L^i D_t^i]', \\ &\qquad \forall (t, K, L) \in [0, T] \times \mathcal{S}^n \times (\mathcal{S}^n)^d, \\ \widehat{A}_t &:= A_t + B_t\Gamma(t, K_t, L_t), \\ \widehat{C}_t^i &:= C_t^i + D_t^i\Gamma(t, K_t, L_t), \quad i = 1, \ldots, d. \end{aligned}$$

The following theorem provides the formulas of the solutions of HS (10) and the optimal control in terms of Riccati equations (9) and (13).

Theorem 5.5. *Let assumptions (A1), (A2) (or (A2)'), and (A3) be satisfied. If (K, L) is the unique adapted solution of Riccati equation (9) and (ψ, ϕ) with $\phi := (\phi^1, \cdots, \phi^d)$ is the unique adapted solution of BSDE (13), then we have*

$$\begin{aligned} \widetilde{y}_t &= K_t\widetilde{x}_t + \psi_t, \quad \widetilde{z}_t := (\widetilde{z}_t^1, \cdots, \widetilde{z}_t^d), \\ \widetilde{z}_t^i &= [L_t^i\widetilde{x}_t + K_t(C_t^i\widetilde{x}_t + D_t^i\widetilde{u}_t + g_t^i)] + \phi_t^i, \quad i = 1, \ldots, d \end{aligned} \tag{14}$$

and the optimal control is

$$\begin{aligned} \widetilde{u}_t(h) &= \Gamma(t, K_t, L_t)\widetilde{x}_t \\ &\quad -[N_t + \sum_{i=1}^{d}(D_t^i)'KD_t^i]^{-1}[\beta_t + B_t'\psi_t + \sum_{i=1}^{d}(D_t^i)'(\phi_t^i + K_t g_t^i)]. \end{aligned}$$

The value function V has the following explicit formula

$$V(t, x) = \frac{1}{2}\langle K(t)x, x\rangle + \langle \psi(t), x\rangle + \frac{1}{2}V^0(t), \quad (t, x) \in [0, T] \times R^n \tag{15}$$

where

$$V^0(t) := 2E^{\mathcal{F}_t}\int_t^T \langle \psi, f\rangle\, ds + E^{\mathcal{F}_t}\int_t^T \sum_{i=1}^d (\langle K g^i, g^i\rangle + 2\langle \phi^i, g^i\rangle)\, ds$$
$$-E^{\mathcal{F}_t}\int_t^T \langle (N+\sum_{i=1}^d (D^i)' K D^i) u^0, u^0\rangle\, ds.$$

and

$$u^0 := [N + \sum_{i=1}^d (D^i)' K D^i]^{-1}[\beta + B'\psi + \sum_{i=1}^d (D^i)'(\phi^i + K g^i)]. \qquad (16)$$

The proof consists of straightforward verification arguments using Itô's formula.

6 Conclusion

Under assumptions (A1), (A2) (or (A2)'), and (A3), the results in Section 5 give the following interrelationship from a mechanical viewpoint:

Stochastic LQ problem for arbitrary initial state

$\Updownarrow$ (a)

Stochastic Maximum Principle for arbitrary initial state
(Linear Stochastic Hamilton System for arbitrary initial state)

$\Updownarrow$ (b)

Stochastic Riccati equation
(Hamilton-Jacobi-Bellman Equation)
Bismut-Peng's Open Problem

In particular, the "down-arrow" relation in the equivalence (b) is completely new, and contains at least the following three contributions:

(i) It gives a rather satisfactory solution to Bismut-Peng's problem on the solution of Riccati equation (9).

(ii) It completes the interrelationship between the stochastic LQ problem, the linear stochastic Hamilton system, and the Riccati equation (9)—a special version of the HJB equation.

(iii) It shows that no gap exists between the stochastic MP and the stochastic Riccati equation—the same fact as demonstrated in the deterministic case—at least in the above-concerned situation, which is contrast to what was argued in Section 5 of Chen, Li, and Zhou[3] and which therefore might be very surprising from their viewpoint.

Acknowledgments

This work is supported by a Research Fellowship from the Alexander von Humboldt Foundation and by the National Natural Science Foundation of China under Project 79790130. It is completed during his visit to the Center of Finance and Econometrics and the Department of Mathematics and Statistics, University of Konstanz, 78457 Konstanz, Germany. (tang@fmi.uni-konstanz.de). The support from the center and the department is also gratefully acknowledged. The author would like to thank Prof. Michael Kohlmann for his hospitality.

1. J. M. Bismut, Linear quadratic optimal stochastic control with random coefficients, *SIAM J. Control Optim.* **14** (1976) 419–444.
2. J. M. Bismut, Controle des systems lineares quadratiques: applications de l'integrale stochastique, in: C. Dellacherie, P. A. Meyer et M. Weil, eds., *Séminaire de Probabilités XII, LNM 649* (Springer-Verlag, Berlin, 1978) 180–264.
3. S. Chen, X. Li and X. Zhou, Stochastic linear quadratic regulators with indefinite control weight costs, *SIAM J. Control Optim.* **36** (1998) 1685–1702.
4. S. Chen and J. Yong, Stochastic linear quadratic optimal control problems, *Appl. Math. Optim.* **43** (2001) 21–45.
5. S. Chen and J. Yong, Stochastic linear quadratic optimal control problems, II, *preprint*
6. S. Chen and J. Yong, Solvability of a stochastic linear quadratic optimal control problem, *preprint*
7. S. Chen and X. Zhou, Stochastic linear quadratic regulators with indefinite control weight costs, II, *SIAM J. Control Optim.* **39** (2000) 1065–1081.
8. F. Delbaen and W. Schachermayer, The variance-optimal martingale measure for continuous processes, *Bernoulli* **2** (1996), 81–105
9. D. Duffie and H. R. Richardson, Mean-variance hedging in continuous time, *Ann. Appl. Prob.* **1** (1991) 1–15.
10. L. I. Gal'chuk, Existence and uniqueness of a solution for stochastic equations with respect to semimartingales, *Theory of Probability and Its Applications* **23** (1978) 751–763.

11. C. Gourieroux, J. P. Laurent, and H. Pham, Mean-variance hedging and numéraire, *Mathematical Finance* **8** (1998) 179–200.
12. S. Heston, A closed-form solution for option with stochastic volatility with applications to bond and currency options, *Rev. Financial Studies* **6** (1993) 327–343.
13. J. Hull and A. White, The pricing of options on assets with stochastic volatilities, *J. Finance* **42** (1987) 281–300.
14. M. Kobylanski, Résultats d'existence et d'unicité pour des équations différentielles stochastiques rétrogrades avec des générateurs á croissance quadratique, *C. R. Acad. Sci. Paris* **324, Série I** (1997) 81–86.
15. M. Kohlmann and S. Tang, Minimization of risk, LQ theory, and mean-variance hedging and pricing, *submitted*
16. M. Kohlmann and S. Tang, Global adapted solution of one-dimensional backward stochastic Riccati equations with application to the mean-variance hedging, *submitted*
17. M. Kohlmann and S. Tang, Multi-dimensional backward stochastic Riccati equations and applications, *submitted*
18. M. Kohlmann and S. Tang, New developments in backward stochastic Riccati equations and their applications, in: M. Kohlmann, S. Tang, eds., *Mathematical Finance* (Birkhaeuser, Basel, Boston, Berlin, 2001) 194–214.
19. M. Kohlmann and X. Zhou, Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach, *SIAM J. Control Optim.* **38** (2000) 1392–1407.
20. A. E. B. Lim and X. Zhou, LQ control and mean-variance portfolio selections: the stochastic parameter case, in: M. Kohlmann, S. Tang, eds., *Mathematical Finance* (Birkhaeuser, Basel, Boston, Berlin, 2001) 261–270.
21. J. P. Laurent and H. Pham, Dynamic programming and mean-variance hedging, *Finance and Stochastics* **3**(1999) 83–110.
22. J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations* (Springer-Verlag, Berlin Heidelberg, New York, 1971).
23. E. Pardoux and S. Peng, Adapted solution of backward stochastic equation, *Systems Control Lett.* **14** (1990) 55–61.
24. S. Peng, Stochastic Hamilton-Jacobi-Bellman equations, *SIAM J. Control Optim.* **30** (1992) 284–304.
25. S. Peng, Some open problems on backward stochastic differential equations,in: S. Chen, X. Li, J. Yong, X. Zhou, eds., *Control of distributed parameter and stochastic systems* (Kluwer Academic Publishers, 1999) 265–273.

26. S. Peng, A general stochastic maximum principle for optimal control problems, *SIAM J. Control Optim.* **28** (1990) 966–979.
27. H. Pham, T. Rheinländer, and M. Schweizer, Mean-variance hedging for continuous processes: New proofs and examples, *Finance and Stochastics* **2** (1998) 173–198.
28. M. Schweizer, Mean-variance hedging for general claims, *Ann. Appl. Prob.* **2** (1992) 171–179.
29. M. Schweizer, Approaching random variables by stochastic integrals, *Ann. Probab.* **22** (1994) 1536–1575.
30. M. Schweizer, Approximation pricing and the variance-optimal martingale measure, *Ann. Probab.* **24** (1996) 206–236.
31. E. M. Stein and J. C. Stein, Stock price distributions with stochastic volatility: an analytic approach, *Rev. Financial Studies* 4 (1991) 727–752.
32. S. Tang, General linear quadratic optimal stochastic control problems with random coefficients: linear stochastic Hamilton systems and backward stochastic Riccati equations, *submitted*
33. S. Tang and X. Li, Necessary conditions for optimal control of stochastic systems with random jumps, *SIAM J. Control Optim.* **32** (1994) 1447–1475.
34. J. Yong and X. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB equations* (Springer-Verlag, Berlin, New York, 1999).
35. W. M. Wonham, On a matrix Riccati equation of stochastic control, *SIAM J. Control Optim.* **6** (1968) 312–326.

Options on Dividend Paying Stocks

Reimer Beneder and Ton Vorst
Econometric Institute, Erasmus University Rotterdam
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
E-mail: vorst@few.eur.nl

In this paper we describe arbitrage opportunities that result from applying a standard price methodology for options on stocks paying discrete dividends. The main reason is the reduction of volatility that comes with the use of clean stock prices in calculating option prices. We propose a method that adjusts the volatility. The accuracy of this method is assessed by comparing the valuation of options with those generated by Monte Carlo simulation. Overall, the volatility adjustment leads to a significant increase in accuracy compared with the application of the straightforward Black-Scholes formula.

1 Introduction

A popular way to value European options on stocks that pay discrete dividends is to adjust the stock price by subtracting the present value of future dividends that are paid before option maturity. Next, this adjusted or clean stock price is substituted in the Black-Scholes formula to derive the option value.

American option values are calculated in a similar way. A binomial tree is built for the clean stock price. Next, for each time layer in the tree the prevalent present values of not yet paid dividends are added back to the stock price in each point of the time layer. In this way a recombining binomial tree results.

By using the familiar backward procedure on this tree one can calculate American option prices. This is consistent with the described way to calculate European option prices. I.e. European prices calculated with the tree are, except for the usual convergence inaccuracies of binomial trees, equal to the Black-Scholes prices.

This approach leads to a most attractive cash and carry arbitrage opportunity. Using software that comes with one of the most popular academic books on option pricing theory and uses the above described method to adjust for dividends we found the following values for American call options: For a stock that pays a dividend of 4 after 51 weeks, with a spot price of 100, a one year at-the-money American call costs 13.60, while a 50 week call with the same strike price costs 13.89.

By buying the one-year call and shorting the 50 week call one makes an instantaneous profit. By exercising the long call whenever the short one is exercised, future net cash flows will be zero or even positive if the short call is not exercised and the long call can be profitably exercised after the maturity of the short call. The reason behind this inconsistency is that in case of adjustment for dividends which are added back as long as they are not yet paid reduces the volatility of the binomial tree, resulting in lower option prices. Also for European options we will show some remarkable price differences, although these can not be arbitraged with a cash and carry strategy, since the above described strategy heavily depends on the American character of the options.

To circumvent the reduction in volatility in the tree one might build a standard tree in which dividends are subtracted at the moment they are paid. In this way a non-recombining tree results, which can be computationally intensive. Hence, usually some fixing of the tree is done around dividend payments to maintain a recombining tree. With these enhancements the pricing inconsistencies, i.e. arbitrage opportunities disappear for American options. Of course, one can also calculate European option prices through these enhanced trees. However, in this paper we propose a method for European options that is a simple extension of the Black-Scholes formula using the clean stock price and adjusted volatilities. Hence, one does not have to use a computationally intensive tree. We will also show how this method can be extended to different kinds of exotic options.

This paper is organised as follows. In the next section we will show the cause of the arbitrage opportunities in the American call option prices. In section 3 we will discuss how this is remedied in the literature for American options. Section 4, which is the main part of the paper, describes a method to adjust volatilities for European options in order to eliminate the pricing inconsistencies. Finally the last section concludes the paper.

2 Volatility Reduction

In this section we will explain the reason behind the arbitrage opportunity for American options as described in the introduction. We will use the familiar Cox, Ross and Rubinstein (1979) tree approach. Figure 1 shows a binomial tree for the option that expires in 50 weeks, just before the dividend payment. The spot price of the underlying is 100, the continuous interest rate is 5%, and the volatility of the underlying is 30%. We have taken a tree with 100 steps, hence each step represents half a week. The tree has been build in the familiar way, where a stock price S can go to Su or Sd, with $d = 1/u$ and

$$u = e^{\sigma\sqrt{\Delta t}}, \tag{1}$$

with $\Delta t = 1/104$. Hence, in our case $u = 1.0299$

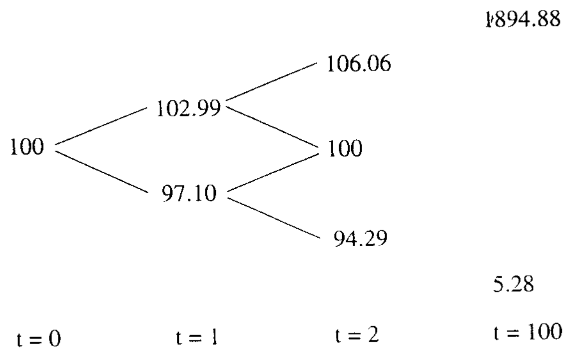


Figure 1: Binomial tree without dividends

Figure 2 shows a binomial tree for the option on the same underlying that expires in 52 weeks, just after the dividend of 4 that is paid one week earlier. For this tree we use 104 steps, hence again each step represents half a week. For this tree we subtract the present value of the dividend from the spot price, to get the clean spot price S^*, with

$$S^* = S - PV(Div) \tag{2}$$

Next we build a tree for the clean stock price by again multiplying by u and d. Finally, in each point of the tree we add back the then present value of the dividend.

When we look at the first few steps of the tree, we see that the spread in stock prices is smaller for the tree with dividend than the one without. On the right hand side of Figure 1 we see the extreme values of the stock price after 100 steps, i.e. after 50 weeks, which is the maturity of the first option. In Figure 2 the values at the right represent not the extreme final values, but again the extreme values after 50 weeks. Also here we see a much larger spread for the option that matures just before the dividend. In fact for up to 50 weeks the trees should represent the same stock price, but we see remarkable differences. It is no surprise that the first tree leads to higher option prices, since it obviously has a higher volatility. This explains why the 50 weeks American option has a value of 13.89 and the 52 weeks American option has a value of 13.60. If we would take a dividend of 10, the one year option price would be 12.95. The parameter combination in the example is not very specific and the results hold for a large range of parameters. A possibility to remove the arbitrage opportunity would be to value the 50 weeks option

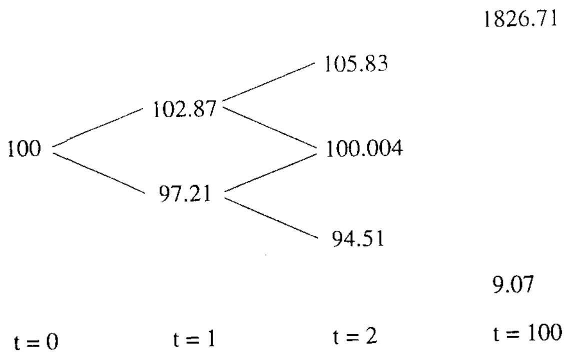


Figure 2: Binomial tree with dividends

with the tree from Figure 2. This would result in an option value of 13.38. The arbitrage opportunity no longer exists. However, the option valuation becomes dependent on dividends that are paid after the option maturity. This would raise the question, how many future dividends have to be taken into account and how precisely these can be estimated.

For the European counterparts we find values of 13.92 and 11.95 respectively for the 50 week and one year calls. The 13.92 is roughly equal to the binomial value of 13.89 since American calls are never exercised before maturity if no dividends are paid. However, once again the 11.95 is low compared to the 13.92.

It is also interesting to compare the binomial tree value of 13.60 for the one year option with the so-called Black approximation (see Hull, 1999). The Black approximation gives a lower bound for the American option value by taking the maximum of the European option value and the value of a similar European option, which matures just before the stock goes ex-dividend. The latter value in this case is 14.07 while the first one is 11.95 as specified earlier on. Hence, the Black lower bound would be 14.07, indicating that the tree gives a mispricing of at least 47 cents, which is quite substantial.

3 American Options

There are several ways to circumvent the arbitrage opportunities in pricing American options described in the previous section. One way, as described e.g. in Ødegaard (1999), is to build a standard binomial tree with the desired volatility until the first ex-dividend date. In the time layer of the first dividend, the dividend is subtracted and the tree is build further from these points. However, from this point on the tree is no longer recombining. Consider a point in the tree just before the dividend with stock price $\tilde{S}$. The point just above has value $\tilde{S}u^2$. After the dividend payment these points become $(\tilde{S}-D)$ and $(\tilde{S}u^2 - D)$. If the first point moves up with factor u and the second goes down with a factor d the values $(\tilde{S}u - Du)$ and $(\tilde{S}u^2d - Dd) = (\tilde{S}u - Dd)$ result. These are not equal, which should hold for a recombining tree. Hence, if the total number of steps in the tree is n and the dividend is after m steps, one finds $(m+1)$ new trees from there onwards, each with $(n-m)$ steps. If there is more than one dividend each time the tree splits further. With this method one can input the desired volatility everywhere, however, the method is very computationally intensive.

An alternative method for American options is where the tree is forced to recombine after each dividend payment. Such a method is described in e.g. Wilmott, Dewynne and Howison (1996), where after the dividend payment new u's and d's are defined. In order to make a tree recombining the relative distances between consecutive points at the same time the layer should be equal to u^2. Hence, if we have n points in the layer at the dividend date one usually specifies the next layer with $m+1$ points, where the relative distances are again u^2 (see figure 3). In the first step immediately after the dividend payment one no longer has upshifts with a factor u and d but these u's and d's depend on the nodes, as indicated in figure 3. In the familiar backward procedure the risk neutral probabilities in this step are adjusted to reflect the different u's and d's. With these probabilities the option values in the $S_{i,j} - D$ points are calculated. The values in the $S_{i,j}$ points just above, are derived by taking the maximum of the value in $S_{i,j} - D$ and early exercising against a price of $S_{i,j}$.

Alternatively, one might calculate the values in the $S_{i,j} - D$ points by multiplying these points with factor u and d which will result in values of the stock price which are in between values in the next layer. Now for these new points, option values are calculated by interpolating the option values in the existing tree points and next to these interpolated values the familiar backward procedure is applied with the standard risk neutral probabilities. Once again values in $S_{i,j}$ are calculated by considering the feasibility of early exercise. The problem with both methods is that one might not be able to construct the next layer such that the $S_{i,j} - D$ points everywhere fall in between two discounted values for the next time layer. If this is not the case for some point, negative risk neutral probabilities will result at that point.

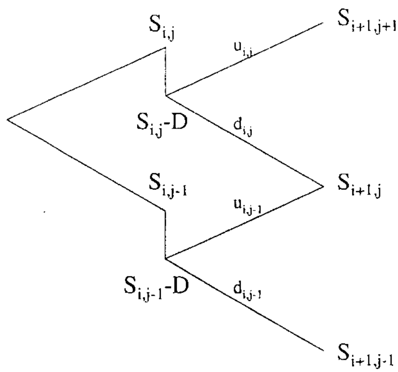


Figure 3: Recombining tree

4 European Options through Volatility Adjustments

All methods described in the previous section come with some problems, being computation time or negative probabilities. Hence, it is important for European options that we do not have to rely on trees. In this section we show an alternative method that can be used to value European options on stocks paying discrete dividend avoiding the usual volatility reduction. We will mainly focus on call options. As before the stockprice can be divided into two components: a riskless component that equals the discounted dividend and a risky component that equals the clean stock price S^* described by equation (2). Our approach consists of specifying a volatility that is a weighted average of an adjusted and an unadjusted volatility where the weighting depends on the timing of the dividend.

The standard assumption in option pricing theory is that the stock price follows a geometric Brownian motion. At the ex-dividend date the stock price drops by the dividend amount (without taking fiscal factors into account). However, to value a European call option we will assume that the risky component follows a geometric Brownian motion. Hence, we write the stochastic process for the risky component as follows

$$dS^* = rS^* dt + \hat{\sigma} S^* dZ \tag{3}$$

The volatility of the risky component is calculated by interpolating between the volatility before and after the dividend payment. The volatility of the risky component before the dividend payment is obtained by requiring a similar impact of the Brownian motion dZ as in the stock price process. Therefore, we multiply the stock volatility by S/S^*. The volatility after the dividend payment is equal to the volatility of the stock. The weighting of the volatilities depends on the timing of the dividend and is given by formula (4), where τ is the time of the dividend payment.

$$\hat{\sigma} = \sqrt{\frac{(\sigma S/S^*)^2\tau + \sigma^2(T-\tau)}{T}} \tag{4}$$

Formula (4) can be justified since it implies that we add the variance of the stock price in the period before the dividend payment to the variance of the stock price over the period after the dividend payment. This is justified under the assumption of time independence of stock price movements which comes with a geometric Brownian motion.

In order to value the option the Black-Scholes formula can be used with S^* and the volatility for the risky component $\hat{\sigma}$.

$$c = S^* N(d_1) - Xe^{-rt}N(d_2) \tag{5}$$

where,

$$d_1 = \frac{\ln(S^*/X) + (r + \hat{\sigma}^2/2)T}{\hat{\sigma}\sqrt{T}} \quad \text{and} \quad d_2 = d_1 - \hat{\sigma}\sqrt{T}$$

The results from this method are compared with results from the Monte Carlo simulation. In each run the stockprice is simulated until the ex-dividend date. The dividend is subtracted and the stockprice is further simulated until maturity of the option. In this way we indeed always have the correct volatility. The option value is based on 10 million runs to minimize the standard error of the simulation.

In table 1 results for five examples are summarized, where τ is the dividend date. The call option values in the second and third column are calculated by using formula (5) without and with adjusting the volatility. The t-value in the last column is the difference between the Monte Carlo value and the analytic* value divided by the standard error of the simulation. Almost all of the simulation results are within twice the standard error bound around the analytic* results.

We indeed see that the adjusted volatility based on the Black-Scholes formula in all cases better approximates the Monte Carlo simulated value than

Table 1: Call Option Values for Dividend Paying Stocks

τ	Analytic	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
$S = 100, X = 100, D = 5, T = 1, r = 0.05, \sigma = 0.3$						
τ	Analytic	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.1	11.287	11.348	0.3016	11.348	0.0063	0.06
0.5	11.342	11.634	0.3078	11.637	0.0065	0.34
0.8	11.383	11.840	0.3122	11.843	0.0066	0.58
$S = 100, X = 100, D = 5, T = 1, r = 0.05, \sigma = 0.15$						
τ	Analytic	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.1	5.635	5.666	0.1508	5.666	0.00288	0.05
0.5	5.687	5.835	0.1539	5.835	0.00296	0.07
0.8	5.726	5.956	0.1561	5.956	0.00301	0.13
$S = 100, X = 100, D = 5, T = 1, r = 0.10, \sigma = 0.30$						
τ	Analytic	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.1	13.492	13.549	0.3016	13.550	0.00680	0.09
0.5	13.613	13.886	0.3076	13.892	0.00697	0.85
0.8	13.701	14.122	0.3117	14.135	0.00707	1.81
$S = 100, X = 100, D = 10, T = 1, r = 0.05, \sigma = 0.30$						
τ	Analytic	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.1	8.685	8.810	0.3035	8.810	0.0055	0.03
0.5	8.782	9.380	0.3166	9.376	0.0059	-0.65
0.8	8.853	9.782	0.3258	9.770	0.0061	-1.87
$S = 100, X = 90, D = 50, T = 1, r = 0.05, \sigma = 0.30$						
τ	Analytic	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.1	16.115	16.169	0.3016	16.170	0.00727	0.11
0.5	16.184	16.444	0.3078	16.454	0.00742	1.31
0.8	16.234	16.642	0.3122	16.664	0.00752	2.91

the straightforward Black-Scholes formula, especially in cases where the dividend is paid close to maturity. In these cases we also see that the volatility adjustment as given in column (4) is quite substantial.

The adjustment assumes that just one dividend is paid before option maturity. However, it is clear that the method can also be extended to cases where more dividends are paid. We have applied the volatility adjustment for a flat volatility structure. However, also with a term structure of volatilities the method can be applied by calculating forward volatilities, adjusting all forward volatilities up to the dividend rate and then once again aggregating to a total volatility over the life time of the option as in formula (4).

The method can not only be applied to standard options, but definitely also to some exotic options.

Consider for example an option to exchange stock U for stock V at time T. These options are usually European since they come in packages with similar and other options that have to be exercised together. Margrabe (1978) derived an analytic formula to value such a European option, which is in fact based on the change of a numeraire method as described by Geman, El Karoui and Rochet (1995), where U can be seen as numeraire. Now suppose the stocks will pay known dividends D_U and D_V at τ_U and τ_V respectively. Based on the change of numeraire technique one might build a one dimensional binomial tree. However, it is not clear by how much the ratio of the two stock prices will drop (or rise) at a dividend payment of one the two stocks. Furthermore, given the mainly European character a kind of Black-Scholes formula would be helpful. As with a call option the analytic formula can still be used with the discounted dividends subtracted of the stock prices and an adjusted volatility.

$$c = S_V^* N(d_1) - S_U^* N(d_2) \tag{6}$$

where,

$$d_1 = \frac{\ln(S_V^*/S_U^*) + \hat{\sigma}^2 T/2}{\hat{\sigma}\sqrt{T}} \quad \text{and} \quad d_2 = d_1 - \hat{\sigma}\sqrt{T}$$

Furthermore,

$$S_V^* = S_V - D_V e^{-r\tau_V} \quad \text{and} \quad S_U^* = S_U - D_U e^{-r\tau_U}$$

The procedure to calculate the volatility σ for non-dividend paying stocks is again based on independence over time, which allows adding variances as follows:

$$\sigma = \sqrt{\sigma_U^2 + \sigma_V^2 - 2\rho\sigma_U\sigma_V} \tag{7}$$

Suppose $\tau_U < \tau_V$, then we consider the intervals $[0, \tau_U]$, $[\tau_U, \tau_V]$ and $[\tau_V, T]$. For each interval the volatility is calculated by using the volatilities of the risky components of stock U and V in formula (7). The adjusted volatility for dividend paying stocks is calculated by interpolating the volatilities of the three intervals.

$$\hat{\sigma} = \sqrt{\frac{\sigma_1^2 \tau_U + \sigma_2^2(\tau_V - \tau_U) + \sigma_3^2(T - \tau_V)}{T}} \tag{8}$$

where,

$$\begin{aligned}
\sigma_1^2 &= (\sigma_U S_U/S_U^*)^2 + (\sigma_V S_V/S_V^*)^2 - 2\rho(\sigma_U S_U/S_U^*)(\sigma_V S_V/S_V^*) \\
\sigma_2^2 &= \sigma_U^2 + (\sigma_V S_V/S_V^*)^2 - 2\rho\sigma_U(\sigma_V S_V/S_V^*)^2 \\
\sigma_3^2 &= \sigma_U^2 + \sigma_V^2 - 2\rho\sigma_U\sigma_V
\end{aligned}$$

The argument is analogous for $\tau_U > \tau_V$.

To calculate the value of the exchange option with Monte Carlo simulation the life of the option is divided in the same three intervals. In each run the stockprices are simulated until the first ex-dividend date, then until the second ex-dividend date and then until maturity. Correlated samples from the standard normal distribution are required for the simulation. The value of the exchange option is the average value after 10 million runs.

Table 2 provides six examples where the option values are calculated by the analytic formula as well as by simulation. τ_U and τ_V are the dividend payment dates.

Again, the difference between the Monte Carlo value and the analytic* value is for all but one example less than twice the standard error of the simulation. As before the adjusted volatility approach leads to much better results than without adjustments. Also in this case the method can be easily extended to cases with more dividend payments and term structures of volatilities.

Another application of the technique is for compound options. A compound option is an option on an option. In the following example a call on a call is considered, but for a call on a put, a put on a call or a put on a put the same reasoning can be used. A compound option has two exercise dates and two strike prices. So a call on a call gives the holder the right to buy at the first exercise date T_1 a European call option for an amount X_1. The call option expires at the second exercise date T_2 (obviously $T_2 > T_1$) and has a strike price of X_2. Geske (1979) has given an analytic formula to value a compound option. When the risky component of the stock price and adjusted volatilities are inserted in this formula it can be used to value a compound option on dividend paying stocks. Using M as the cumulative bivariate normal distribution function, this means that the value of a call on a call is

$$S^* M(a_1, b_1, \sqrt{T_1/T_2}) - X_2 e^{-rT_2} M(a_2, b_2, \sqrt{T_1/T_2}) - e^{-rT_1} X_1 N(a_2) \tag{9}$$

Table 2: Option to Exchange U for V, both Dividend Paying Stocks

$S_U = 100, S_V = 100, D_U = 2, D_V = 5, T = 1, \sigma_U = 0.20, \sigma_V = 0.40, \rho = 0.50, r = 0.05$								
τ_U	τ_V	Analytic	σ	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.8	0.3	11.831	0.3464	12.039	0.3519	12.037	0.00753	-0.27
0.6	0.5	11.865	0.3464	12.206	0.3554	12.210	0.00763	0.55
0.4	0.7	11.899	0.3464	12.370	0.3589	12.371	0.00772	0.12
$S_U = 100, S_V = 100, D_U = 2, D_V = 5, T = 1, \sigma_U = 0.10, \sigma_V = 0.50, \rho = 0.50, r = 0.05$								
τ_U	τ_V	Analytic	σ	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.8	0.3	16.040	0.4583	16.312	0.4655	16.313	0.01109	0.08
0.6	0.5	16.075	0.4583	16.538	0.4706	16.546	0.01122	0.71
0.4	0.7	16.110	0.4583	16.758	0.4756	16.765	0.01135	0.55
$S_U = 100, S_V = 100, D_U = 2, D_V = 5, T = 1, \sigma_U = 0.20, \sigma_V = 0.40, \rho = -0.30, r = 0.05$								
τ_U	τ_V	Analytic	σ	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.8	0.3	17.525	0.4980	17.817	0.5058	17.818	0.01025	0.01
0.6	0.5	17.560	0.4980	17.975	0.5091	17.985	0.01034	1.02
0.4	0.7	17.596	0.4980	18.127	0.5123	18.133	0.01043	0.61
$S_U = 100, S_V = 100, D_U = 2, D_V = 5, T = 1, \sigma_U = 0.20, \sigma_V = 0.40, \rho = 0.50, r = 0.10$								
τ_U	τ_V	Analytic	σ	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.8	0.3	11.840	0.3464	12.045	0.3518	12.043	0.00754	-0.27
0.6	0.5	11.907	0.3464	12.239	0.3552	12.243	0.00764	0.56
0.4	0.7	11.972	0.3464	12.427	0.3584	12.428	0.00774	0.14
$S_U = 100, S_V = 100, D_U = 10, D_V = 5, T = 1, \sigma_U = 0.20, \sigma_V = 0.40, \rho = 0.50, r = 0.05$								
τ_U	τ_V	Analytic	σ	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.8	0.3	15.227	0.3464	15.438	0.3523	15.442	0.00828	0.48
0.6	0.5	15.305	0.3464	15.634	0.3556	15.647	0.00840	1.50
0.4	0.7	15.383	0.3464	15.833	0.3589	15.847	0.00851	1.62
$S_U = 90, S_V = 100, D_U = 2, D_V = 5, T = 1, \sigma_U = 0.20, \sigma_V = 0.40, \rho = 0.50, r = 0.05$								
τ_U	τ_V	Analytic	σ	Analytic*	$\hat{\sigma}$	MC Sim	Std Error	t-value
0.8	0.3	16.393	0.3464	16.586	0.3519	16.591	0.00856	0.58
0.6	0.5	16.435	0.3464	16.751	0.3554	16.766	0.00866	1.75
0.4	0.7	16.477	0.3464	16.914	0.3589	16.932	0.00876	2.10

where,

$$a_1 = \frac{\ln(S^*/\bar{S}) + (r + \hat{\sigma}_1^2/2)T_1}{\hat{\sigma}_1\sqrt{T_1}} \quad \text{and} \quad a_2 = a_1 - \hat{\sigma}_1\sqrt{T_1}$$

$$b_1 = \frac{\ln(S^*/X_2) + (r + \hat{\sigma}_2^2/2)T_2}{\hat{\sigma}_2\sqrt{T_2}} \quad \text{and} \quad b_2 = b_1 - \hat{\sigma}_2\sqrt{T_2}$$

The variable $\bar{S}$ is the stock price for which the call option price at time T_1 equals X_1, minus the present value at T_1 of the dividend. So if the ex-dividend date is after the first exercise date ($\tau > T_1$) formula (5) for pricing a call option on a dividend paying stock can be used in the calculation of $\bar{S}$. When the dividend is already paid at T_1 ($\tau < T_1$) then $\bar{S}$ is simply the stock price for which the value of the call option at T_1 equals X_1. The adjusted volatilities $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are calculated by using formula (4) with respectively T_1 and T_2 instead of T. In case the ex-dividend date is after the first exercise date then $\hat{\sigma}_1$ equals the volatility of the risky component before dividend payment ($\sigma S/S^*$).

To obtain the value of a compound option by Monte Carlo simulation, the stock price in each run is simulated until the first exercise date. For each stock price at T_1 the option value is calculated by using the Black-Scholes formula if the dividend is already paid or by using formula (5) if the ex-dividend date is after the first exercise date. The option value is the average value after 10 million runs.

The results for some examples are summarized in table 3. Once again the results are very satisfying and are definitely an improvement on the standard method without adjusting the volatility.

5 Conclusion

In this paper we have described arbitrage opportunities that result from applying a standard price methodology for American options on stocks paying discrete dividends. The main reason is the reduction of volatility that comes with the use of clean stock prices in calculating option prices. We propose a method that adjusts the volatility. The accuracy of this method is assessed by comparing the valuation of options with those generated by Monte Carlo simulation. Overall, the volatility adjustment leads to a significant increase in accuracy compared with the application of the straightforward Black-Scholes formula.

The price and volatility adjustments are quite substantial. One might wonder why market participants that use different methods with the same implied volatility can ever trade with each other given the price differences. A

Table 3: Compound Option Values for Dividend Paying Stocks

$S = 100, X_1 = 4, X_2 = 100, D = 5, T_1 = 0.3, T_2 = 1, r = 0.05, \sigma = 0.30$							
τ	Analytic	Analytic*	$\hat{\sigma}_1$	$\hat{\sigma}_2$	MC Sim	Std Error	t-value
0.1	7.719	7.794	0.305	0.302	7.795	0.00291	0.19
0.5	7.771	8.086	0.315	0.308	8.086	0.00301	0.02
0.8	7.809	8.263	0.315	0.312	8.259	0.00302	-1.23
$S = 100, X_1 = 4, X_2 = 100, D = 5, T_1 = 0.7, T_2 = 1, r = 0.05, \sigma = 0.30$							
τ	Analytic	Analytic*	$\hat{\sigma}_1$	$\hat{\sigma}_2$	MC Sim	Std Error	t-value
0.1	8.730	8.797	0.302	0.302	8.797	0.00472	0.13
0.5	8.779	9.103	0.311	0.308	9.103	0.00488	-0.11
0.8	8.816	9.297	0.315	0.312	9.293	0.00495	-0.75
$S = 100, X_1 = 4, X_2 = 100, D = 5, T_1 = 0.3, T_2 = 1, r = 0.05, \sigma = 0.15$							
τ	Analytic	Analytic*	$\hat{\sigma}_1$	$\hat{\sigma}_2$	MC Sim	Std Error	t-value
0.1	2.490	2.534	0.153	0.151	2.534	0.00120	0.17
0.5	2.530	2.691	0.158	0.154	2.690	0.00126	-1.31
0.8	2.560	2.771	0.158	0.156	2.767	0.00127	-2.88
$S = 100, X_1 = 4, X_2 = 100, D = 5, T_1 = 0.3, T_2 = 1, r = 0.10, \sigma = 0.30$							
τ	Analytic	Analytic*	$\hat{\sigma}_1$	$\hat{\sigma}_2$	MC Sim	Std Error	t-value
0.1	9.837	9.906	0.305	0.302	9.907	0.00325	0.29
0.5	9.953	10.248	0.315	0.308	10.253	0.00336	1.43
0.8	10.037	10.462	0.315	0.312	10.467	0.00337	1.37
$S = 100, X_1 = 4, X_2 = 90, D = 5, T_1 = 0.3, T_2 = 1, r = 0.05, \sigma = 0.30$							
τ	Analytic	Analytic*	$\hat{\sigma}_1$	$\hat{\sigma}_2$	MC Sim	Std Error	t-value
0.1	12.310	12.373	0.305	0.302	12.374	0.00359	0.28
0.5	12.376	12.655	0.315	0.308	12.664	0.00370	2.41
0.8	12.425	12.839	0.315	0.312	12.851	0.00370	3.17
$S = 100, X_1 = 2, X_2 = 100, D = 5, T_1 = 0.3, T_2 = 1, r = 0.05, \sigma = 0.30$							
τ	Analytic	Analytic*	$\hat{\sigma}_1$	$\hat{\sigma}_2$	MC Sim	Std Error	t-value
0.1	9.388	9.453	0.305	0.302	9.453	0.00300	0.13
0.5	9.442	9.743	0.315	0.308	9.745	0.00310	0.71
0.8	9.482	9.940	0.315	0.312	9.941	0.00310	0.60

possible explanation for them still trading with each other might be that they use different at-the-money implied volatilities in order to arrive at the same prices. These different implieds are then used to set implieds for out and in the money options. It can be expected that based on these implieds, also the price differences in these options are small.

References

1. Black, Fischer and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 1973, 81, 637-54.
2. Cox. J., S. Ross and M. Rubinstein, "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 1979, 7, 229-264.
3. Geman, H., N. El Karoui and J-C. Rochet, "Changes of Numeraire, Changes of Probability and Option Pricing," *Journal of Applied Probability*, 1995, 32, 443-458.
4. Geske, Robert, "The Valuation of Compound Options," *Journal of Financial Economics*, 1979, v7(1), 63-82.
5. Hull, John C., "Options, Futures, & Other Derivatives," 1999, Prentice Hall.
6. Margrabe, William, "The Value of an Option to Exchange one Asset for Another," *Journal of Finance*, 1978, v33(1), 177-186.
7. Ødegaard, B.A., "Financial Numerical Recipes," http://finance.bi.no /~bernt//gcc_prog/algoritms/algoritms/algoritms.html
8. Wilmott, P., J. Dewynne and S. Howison, "Option Pricing: Mathematical Models and Computation, 1996, Oxford Financial Press.

Some Remarks on Arbitrage Pricing Theory

Jianming Xia and Jia-An Yan[a]
Institute of Applied Mathematics
Academy of Mathematics and System Sciences
Chinese Academy of Sciences, Beijing 100080, China

Abstract

In this note we report main results in a recent paper by the authors, in which we established a version of Kramkov's optional decomposition theorem in the setting of equivalent martingale measures and using this theorem we clarified some basic concepts and results in arbitrage pricing theory: superhedging, fair price, replicatable contingent claim, complete markets.

Keywords allowable strategy, martingale measure, no free lunch, numeraire, superhedging, replicating.

1 Introduction

Delbaen and Schachermayer(1994) introduced the notion of No Free Lunch with Vanishing Risk (or NFLVR in short) and showed the fundamental theorem of asset pricing in a general version as following: if the deflated price process of assets in a financial market is a locally bounded vector-valued semimartingale, the condition of NFLVR is equivalent to the existence of an equivalent local martingale measure for the deflated price process. In that paper and in Delbaen and Schachermayer(1998) which extended their result to the case of unbounded processes, the notion of NFLVR was defined w.r.t. admissible (or tame) strategies, and consequently, depends on the choice of numeraire.

Yan (1998) proposed to work with a financial market in which if we take an asset as the numeraire the deflated price processes of other assets admit an equivalent martingale measure instead of a local martingale measure, and called such a market a fair market. By introducing the notion of allowable trading strategy, which is independent of the choice of numeraire, and based on the basic result of Delbaen and Schachermayer(1994), Yan (1998) showed that the fairness of a market is equivalent to the condition of NFLVR w.r.t. allowable strategies. Both the fairness and NFLVR property w.r.t. allowable strategies are independent of the choice of numeraire.

[a]Yan's work was supported by the 973 project on mathematics, the Ministry of Science and Technology. Yan's E-mail address is jayan@mail.amt.ac.cn. Xia's E-mail address is jmxia_sh@yahoo.com.

In a recent paper by the authors we studied fair markets. We established a version of Kramkov's optional decomposition theorem in the setting of equivalent martingale measures and using this theorem we clarified some basic concepts and results in option pricing theory: superhedging, fair price, replicatable contingent claim, complete market. In this note we report these results without giving proofs.

2 The semimartingale model and some existing results

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, and $(\mathcal{F}_t)_{0\leq t\leq T}$ a filtration satisfying usual conditions. $\mathcal{F}_0 = \sigma\{\emptyset, \Omega\}, \mathcal{F}_T = \mathcal{F}$, where positive number T is a fixed and finite time horizon. We consider a financial market which consists of $m+1$ assets. Their price processes S^j, $j = 0, 1, \cdots, m$, are assumed to be strictly positive semimartingales with càdlàg paths. We take asset 0 as the numeraire and call $(S^0)^{-1}$ the deflator process. For notational convenience, we set $S = (S^1, \cdots, S^m), \widetilde{S} = \left(S^1/S^0, \cdots, S^m/S^0\right)$, and call $\widetilde{S}$ the deflated price process of the assets. Note that the deflated price process of asset 0 is the constant 1.

A trading strategy is an $\mathbb{R}^{m+1}$-valued $\mathcal{F}_t$-predictable process $\psi = \{\varphi^0, \varphi\}$ such that ψ is integrable w.r.t. semimartingale (S^0, S), where $\varphi = (\varphi^1, \cdots, \varphi^m)$ and φ_t^j represents the number of units of the asset j held at time t, $0 \leq j \leq m$. The wealth $V_t(\psi)$ of a trading strategy $\psi = \{\varphi^0, \varphi\}$ at time t is $V_t(\psi) = \varphi_t^0 S_t^0 + \varphi_t \bullet S_t$, where $\varphi_t \bullet S_t = \sum_{j=1}^m \varphi_t^j S_t^j$. A trading strategy $\psi = \{\varphi^0, \varphi\}$ is said to be self-financing, if

$$V_t(\psi) = V_0(\psi) + \int_0^t \psi_u d(S_u^0, S_u), \quad t \in [0, T],$$

where $\int_0^t \psi_u d(S_u^0, S_u)$ denotes a vector stochastic integral. We refer the read to Jacod (1979) for the properties of vector semimartingale integrals.

The following result about vector semimartingale integrals seems to be new.

Theorem 2.1 *Let X be an $\mathbb{R}^n$-valued semimartingale and H an $\mathbb{R}^n$-valued predictable process. If $H \in L(X)$ and*

$$H_t \bullet X_t = H_0 \bullet X_0 + \int_0^t H_s dX_s, \tag{2.1}$$

where $\bullet$ denotes the inner product of two vectors and $L(X)$ the class of all $\mathbb{R}^n$-valued predictable processes integrable w.r.t. X, then for any real-valued *semimartingale* y, $H \in L(yX)$ *and*

$$y_t(H \bullet X)_t = y_0(H \bullet X)_0 + \int_0^t H_s d(yX)_s. \tag{2.2}$$

¿From Theorem 2.1 we deduce immediately the following

Lemma 2.1 *A strategy* $\psi = \{\varphi^0, \varphi\}$ *is self-financing if and only if its wealth process* $(V_t(\psi))$ *satisfies*

$$d\left(\frac{V_t(\psi)}{S_t^0}\right) = \varphi_t d\left(\frac{S_t}{S_t^0}\right).$$

The following lemma is an easy consequence of Lemma2.1.

Lemma 2.2 *For any given* $\mathbb{R}^m$*-valued predictable process* φ *which is integrable w.r.t.* S *and a real number* x *there exists a real-valued predictable process* φ^0 *such that* $\{\varphi^0, \varphi\}$ *is a self-financing strategy with initial wealth* x.

Definition 2.1 *We take asset 0 as a numeraire. A trading strategy* ψ *is said to be admissible (or tame) if* ψ *is self-financing and there exists a positive constant* c *such that the wealth* $V_t(\psi)$ *at any time* t *is bounded from below by* $-cS_t^0$.

Definition 2.2 *Let* X *be a vector-valued semimartingale. A probability measure* $\mathbb{Q}$ *is said to be an equivalent (local) martingale measure for* X *if* $\mathbb{Q}$ *is equivalent to* $\mathbb{P}$ *and* X *is a* $\mathbb{Q}$*-(local) martingale.*

Put

$$K = \left\{ \frac{V_T(\psi)}{S_T^0} : \ \psi \text{ is admissible and } V_0(\psi) = 0 \right\}, \ C = (K - L_+^0) \cap L^\infty.$$

The market is said to satisfy the NFLVR (resp. NA) property w.r.t. admissible strategies if

$$\overline{C} \cap L_+^\infty = \{0\} \,(\text{resp. } C \cap L_+^\infty = \{0\}),$$

where NA stands for No-Arbitrage, $\overline{C}$ denotes the closure of C taken in the supnorm topology of L^∞. Note that the NA condition $C \cap L_+^\infty = \{0\}$ is equivalent to the more convincing condition $K \cap L_+^0 = \{0\}$.

The fundamental theorem of asset pricing in the locally bounded case, as in Delbaen and Schachermayer(1994) Theorem 1.1, can now be formulated as follows:

Theorem 2.2 *If* $\widetilde{S}$ *is locally bounded, then the market satisfies the NFLVR property w.r.t. admissible strategies if and only if there exists an equivalent local martingale measure for* $\widetilde{S}$. *In this case the set* C *is already weak*(i.e.* $\sigma(L^\infty, L^1)$*) closed in* L^∞.

As we can see that the notion of admissible strategies is variant under the change of numeraire, so the NFLVR property w.r.t admissible strategies depends on the choice of numeraire. In fact, Delbaen and Schachermayer(1995a,1995b) gave an example $S = (1, R)$, where R is the Bessel(3) process. They showed that the market satisfies the NFLVR property w.r.t. admissible strategies if R is chosen to be numeraire, while the market allows arbitrage w.r.t. admissible strategies if 1 is chosen to be numeraire.

Yan(1998) gave the following

Definition 2.3 *The market is said to be fair if there is an equivalent martingale measure for the deflated price process* $\widetilde{S}$.

Denoted by $\mathcal{M}^j$(resp. $\mathcal{M}^j_{loc}$) the set of all equivalent martingale measures(resp. equivalent local martingale measures) for the deflated price process $(\frac{S^0}{S^j}, \frac{S^1}{S^j}, \cdots, \frac{S^m}{S^j})$, if asset j is taken as numeraire.

Yan(1998) Theorem 2.2 showed that the fairness of the market is independent of the choice of numeraire. In fact, let $\mathcal{M}^0 \neq \emptyset$. For a given j and every $\mathbb{Q}^0 \in \mathcal{M}^0$, $\frac{S^j}{S^0}$ is a $\mathbb{Q}^0$-martingale. We can define an equivalent probability measure $\mathbb{Q}^j$ by

$$\frac{d\mathbb{Q}^j}{d\mathbb{Q}^0} = \frac{S^0_0}{S^j_0} \cdot \frac{S^j_T}{S^0_T}.$$

By Beyes' rule, $\mathbb{Q}^j \in \mathcal{M}^j$ and $\mathbb{Q}^0 \longrightarrow \mathbb{Q}^j$ is a bijection from $\mathcal{M}^0$ onto $\mathcal{M}^j$.

In order to characterize the fairness of the market, Yan(1998) introduced the following definition.

Definition 2.4 *A trading strategy ψ is said to be allowable if ψ is self-financing and there exists a positive constant c such that the wealth $V_t(\psi)$ at any time t is bounded from below by* $-c\sum_{j=0}^{m} S^j_t$.

It can be see that the deflated wealth process of an allowable self-financing strategy is a local $\mathbb{Q}$-martingale and a $\mathbb{Q}$-supermartingale for any $\mathbb{Q} \in \mathcal{M}^0$.

The following theorem is due to Yan(1998). Its proof was based on Theorem 2.2 which is due to Delbaen and Schachermayer(1994).

Theorem 2.3 *The market is fair if and only if there is no sequence (ψ_n) of allowable strategies with initial wealth 0 such that $V_T(\psi_n) \geq -\frac{1}{n}\sum_{j=0}^{m} S^j_T$, a.s., for all $n \geq 1$ and $V_T(\psi_n)$, a.s., tends to a non-negative random variable ξ with* $\mathbb{P}(\xi > 0) > 0$.

Remark 2.1 *1) The notion of allowable strategies does not depend on the choice of numeraire.*

2) According to Delbaen and Schachermayer(1994), the condition in Theorem 2.3 can be called NFLVR w.r.t. allowable strategies. The NFLVR property w.r.t. allowable strategies is independent of the choice of numeraire.

3 Optional decomposition theorem and superhedging

It is well known that optional decomposition theorem is very useful in mathematical finance. But in the existing literature, such as Kramkov(1996) and Föllmer and Kabanov(1998), the theorem was given w.r.t. the equivalent local martingale measures. The following theorem is a version of Kramkov's optional decomposition theorem under the setting of equivalent martingale measures.

Theorem 3.1 *Let S be a vector-valued semimartingale with non-negative components defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, where $(\mathcal{F}_t)$ satisfies the usual conditions. Let the set $\mathcal{M}$ of all equivalent martingale measures for S not be empty. If X is a non-negative $\mathcal{M}$-supermartingale, i.e. $\mathbb{Q}$-supermartingale for all $\mathbb{Q} \in \mathcal{M}$, then there are an adapted, right continuous and increasing process C with $C_0 = 0$, and an S-integrable predictable process φ such that*

$$X = X_0 + \varphi.S - C.$$

Moreover, $\varphi.S$ is a local martingale.

Let ξ be a contingent claim at time T. In general, one can not find a self-financing strategy to replicate ξ, but one can find a minimal value at any time t with which one can cover the claim ξ by a self-financing strategy with non-negative wealth on the time interval $[t, T]$. This minimal value is called the *cost at time t of superhedging ξ*, and is defined as essinf V_t, where V_t runs over the class of $\mathcal{F}_t$-measurable and non-negative random variables such that

$$V_t + \int_t^T \psi d(S^0, S^1, \cdots, S^m) \geq \xi, \text{ a.s.}$$

and

$$V_t + \int_t^u \psi d(S^0, S^1, \cdots, S^m) \geq 0, \text{ a.s., for all } u \in [t, T]$$

for some self-financing strategy ψ.

Here the cost of superhedging does not involve the numeraire. In the literature, for example in Kramkov(1996) and Föllmer and Kabanov(1998), this problem was solved by using the optional decomposition theorem (due to Kramkov) based on the equivalent local martingale measures. The result can be state as follows: We take S^0 as the numeraire and let $\mathcal{M}^0_{loc}$ denote the set of all equivalent local martingale measures for $(\frac{S^1}{S^0}, \cdots, \frac{S^m}{S^0})$. Assume $\mathcal{M}^0_{loc}$ is non-empty. Then the cost at time t of superhedging the claim ξ is given by

$$U_t = \operatorname{esssup}_{\mathbb{Q}\in\mathcal{M}^0_{loc}} S^0_t \mathbb{E}_{\mathbb{Q}}\left[\frac{\xi}{S^0_T}\middle|\mathcal{F}_t\right] \tag{3.1}$$

if

$$\sup_{\mathbb{Q}\in\mathcal{M}^0_{loc}} S^0_0 \mathbb{E}_{\mathbb{Q}}\left[\frac{\xi}{S^0_T}\right] < \infty.$$

However, in such a market, for some numeraire the corresponding local martingale measure may do not exist. The model $S = (1, R)$ in Delbaen and Schachermayer(1995a, 1995b), cited in Section 2, is such an example: if the Bessel(3) process R is taken as the numeraire, the deflated price process $\frac{1}{R}$ is a local martingale and the cost of superhedging is well-expressed by (3.1); if we choose 1 as the numeraire, then the process R admits no equivalent local martingale measure and hence the cost of superhedging has no similar expression as (3.1). This is a paradox.

We show how to express the cost of superhedging in a fair market. First of all, we augment the market by a new asset with price process $S^{m+1} = \sum_{j=0}^{m} S^j$. It is obvious that any self-financing strategy in the augmented market $(S^0, S^1, \cdots, S^{m+1})$ can be expressed as the one in the original market $(S^0, S^1, \cdots, S^m)$. Thus the cost of superhedging in these two markets are the same. Denoted by $\mathcal{M}^{m+1}$ the set of all equivalent martingale measures corresponding to the numeraire S^{m+1}. Since $(\frac{S^0}{S^{m+1}}, \frac{S^1}{S^{m+1}}, \cdots, \frac{S^m}{S^{m+1}})$ is uniformly bounded, $\mathcal{M}^{m+1}$ is just the set of all equivalent local martingale measures for it. Thus, by (3.1), the cost at time t of superhedging the claim ξ is given by

$$U_t = \operatorname{esssup}_{\mathbb{Q}\in\mathcal{M}^{m+1}} S^{m+1}_t \mathbb{E}_{\mathbb{Q}}\left[\frac{\xi}{S^{m+1}_T}\middle|\mathcal{F}_t,\right]$$

if

$$\sup_{\mathbb{Q}\in\mathcal{M}^{m+1}} S^{m+1}_0 \mathbb{E}_{\mathbb{Q}}\left[\frac{\xi}{S^{m+1}_T}\right] < \infty.$$

For every $\mathbb{Q} \in \mathcal{M}^{m+1}$, $\frac{S^0}{S^{m+1}}$ is a $\mathbb{Q}$-martingale. Now let $0 \le j \le m$. We can define a probability measure $\mathbb{Q}^j$ by

$$\frac{d\mathbb{Q}^j}{d\mathbb{Q}} = \frac{S^{m+1}_0}{S^j_0} \cdot \frac{S^j_T}{S^{m+1}_T}.$$

It is easy to see that $\mathbb{Q} \longrightarrow \mathbb{Q}^j$ is a bijection from $\mathcal{M}^{m+1}$ onto $\mathcal{M}^j$. By Beyes' rule, we have

$$U_t = \mathrm{esssup}_{\mathbb{Q}\in\mathcal{M}^j} S_t^j \mathbb{E}_{\mathbb{Q}} \left[\frac{\xi}{S_T^j} \middle| \mathcal{F}_t \right]. \tag{3.2}$$

Clearly, in a fair market, the expression (3.2) doesn't depend on the choice of numeraire.

4 Pricing of contingent claims and completeness of a fair market

First of all we show that even in the Black-Scholes economy, the principle of NFLVR w.r.t. admissible strategies can not determine uniquely the price of a contingent claim. In fact, let W be a standard Brownian motion defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $(\mathcal{F}_t)$ the usual augmentation of the natural filtration of W and $\mathcal{F} = \mathcal{F}_T$. We set $S^0 \equiv 1$ and

$$dS_t^1 = S_t^1(\mu dt + \sigma dW_t),$$

where μ and σ are constants. Define $Z = \mathcal{E}(-\frac{\mu}{\sigma}W)$ and $d\mathbb{Q} = Z_T d\mathbb{P}$, then

$$\mathcal{M}_{loc}^0 = \mathcal{M}^0 = \{\mathbb{Q}\}.$$

In this market, from the results in Section 2, the NFLVR property w.r.t. admissible strategies holds whenever either of the two assets is chosen as numeraire. We can find a non-negative $\mathbb{Q}$-local martingale X such that X is not a $\mathbb{Q}$-martingale. It is clear that X is a $\mathbb{Q}$-supermartingale and $\mathbb{E}_{\mathbb{Q}}[X_T] < X_0$. Let $Y_t = \mathbb{E}_{\mathbb{Q}}[X_T|\mathcal{F}_t]$. It is clear that $X \neq Y$. Both (S^1, X) and (S^1, Y) are $\mathbb{Q}$-local martingales. Thus both X and Y can be candidates for the price process of the contingent claim X_T, under the principle of NFLVR w.r.t. admissible strategies. Which one should we chose? The principle of NFLVR w.r.t. admissible strategies can not give the answer. But the principle of NFLVR w.r.t. allowable strategies can do it! Under the principle of NFLVR w.r.t. allowable strategies, the price process of the claim X_T should be a $\mathbb{Q}$-martingale. So it must be Y.

This example suggests that we should study the pricing of European contingent claims under the principle of NFLVR w.r.t. allowable strategies. That means we should work in a fair market, as described in Section 2. By a *(European) contingent claim* at time T we mean a non-negative $\mathcal{F}_T$-measurable random variable. Let ξ be a contingent claim. Assume that $(S_T^j)^{-1}\xi$ is $\mathbb{Q}$-integrable for some $0 \leq j \leq m$ and some $\mathbb{Q} \in \mathcal{M}^j$. We put

$$V_t = S_t^j \mathbb{E}_{\mathbb{Q}} \left[(S_T^j)^{-1}\xi \mid \mathcal{F}_t \right]. \tag{4.1}$$

If we consider (V_t) as the price process of an asset, then the market augmented with this derivative asset is still fair, because when we take asset j as the numeraire the deflated price process of this derivative asset is a $\mathbb{Q}$-martingale. So it seems that (V_t) can be considered as a candidate for a"fair" price process of ξ. However, generally speaking, if the martingale measure is not unique we can not define uniquely the "fair" price of a contingent claim.

It is clear that a "fair price" of ξ needs not to be the "cost of replicating" ξ by an allowable strategy and the "cost of replicating" ξ by an allowable strategy needs not to be a "fair price" of ξ, either. In general, the fair prices of ξ are not unique. Even if ξ is replicatable by an allowable strategy, the costs of different allowable strategies replicating ξ are usually different. However, we can introduce the following notion of trading strategies:

Definition 4.1 *A trading strategy ψ is said to be regular (resp. strongly regular), if it is allowable and there is a j, $0 \leq j \leq m$, such that $\frac{V(\psi)}{S^j}$ is a $\mathbb{Q}$-martingale (or, equivalently, $\mathbb{E}_{\mathbb{Q}}\left[\frac{V_T(\psi)}{S_T^j}\right] = \frac{V_0(\psi)}{S_0^j}$) for some $\mathbb{Q} \in \mathcal{M}^j$ (resp. for all $\mathbb{Q} \in \mathcal{M}^j$).*

By Beyes' rule, it is easy to see that if ψ is regular (resp. strongly regular), then for each j, $0 \leq j \leq m$, $\frac{V(\psi)}{S^j}$ is a $\mathbb{Q}$-martingale (or, equivalently, $\mathbb{E}_{\mathbb{Q}}\left[\frac{V_T(\psi)}{S_T^j}\right] = \frac{V_0(\psi)}{S_0^j}$) for some $\mathbb{Q} \in \mathcal{M}^j$ (resp. for all $\mathbb{Q} \in \mathcal{M}^j$).

Lemma 4.1 *Let ψ and ψ' be regular strategies with $V_T(\psi) = V_T(\psi')$. Then $V_t(\psi) = V_t(\psi')$ for all $t \in [0,T]$.*

By Theorem 3.1, we can easily deduce the following characterization for contingent claims that can be replicated by regular strategies.

Theorem 4.1 *Let a contingent claim ξ satisfy*

$$\sup_{\mathbb{Q}\in\mathcal{M}^0} \mathbb{E}_{\mathbb{Q}}\left[\frac{\xi}{S_T^0}\right] < \infty$$

(or, equivalently, for all j, $0 \leq j \leq m$, $\sup_{\mathbb{Q}\in\mathcal{M}^j} \mathbb{E}_{\mathbb{Q}}\left[\frac{\xi}{S_T^j}\right] < \infty$). Then the following conditions are equivalent:

1) There is a $j, 0 \leq j \leq m$, and some $\mathbb{Q}^j \in \mathcal{M}^j$ such that

$$\sup_{\mathbb{Q}\in\mathcal{M}^j} \mathbb{E}_{\mathbb{Q}}\left[\frac{\xi}{S_T^j}\right] = \mathbb{E}_{\mathbb{Q}^j}\left[\frac{\xi}{S_T^j}\right];$$

2) For all $j, 0 \leq j \leq m$, there exists some $\mathbb{Q}^j \in \mathcal{M}^j$ such that

$$\sup_{\mathbb{Q}\in\mathcal{M}^j} \mathbb{E}_{\mathbb{Q}}\left[\frac{\xi}{S_T^j}\right] = \mathbb{E}_{\mathbb{Q}^j}\left[\frac{\xi}{S_T^j}\right];$$

3) $\xi = V_T(\psi)$ *for some regular strategy* ψ.

As a consequence of the above theorem, we obtain

Theorem 4.2 *Let* ξ *be a contingent claim. Then the following conditions are equivalent:*

1) There is a $0 \le j \le m$ *such that for all* $Q \in \mathcal{M}^0$ $I\!E_Q\left[\frac{\xi}{S_T^j}\right]$ *are the same constant;*

2) For all $0 \le j \le m$ *and all* $Q \in \mathcal{M}^j$, $I\!E_Q\left[\frac{\xi}{S_T^j}\right]$ *are the same constant;*

3) $\xi = V_T(\psi)$ *for some strongly regular strategy* ψ.

Now we turn to study the completeness of a market. The following theorem is the key for this study.

Theorem 4.3 *Let* ξ *satisfy* $0 \le \xi \le c\sum_{j=0}^{m} S_T^j$ *for some constant* c. *If* ψ *is a regular strtegy which replicates* ξ, *then* ψ *is strongly regular.*

We propose the following definition of complete markets which doesn't depend on the choice of numeraire (see Theorem below).

Definition 4.2 *The market is said to be complete, if any contingent claim* ξ *at time* T *satisfying* $0 \le \xi \le \sum_{j=0}^{m} S_T^j$ *can be replicated by a regular strategy.*

The following theorem gives a characterization for the completeness of a market.

Theorem 4.4 *The following conditions are equivalent:*

1) The market is complete;

2) $\mathcal{M}^0$ *is a singleton;*

3) Any contingent claim ξ *at time* T *satisfying*

$$\sup_{Q \in \mathcal{M}^0} I\!E_Q\left[\frac{\xi}{S_T^0}\right] < \infty$$

(or, equivalently, for all j, $0 \le j \le m$, $\sup_{Q \in \mathcal{M}^j} I\!E_Q\left[\frac{\xi}{S_T^j}\right] < \infty$*) can be replicated by a regular strategy;*

4) There is a j, $0 \le j \le m$, *such that any contingent claim* ξ *satisfying* $0 \le \xi \le S_T^j$ *can be replicated by a regular strategy.*

Acknowledgement. This note is based on an invited talk presented by the second author (J.A. Yan) at the International Conference on Mathematical Finance, held in Fudan University, Shanghai, China. This author is grateful to Fudan University for warm hospitality.

References

Delbaen, F. and W., Schachermayer(1994): A general version of the fundamental theorem of asset pricing, Math. Annalen., **300**, 463-520.

Delbaen, F. and W., Schachermayer(1995a): Arbitrage possibilities in Bessel processes and their relation to local martingales, Probability Theory and Related Fields, **102**, 357-366.

Delbaen, F. and W., Schachermayer(1995b): The no-arbitrage property under a change of numeraire, Stochastics and Stochastic Reports, **53**, 213-226.

Delbaen, F. and W., Schachermayer(1998): The fundamental theorem of asset pricing for unbounded stochastic processes, Math. Annalen., **312**, 215-250.

Föllmer, H. and Yu. M., Kabanov(1998): Optional decomposition theorem and Lagrange multipliers, Finance and Stochastics, **2**(**1**), 69-81.

Jacod, J. (1979): *Calcul stochastique et problémes de martingales*, LN in Math. 714, Springer.

Kramkov, D.O.(1996): Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets, Probability Theory and Related Fields, **105**, 459-479.

Xia, J.M. and J.A. Yan(2001): Clarifying some basic concepts and results in arbitrage pricing theory, Preprint.

Yan, J. A.(1998): A new look at the fundamental theorem of asset pricing, *J. Korean Math. Soc.*, **35**(3), 659-673.

Risk: From Insurance to Finance[a]

Hailiang Yang
Department of Statistics and Actuarial Science
The University of Hong Kong
Pokfulam Road, Hong Kong
email: hlyang@hkusua.hku.hk

Risk and risk management are the most important concerns for insurance companies and financial institutions. In this paper, we will overview some of our research work on this area. In the insurance risk case, we will use ruin probability as a risk measure and discuss the issue of how to estimate it. We will also post some research problems on this subject. For finance risk, we will briefly discuss the risk measures in literature and summarize some of our recent results on coherent risk measures for derivatives. Then we will illustrate how to use some actuarial science techniques to measure financial risk, in particular, credit risk. In this paper, we will focus on the interplay between finance and actuarial science.

Keywords: Risk measures, Insurance risk, Ruin theory, Financial risk, Risk measures for derivatives, Credit rating, Default probability, Default time.

1 Introduction

According to the American Heritage Dictionary, risk can be defined as "the possibility of suffering harm or loss". In Holton (1997), risk is defined as "the exposure to uncertainty". It is well understood, at least in the financial world, that risk has three components. Namely uncertainty, exposure to uncertainty, and the attitude toward risk of a consumer of risk information. Risk management is an important practical issue which also creates many interesting theoretical problems for risk researchers. Alfred Steinherr (1988) stressed the significance of modern risk management by calling it "one of the most important innovations of the 20th Century."

In order to have a better management of risk, as a first step, we need to measure the risk. Many forms of risk measure have already been proposed and applied in practice. For example, volatility has been used by the finance community for a long time while ruin probability is an important risk measure in actuarial science. Value-at-Risk (VaR) has become a very popular measure of risk over the last decade. However, recently VaR as a risk measure has been criticized for not being sub-additive (therefore VaR is not a coherent risk measure). The first papers to introduce the notion of coherent risk measure and

[a]This work was supported by a grant from Research Grants Council of HKSAR (Project No: HKU 7139/01H)

point out the theoretical problems with VaR were Artzner et al. (1997,1999). Both Vorst (2000) working in discrete time and Basak and Shapiro (2001) working in continuous time found that if managers really do act as if VaR is a binding constraint they will make some bizarre and sub-optimal decisions. There are some other risk measures that have a more solid theoretical base than VaR. One of them is called expected shortfall (ES). Based on our studies of risk measures and risk management, we believe that there is no "best risk measure" for the purpose of risk management in an absolute sense. All risk measures in the literature have their own merits, depending on the role and purpose of the risk measure in the risk management.

This paper is intended to serve as a review of our recent work on risk measures in both actuarial science and finance. We will consider the insurance risk first and will use ruin probability as a risk measure. We will also summarize some of our research results and point out some research problems in this area. In the second part of this paper, we will briefly summarize our recent results on coherent risk measures for derivatives. The third part will focus on the interplay between actuarial science and finance. In particular, we will provide some ideas on how to measure and manage financial risk by using ruin theory techniques.

2 Insurance risk: Ruin probability

For a long period of time, ruin probabilities have been of major interest in mathematical insurance and have been investigated by many authors. The early work on this problem can, at least, track back to Lundberg (1903).

For mathematical simplicity, the models used in risk theory are idealized. The most commonly used model in actuarial science is the compound Poisson model: Let $\{U(t); t \geq 0\}$ denote the surplus process, which measures the surplus of the portfolio at time t, and let $U(0) = u$ be the initial surplus. The surplus at time t can be written as:

$$U(t) = u + pt - X(t), \tag{1}$$

where $p > 0$ is a constant, representing the premium rate, $X(t) = \sum_{i=1}^{N(t)} Y_i$ is the claim process, $\{N(t); t \geq 0\}$ is the number of claims up to time t while the sequence $\{Y_1, Y_2, \cdots\}$ are independent and identically distributed (i.i.d.) variables with the same distribution $F(x)$ which has mean μ and $\{Y_1, Y_2, \cdots\}$ is independent of $\{N(t); t \geq 0\}$ and $N(t)$ is a homogeneous Poisson process with intensity λ.

Another way of modelling the insurance risk process is by using a discrete time model. Let U_t be the surplus at time t and let x be the initial surplus.

Then,

$$U_t = x + \sum_{i=1}^{t} X_i - \sum_{i=1}^{t} Y_i, \tag{2}$$

where $\{Y_1, Y_2, \ldots\}$ is a sequence of independent and identically distributed positive random variables, Y_t denotes the claim size within the time period t, $\{X_1, X_2, \ldots\}$ is another i.i.d. sequence, $\{X_t\}$ and $\{Y_t\}$ are independent, and X_t denotes the premium income within the time period t.

Define

$$\psi(u) = P\{\bigcup_{t\geq 0} \{U(t) < 0\} | U(0) = u\} = P\{T < \infty | U(0) = u\} \tag{3}$$

as the probability of ruin with initial surplus u, where $T = \inf\{t \geq 0 : U(t) < 0\}$ is called the ruin time.

The main classical results about ruin probability for the classical risk model are due to Lundberg (1926) and Cramer (1930). The main problem with ruin theory is how to estimate or calculate the ruin probability. In Yang (1999), by constructing a martingale and using the martingale inequality, some exponential and non-exponential bounds were obtained.

We say a distribution function $B(x)$ is a new worse than used (NWU) distribution, if $B(x)$ is a d.f. of a non-negative random variable and $\bar{B}(x) = 1 - B(x)$, $\bar{B}(x)\bar{B}(y) \leq \bar{B}(x+y)$ for $x \geq 0$ and $y \geq 0$. We say that $B(x)$ is a new better than used (NBU) if $\bar{B}(x)\bar{B}(y) \geq \bar{B}(x+y)$ for $x \geq 0$ and $y \geq 0$.

Lemma Suppose that $B_1(x)$ is a NWU d.f. and $B_2(x)$ is a NBU d.f.

$$E\{\frac{1}{\bar{B}_1(Y)}\bar{B}_2(X)\} \leq 1 \tag{4}$$

and

$$\bar{B}_1(y-x) \geq \bar{B}_1(y)\{\bar{B}_2(x)\}^{-1} \quad \text{for} \quad y \geq x. \tag{5}$$

Then,

$$\psi(x) \leq \bar{B}_1(x). \tag{6}$$

Proof: See Yang (1999).

The following result is a special case of the above Lemma.

Corollary 1 Suppose that $B(x)$ is a NWU d.f. and satisfies:

$$i). \quad \bar{B}(y-x) \geq \bar{B}(y)e^{\mu x} \quad \text{for} \quad y \geq x \tag{7}$$

$$ii). \quad E\{\frac{1}{\bar{B}(Y)}\}E(e^{-\mu X}) \leq 1 \tag{8}$$

The condition $i)$ is true if the failure rate of $\bar{B}$ satisfies $\mu_B(t) \geq \mu$. Then

$$\psi(x) \leq \bar{B}(x). \tag{9}$$

The above classical model does not consider the investment effect. Insurance companies began paying more attention to the investment income about ten years ago. In the past, insurance companies only focused on underwriting income. In fact, large portions of the surplus of insurance companies come from the investment income. If we include a constant interest rate in our model, the model becomes: Let r be the compound interest rate and assume that r is a constant here $(r \geq 0)$. Then,

$$U_t = x(1+r)^t + \sum_{i=1}^{t} X_i(1+r)^{t-i+1} - \sum_{i=1}^{t} Y_i(1+r)^{t-i}, \tag{10}$$

where all the notations are the same as before and we assume that the premium is paid at the beginning of the time period and the claims are paid at the end.

For this model, we define the ruin probability

$$\begin{aligned}\psi(x) &= P\{U_n < 0\,, \text{ for some n}\}\\ &= P\left\{\sum_{i=1}^{n} Y_i(1+r)^{-i} - \sum_{i=1}^{n} X_i(1+r)^{-i+1} > x\,, \text{ for some n}\right\}.\end{aligned}$$

The same results as in the case without interest rate can been obtained. For details, see Yang (1999).

Another important problem in risk theory is the correlated risk. There are two kinds of correlation. One is correlations among different business lines and the other is that where premium and claims are dependent on history. The first kind of correlation is very difficult to deal with. In the case of two different lines of business, the problem can be formulated as follows:

$$\mathbf{U}_n(\mathbf{u}) = \mathbf{u} + \mathbf{c}n - \mathbf{S}_n.$$

Here, $\mathbf{U}_n$, $\mathbf{u}$ and $\mathbf{c}$ are (column) vectors, and $\mathbf{S}_n$ are the aggregate vector-valued claims. Write $S_{in} = \sum_{j=1}^{n} X_{ij}$ and the problem is how to define and estimate the ruin probability when X_{ij} and X_{kj} are correlated.

The continuous version of the above formulation is:

$$\mathrm{U}(t) = \mathbf{u} + \mathbf{c}t - \mathbf{S}_t.$$

Here $\mathbf{U}(t)$, $\mathbf{u}$, $\mathbf{c}$ and $\mathbf{S}(t)$ are (column) vectors. By assuming different forms of $\mathbf{S}(t)$, we introduce different correlated risks. One particular form is assuming that $\mathbf{S}(t) = \sum_{j=1}^{N(t)} \mathbf{X}_j$, where $\mathbf{X}_j$ are vectors and $N(t)$ is a Poisson process with parameter λ. We assume, for fixed i, that $\{X_{ij},\ j = 1, 2, \ldots\}$ are i.i.d r.v's as in the univariate risk model, independent of $N(t)$. (Of course, we can also assume that the X_{ij} are correlated for different j. More generally, we can even assume that different components of $\mathbf{S}(t)$ have different claim number processes and the claim processes are correlated).

In the following, we assume that $EX_{ij} = \mu_i$, $c_i > \lambda\mu_i$, $i = 1, \ldots, n$, and the corresponding relative security loading vector is $(\theta_1, \ldots, \theta_n)'$. A first natural question is how to calculate ruin probability for model (2). The word "*ruin*" in the multivariate case may have different meanings. The following three types of time of ruin are the most interesting:

$$T_1 = \inf\{t \mid \min_{1 \le i \le n}\{U_i(t), i = 1, \ldots, n\} < 0\}, \tag{11}$$

$$T_2 = \inf\{t \mid \max_{1 \le i \le n}\{U_i(t), i = 1, \ldots, n\} < 0\}, \tag{12}$$

$$T_3 = \inf\{t \mid [\sum_{i=1}^{n} U_i(t)] < 0\}. \tag{13}$$

With the time of ruin defined, the corresponding (infinite time) probability of ruin is denoted by $\psi_i(\mathbf{u}) = P[T_i < \infty \mid \mathbf{U}(0) = (u_1,\ \ldots, u_n)'],\quad i = 1, 2, 3.$ Notice that T_1, T_2 and T_3 have obvious interpretations. For instance, $\{T_1 < \infty\}$ means that ruin occurs if at least one of the $\{U_i(t),\ i = 1, \ldots, n\}$ is below zero. Can we obtain Lundberg type bounds for $\psi_i(\mathbf{u})$ (i = 1, 2, 3)?

For another kind of dependency, we use a time series model to model the problem. Suppose that $\{W_1, W_2, \ldots\}$ is a sequence of independent and identically distributed (i.i.d) non-negative random variables. Let the common distribution function of W_i be $G(x) = \Pr(W \le x)$ and $E(W) < +\infty$, where an arbitrary W_i is denoted by W. We assume that $\{X_1, X_2, \ldots\}$ is a sequence of non-negative random variables, and

$$\begin{aligned} X_k &= W_k + bX_{k-1} \qquad k = 1, 2, 3, \ldots, \\ X_0 &= x_0, \end{aligned}$$

where $-1 < b < 1$. Here, X_i denotes the premium collected during the i^{th} year. We assume that the premiums at the beginning of a subsequent year are an upgrade of last year's premium plus a random noise term.

In addition, assume that $\{Y_1, Y_2, \ldots\}$ is a sequence of non-negative random variables, and

$$\begin{aligned} Y_k &= Z_k + aY_{k-1} \qquad k = 1, 2, 3, \ldots \\ Y_0 &= y_0, \end{aligned}$$

where $\{Z_k\}$ is a sequence of i.i.d non-negative random variables, independent of $\{W_1, W_2, \ldots\}$ and $-1 < a < 1$. Here, Y_i denotes the claims during the i^{th} year.

Let the common distribution function of Z_i be $F(x) = \Pr(Z \leq x)$, where an arbitrary Z_i is denoted by Z and assume that $EZ < +\infty$.

Let U_t be the surplus at time t, r be the compound interest rate (we assume that $r > 0$ is a constant here), and x denote the initial surplus. Then,

$$U_t = x(1+r)^t + \sum_{i=1}^{t} X_i (1+r)^{t-i+1} - \sum_{i=1}^{t} Y_i (1+r)^{t-i}.$$

Here, we assume that the claim Y_i is paid at the end of the time period and the premium X_i is paid at the beginning of the time period.

$$\psi(x, y_0, x_0) = P\Big(\bigcup_{t=1}^{\infty} (U_t \leq 0) \,\Big|\, U_0 = x, Y_0 = y_0, X_0 = x_0\Big)$$

is the ruin probability for this model.

Using exactly the same approach as that in Yang (1999), we can obtain both exponential and non-exponential bounds for the ruin probability. For details, see Yang and Zhang (2001).

3 Financial Risk: Coherent Risk Measures for Derivative Securities

Risk measurement for derivatives is an important issue in the practice of total risk management. Traditional methods of measuring risks of derivatives rely solely on the Greek letters, namely delta, gamma and rho, etc. Over the last decade, VaR has become a popular tool for measuring risks of derivatives. Developed by J.P. Morgan's Risk-Metric Group, VaR attempts to summarize the total risk of a portfolio by a single number which is a statistical estimation of a portfolio's loss with the property that, with a given (small) probability level, the owner of the portfolio stands to incur that loss or more over a given (typically short) holding period of the portfolio. Basically, there are two common analytical approaches to calculate VaR for derivatives. Both approaches are difficult to implement if a portfolio consists of a significant number of derivatives. Perhaps the Monte Carlo simulation and its cousin, the Quasi-Monte Carlo method, are effective numerical methods to calculate VaR for derivatives. For an overview of the calculation of VaR for derivatives, see Duffie and Pan (1997). Artzner et al. (1997), (1999) proposed a set of four desirable properties for risk measures, namely translation invariance, positive homogeneity, monotonicity and subadditivity. A risk measure satisfying those properties is called a coherent risk measure. It has been pointed out in Artzner et al. (1997), (1999) that VaR does not, in general, satisfy the subadditivity property, especially when the portfolio contains derivatives which are non-linear in nature. This makes the problem of investigating other more conceptually consistent risk measures for derivatives more interesting. Besides VaR, there are two other popular risk measures in the finance and actuarial science literature, namely Expected Shortfall (ES) and Generalized "Scenarios" Expectation (GSE). Artzner et al. (1997), (1999), Acerbi and Tasche (2001) and Yamai and Yoshiba (2001) proposed the use of ES as a more theoretically consistent risk measure instead of VaR. Acerbi and Tasche (2001) proved the subadditivity property (hence coherence) of ES. Loosely speaking, ES is defined as the average loss of a portfolio in the worst $100\alpha\%$ cases, for some (small) probability level α. See Acerbi and Tasche (2001) for a formal definition of ES. For detailed discussions about the practical issues of ES, see Yamai and Yoshiba (2001). GSE is defined as the supremum of the expected loss of a portfolio over a set of probability measures or generalized "scenarios". It can be considered as a generalization of the margin system SPAN (Standard Portfolio Analysis of Risk), developed by the Chicago Mercantile Exchange. In Artzner et al. (1999), it has been shown that the representation form of coherent risk measures is given by GSE.

In Yang and Siu (2001), by following the representation form of coherent risk measures introduced by Artzner et al. (1999), we introduce a scenario-based risk measure which involves the use of the risk-neutral probability ($\mathcal{Q}$-measure) , the physical probability ($\mathcal{P}$-measure), and a family of subjective probability measures. Here, the physical probability $\mathcal{P}$, which is also called the statistical/data-generating probability, is the underlying probability law that drives the realization of the stock-price movement. It is objective and unique. In practice, the underlying probability law is not known but we can estimate it through the use of some statistical techniques. A subjective probability measure is assigned according to an agent's subjective beliefs and risk preference. Its assignment needs not be subjected to a general agreement.

Siu and Yang (2001) proposes a partial differential equation (P.D.E.) approach for calculating coherent risk measures for portfolios of derivatives under the Black-Scholes economy. The P.D.E. approach provides a natural extension of Yang and Siu (2001) and an alternative method for implementing the risk measures in Yang and Siu (2001). It enables us to define the risk measures in a dynamic way and to deal with American options in a relatively effective way.

Siu, Tong and Yang (2001) proposes a model for measuring risks for derivatives. We construct our model within the context of Gerber-Shiu's option-pricing framework (see Gerber and Shiu (1994)). A new concept, namely Bayesian Esscher "scenarios," which generalizes the concept of generalized "scenarios," is introduced via a "Random Esscher Transform."

4 Credit Risk: Actuarial Science Approach

In this section, we consider a firm which could either be a financial corporation or an insurance company. At the beginning of each time interval, a rating agency will provide a credit rating to assess the firm's abilities in meeting its debt obligations (to pay possible claims in an insurance company case). We use a Markov Chain to model the dynamics of the firm's credit ratings. It is a modification of the Jarrow, Lando and Turnbull (JLT) (1997) model. The only difference between our model and the JLT model is that we only consider the non default rating states.

Let I_t be a time-homogeneous Markov Chain with a state space of $N = \{1, 2, \ldots, k\}$, where state 1 represents the highest credit class, and state k represents the lowest. In Moody's ratings, state 1 can be thought of as Aaa and state k as Caa, and in $S\&P$'s, state 1 as AAA and state k as CCC.

Let

$$q_{ij} = P\{I_{t+1} = j \,|\, I_t = i\} , \quad i, j \in N , \quad t = 0, 1, 2, \ldots \tag{14}$$

be the one-step transition probabilities. The transition matrix of the Markov Chain I_t can then be written as

$$Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1k} \\ q_{21} & q_{22} & \cdots & q_{2k} \\ \vdots & & & \\ q_{k1} & q_{k2} & \cdots & q_{kk} \end{bmatrix} . \tag{15}$$

Let u be the initial surplus of the firm, and X_n^i be the portfolio change in the n^{th} time interval if the firm's credit rating in time interval n is of class i. The surplus of the firm at time n can then be written as

$$U_n = u + \sum_{m=1}^{n} X_m^{I_{m-1}} , \tag{16}$$

where we assume that X_m^i $i = 1, \ldots, k$, $m = 1, 2, \ldots$ are independent random variables. We say that ruin or default occurred at time n if $U_n \leq 0$.

Let $T = \inf\{n; U_n \leq 0\}$. The stopping time, T, is called the default time. The probability of default before or at time n is defined as

$$\psi_{ni_0}(u) = P\{T \leq n \,|\, I_0 = i_0, U_0 = u\}. \tag{17}$$

Under this setup, in Yang (2001) a recursive formula for the finite time default probability and default time distribution are obtained. A coupled Volterra type integral equation system is obtained for the ultimate default probability. Yang (2001) also discussed the severity of the default. An integrated risk management method was developed based on this model in Yang (2000).

References

1. Acerbi, C. and Tasche, D. (2001) "On the Coherence of Expected Shortfall", Preprint.

2. Artzner, P., Delbaen, F., Eber, J., and Heath, D. (1997) "Thinking Coherently," *Risk*, 10(11), 68-71.

3. Artzner, P., Delbaen, F., Eber, J., and Heath, D. 1999. "Coherent Measures of Risk," *Mathematical Finance*, 9(3), 203-228.

4. Basak, S. and Shapiro, A. (2001) "Value-at-Risk-Based Risk Management: Optimal Policies and Asset Prices", *The Review of Financial Studies.* **14(2)**, 371-405.

5. Cramer, H. (1930) "On the Mathematical Theory of Risk", *Skandia Jubilee Volume*, Stockholm.

6. Duffie, D. and Pan, J. (1997) "An Overview of Value at Risk", *Journal of Derivatives.* **4(3),** 7-49.

7. Gerber, H. U. and Shiu, E. S. W. (1994) "Option pricing by Esscher transforms", *Transactions of the Society of Actuaries*, **XLVI**, p. 99-191.

8. Holton, G. (1997) "Subjective Value at Risk," *Financial Engineering News*, 1(1), 1, 8-9, 11.

9. Jarrow, R. A., Lando, D. and Turnbull, S.M. (1997) "A Markov model for the term structure of credit risk spread," *Review of Financial Studies*, **10**, p. 481-523.

10. Lundberg, F. (1903) *I. Approximerad Framstallning af Sannolikhetsfunktionen. II. Aterforsakring af Kollektivrisker*, Almqvist and Wiksells, Uppsala.

11. Lundberg, F. (1926) *Forsakringsteknisk Riskutjamning*, F. Englunds boktryckeri A.B., Stockholm.

12. Siu, T.K., and Yang, H. (2001) "A P.D.E. Approach for Risk Measures of Derivatives," *Applied Mathematical Finance*, Accepted.

13. Siu, T.K., Tong, H. and Yang, H. (2001) "Risk Measure for Derivatves by Esscher Transform", *North American Actuarial Journal,* Vol. 5, No. 3, p. 78-91.

14. Steinherr, A., (1988) *Derivatives: The Wild Beast of Finance*, Wiley, New York.

15. Vorst T., (2000) "Optimal Portfolios under a Value at Risk Constraint", Working paper, Erasmus University Rotterdam.

16. Yamai, Y. and Yoshiba, T. (2001) "On the validity of value at risk: comparative analyses with expected shortfall", discussion paper 2001-E-4, Institute for Monetary and Economic Studies, Bank of Japan.

17. Yang, H. (1999) "Non-exponential bounds for ruin probability with interest effect included", *Scandinavian Actuarial Journal*, p.66-79.

18. Yang, H. (2001) "Ruin theory in a financial corporation model with credit risk". Submitted.

19. Yang, H. (2000) "An Integrated Risk Management Method: VaR Approach", *Multinational Finance Journal*, Vol. 4, No. 3 & 4, p. 201-219.

20. Yang, H., and Siu, T. K. (2001) "Coherent Risk Measures for Derivatives Under Black-Scholes Economy," *International Journal of Theoretical and Applied Finance*, 4(5).

21. Yang, H. and Zhang, L. (2001) "Upper Bounds for Ruin Probability in an Autoregressive Model with Constant Interest Rate". Submitted.

Using Stochastic Approximation Algorithms in Stock Liquidation

G. Yin
Department of Mathematics, Wayne State University, Detroit, MI 48202, gyin@math.wayne.edu

Q. Zhang
Department of Mathematics, University of Georgia, Athens, GA 30602, qingz@math.uga.edu

R.H. Liu
Department of Mathematics, University of Georgia, Athens, GA 30602, rliu@alpha.math.uga.edu

Abstract

For hybrid geometric Brownian motion stock liquidation models, it has been proved that the optimal selling policy is of threshold type, which can be obtained by solving a set of two-point boundary value problems. The total number of equations to be solved is the same as that of the numbers of states of the underlying Markov chain. To reduce the computational burden, this work develops Monte Carlo algorithms, which are recursive and are stochastic optimization type, to approximate the optimal threshold values in stock trading. Then asymptotic properties of the proposed algorithms such as the convergence and rates of convergence are developed.

Key words. geometric Brownian motion, hybrid model, stochastic optimization, recursive algorithm

Mathematics Subject Classification. 90A09, 60J10, 60J27, 62L20.

1 Introduction

In this paper, we consider liquidation decision making for a single non-dividend stock. We focus on the case that a selling rule is determined by two threshold levels, a target price and a stop-loss limit. One makes a selling decision whenever the price reaches either the target price or the stop-loss limit.

In finance literature, the celebrated Black-Scholes model based on geometric Brownian motion (GBM) is widely used in the analysis of options pricing and portfolio management; see Merton [9] among others. This model uses a stochastic differential equation with deterministic expected returns and volatility and gives reasonably good description of the market in a short period of time. However, the model has limitation because it cannot catch random parameter changes. In reality, the stock price movements are far from being a "random walk" in longtime horizon. Recognizing the needs, modifications that incorporate the random influence of the parameter changes have been made; see for example, [2,4,5,6,10] and the references therein. In particular, studies on volatility parameter dictated by additional stochastic differential equations can be found in Fouque, Papanicolaou, and Ronnie[3], Hull[4], and Musiela and Rutkowski[10] among others.

One of the factors that affects decision making in a marketplace significantly is the trend of the stock market. Frequently, the systems under consideration are time varying and are associated with movements involving discontinuity influenced by uncertain and exogenous discrete events. As a result, the configurations or parameters of the dynamic systems come from one of several different regimes with transitions among regimes governed by an unobservable hidden process. The behavior of the corresponding dynamic systems at different regimes could be markedly different. Therefore, in contrast to the Black-Scholes model, a promising alternative is to allow for the possibility of sudden, discrete changes in the values of the parameters resulting in a "hybrid" or "switching" Black-Scholes model. In a recent paper of Zhang[14], a hybrid switching GBM model, i.e., a number of GBMs modulated by a finite-state Markov chain, is proposed and developed. Such switching process can be used to represent market trends or the trends of an individual stock. In addition, various economic factors such as interest rates, business cycles etc., which can be modeled by use of a continuous-time Markov chain, could also be incorporated in the model.

For such a hybrid model, the optimal threshold levels are obtained by solving a set of two-point boundary value problems. If the underlying Markov chain has only two states, then the corresponding two-point boundary value problem has an analytical solution and the optimal solution can be obtained in closed-form. However, more detailed market study are often necessary that requires the underlying Markov chain having more than two states ($m > 2$). In this case, the computation becomes much more involved because a set of two-boundary value problems must be solved. In such a setup, a closed-form or analytic solution is difficult to obtain although general existence has been proved[14]. It is thus of practical concern, to develop systematic and efficient algorithms leading to approximation of the optimal policy.

Motivated by the reduction of computational effort, we suggest and develop an alternative approach in this paper. Focusing our attention on threshold selling rules, in lieu of solving a set of boundary value problems, we formulate the problem as a stochastic optimization procedure and design a Monte

Carlo scheme for resolution, which enables us to design a recursive algorithm of stochastic approximation type; for a most up-to-date account of stochastic approximation algorithms, see Kushner and Yin [7]. Although stochastic control techniques have been employed in various finance problems, optimization using stochastic approximation methods for stock liquidation decision making appears to be new to the best of our knowledge.

The rest of the paper is arranged as follows. Section 2 begins with the formulation of the problem. We first present the model and then we propose a recursive algorithm. Regarding the objective of maximizing the expected return as a function of the threshold values, gradient estimates of the objective function is provided via finite difference methods. One of the advantages is the simple form and systematic nature of the algorithm. It is particularly useful for on-line computation. Section 3 then proceeds with the study of the asymptotic properties of the underlying algorithm. Using weak convergence methods, we obtain the convergence of the algorithm and ascertains the convergence rate. These results have also been verified by simulation studies and numerical experiments using real market data [13]. We remark that other factors affecting liquidation decisions such as transaction costs (commissions) and capital gain taxes etc. (see [1] and the references therein) may also be built into the model. In evaluating the rates of convergence, one could incorporate the information of bias and the computational budget as was done in [8]. Finally, variations of the algorithm including random directions finite difference gradient estimates, gradient estimates using sample a single sample path and using averaging of the observations, and projection and truncation procedures can also be considered.

Throughout the paper, we use K to denote a generic positive constant whose values may be different for different usage. For a suitable function g, g_x and g_{xx} denote the gradient and Hessian of g, respectively. For any $z \in \mathbb{R}^{\ell \times \iota}$ with some $\ell, \iota \geq 1$, z' denotes its transpose. For a vector v, v^i denotes its ith component.

2 Formulation

Let $\alpha(t)$ be a finite-state, continuous-time Markov chain having state space $\mathcal{M} = \{1, \ldots, m\}$ (see Chapter 2 of [12]), which represents market trends and other economic factors. For example, when $m = 2$, $\alpha(t) = 1$ and $\alpha(t) = 2$ describe the up and down trends, respectively. Since the actual market may include more complex scenarios, we assume that $\alpha(t)$ has m states with $m \geq 2$.

Suppose that $S(t)$ is the price of the stock satisfying

$$\begin{aligned} &\frac{dS(t)}{S(t)} = \mu(\alpha(t))dt + \sigma(\alpha(t))dw(t), \\ &S(0) = S_0 \ \text{ initial price,} \end{aligned} \tag{2.1}$$

where $w(\cdot)$ is a standard Brownian motion that is independent of $\alpha(\cdot)$. The model can be viewed as a hybrid or switching Black-Scholes model.

Note that in (2.1), both the drift and the diffusion depend on the jump process $\alpha(t)$. Set

$$X(t) = \int_0^t r(\alpha(s))ds + \int_0^t \sigma(\alpha(s))dw(s), \tag{2.2}$$

with

$$r(i) = \mu(i) - \frac{\sigma^2(i)}{2} \ \text{ for each } \ i = 1, \ldots, m. \tag{2.3}$$

We can then rewrite the solution of (2.1) as

$$S(t) = S_0 \exp(X(t)). \tag{2.4}$$

Let $z = (z^1, z^2)' \in (0, \infty) \times (0, \infty)$. We consider two boundaries of the threshold, a lower boundary $z^1 > 0$ and an upper boundary $z^2 > 0$ such that whenever the stock price reaches the upper bound $S_0 \exp(z^2)$ or the lower bound $S_0 \exp(-z^1)$, sell the stock to take profit or to prevent from further loss.

Focusing on threshold type policies, we formulate the underline problem as an optimization procedure. Let τ be the first time that the price hits the boundaries. That is,

$$\tau = \inf\{t > 0 : \ S(t) \notin (S_0 \exp(-z^1), S_0 \exp(z^2))\},$$

or

$$\tau = \inf\{t > 0 : X(t) \notin (-z^1, z^2)\}. \tag{2.5}$$

Our objective is to find the optimal threshold level $z_* = (z_*^1, z_*^2)' \in (0, \infty) \times (0, \infty)$ so that the expected return is maximized. We rewrite the problem as:

$$\text{Problem } \mathcal{P} : \begin{cases} \text{Find argmax } \varphi(z), \\ \varphi(z) = E[\Phi(X(\tau)) \exp(-\widetilde{\varrho}\tau)], \\ z = (z^1, z^2)' \in (0, \infty) \times (0, \infty), \end{cases} \tag{2.6}$$

where $\Phi(\cdot)$ is a suitable real-valued function, and $\widetilde{\varrho} > 0$ is the discount rate.

Although when $m = 2$, an analytic solution can be found, in general, a closed-form solution may be virtually impossible to obtain. Even in the case of $m = 2$, the computation for obtaining the closed-form solutions is a nontrivial task since it involves solutions of two-point boundary value problems.

As an alternative, we design a Monte Carlo stochastic optimization procedure to resolve the problem by constructing a sequence of estimates of the optimal threshold value z_* using

$$z_{n+1} = z_n + \{\text{step size}\} \cdot \{\text{gradient estimate of } \varphi(z)\},$$

where the step size can be either a decreasing sequence of real numbers or a small positive constant.

The approximation procedures will depend on how the gradient estimates of $\varphi_z(z)$ are constructed. Using (2.1), generate a sample path $X(t)$ that is the solution of (2.2). At time n (n being the iteration number), with the threshold value fixed at $z_n = (z_n^1, z_n^2)' \in (0,\infty) \times (0,\infty)$, we compute τ_n the first exit time of $X(t)$ from $I_{z_n} = (-z_n^1, z_n^2)'$ (the interval with the lower and upper boundaries set at $-z_n^1$ and z_n^2, respectively) by

$$\tau_n = \inf\{t > 0 : \zeta(t) \notin I_{z_n}\}. \tag{2.7}$$

Define a combined process ξ_n that includes the random effects from $X(t)$ and the stopping time τ_n as

$$\xi_n = (X(\tau_n), \tau_n)', \tag{2.8}$$

where $X(\tau_n)$ denotes the random process $X(t)$ stopped at τ_n. Henceforth, call $\{\xi_n\}$ the sequence of collective noise.

Let $\widetilde{\varphi}(\cdot)$ and $\widehat{\varphi}(\cdot)$ be real-valued functions defined on $\mathbb{R}^2 \times \mathbb{R}^1$. When the threshold value is set at z, take random samples of size n_0 with random noise sequences $\{\xi_{n,\ell}^{\pm}\}_{\ell=1}^{n_0}$ such that

$$\widehat{\varphi}(z, \xi_n^{\pm}) \stackrel{\text{def}}{=} \frac{\widetilde{\varphi}(z, \xi_{n,1}^{\pm}) + \cdots + \widetilde{\varphi}(z, \xi_{n,n_0}^{\pm})}{n_0}. \tag{2.9}$$

Note that frequently independent random samples are used. To allow a slightly more flexibility, we will not assume the independence here. However, assume that it is stationary and

$$E\widehat{\varphi}(z, \xi_n^{\pm}) = \varphi(z) \quad \text{for each } z. \tag{2.10}$$

Then, for each z, $\widehat{\varphi}(z, \xi_n^{\pm})$ is an estimator of $\varphi(z)$. In our simulation study, using $\widetilde{\varphi}(z, \xi_n^{\pm})$ with independent random samples to estimate the mean of

$\Phi(X(\tau_n))\exp(-\widetilde{\varrho}\tau_n)$, by the well-known law of large numbers $\widehat{\varphi}(z,\xi_n)$ converges to $\varphi(z)$ w.p.1 as $n_0 \to \infty$. In what follows, in lieu of using (2.9) with $\widetilde{\varphi}(z,\xi^{\pm}_{n,\ell})$, we will work with the form $\widehat{\varphi}(z,\xi_n)$, give conditions needed for obtaining convergence and rate of convergence, and derive the asymptotic properties of the underlying algorithms. We shall also validate the conditions under independence assumption of $\{\xi^{\pm}_{n,\ell}\}$.

Consider a stochastic approximation procedure with finite-difference-type gradient estimates. Use $Y^{\pm}_n = (Y^{\pm,1}_n, Y^{\pm,2}_n)' \in \mathbb{R}^2$ to denote the outcomes of two simulation runs taken at the nth iteration, where $Y^{\pm,\iota}_n = Y^{\pm,\iota}(z_n,\xi^{\pm}_n)$ with

$$\begin{aligned} Y^{+,\iota}_n(z,\xi^+) &= \widehat{\varphi}(z+\delta_n e_\iota,\xi^+), \quad \text{for} \quad \iota = 1,2, \\ Y^{-,\iota}_n(z,\xi^-) &= \widehat{\varphi}(z-\delta_n e_\iota,\xi^-), \quad \text{for} \quad \iota = 1,2, \end{aligned} \tag{2.11}$$

e_ι being the standard unit vectors $e_1 = (1,0)'$ and $e_2 = (0,1)'$, and $\xi^{\pm}_n$ being two different (collective) noises taken at the threshold values $z \pm \delta_n e_\iota$, respectively. [For notational simplicity, we have used ξ_n to represent both ξ^+_n and ξ^-_n henceforth.] The gradient estimate is given by

$$D\widehat{\varphi}(z_n,\xi_n) \stackrel{\text{def}}{=} \frac{Y^+_n - Y^-_n}{2\delta_n}.$$

A stochastic optimization algorithm then takes the form

$$z_{n+1} = z_n + \varepsilon_n D\widehat{\varphi}(z_n,\xi_n), \tag{2.12}$$

where $\{\varepsilon_n\}$ is a sequence of real numbers known as step sizes.

To proceed, define

$$\begin{aligned} \rho_n &= (Y^+_n - Y^-_n) - E_n(Y^+_n - Y^-_n), \\ \chi^\iota_n &= [E_n Y^{+,\iota}_n - \varphi(z_n+\delta_n e_\iota)] - [E_n Y^{-,\iota}_n - \varphi(z_n - \delta_n e_\iota)], \quad \iota = 1,2, \\ b^\iota_n &= \frac{\varphi(z_n+\delta_n e_\iota) - \varphi(z_n-\delta_n e_\iota)}{2\delta_n} - \varphi_{z^\iota}(z_n), \quad \iota = 1,2, \end{aligned} \tag{2.13}$$

where E_n denotes the conditional expectation with respect to $\mathcal{F}_n$, the σ-algebra generated by $\{z_1, \xi^{\pm}_j : j < n\}$, $\varphi_{z^\iota}(z) = (\partial/\partial z^\iota)\varphi(z)$, and $\varphi_z(\cdot) = (\varphi_{z^1}(\cdot), \varphi_{z^2}(\cdot))'$ denotes the gradient of $\varphi(\cdot)$. In what follows, when we want to emphsize the dependence on (z,ξ), we spell it out, for example, we use the notation $\chi^\iota(z,\xi)$ similar as that of $Y^{\pm,\iota}_n(z,\xi^{\pm})$ defined in (2.11). Write $\chi_n = (\chi^1_n, \chi^2_n)'$ and $b_n = (b^1_n, b^2_n)'$ and note that $\chi_n = \chi_n(z_n,\xi_n)$, which will be used in what follows. With the noise $\chi_n(z_n,\xi_n)$ and the bias b_n defined above, algorithm (2.12) becomes

$$z_{n+1} = z_n + \varepsilon_n \varphi_z(z_n) + \varepsilon_n \frac{\rho_n}{2\delta_n} + \varepsilon_n b_n + \varepsilon_n \frac{\chi_n(z_n, \xi_n)}{2\delta_n}. \tag{2.14}$$

3 Convergence and Rates of Convergence

We will study the algorithm as a stochastic approximation problem in general and provide a set of conditions, under which we derive the asymptotic properties. The main assumptions and results are presented with the detailed proofs being relegated to [13]. It should be mentioned that the conditions posed turn out be all verifiable when Monte Carlo methods and i.i.d. samples are used (see the simulation study presented in the aforementioned paper).

To study the convergence of the underlying algorithm, in lieu of dealing with the discrete iterations, we take a continuous-time interpolation leading to a mean ordinary differential equation (ODE). The stationary points of the ODE are the threshold values that we are searching for. Then the rate of convergence is studied via an appropriate scaling. We show that a suitably scaled sequence of the estimation errors converges weakly to a diffusion process. The scaling factor together with the asymptotic covariance of the limit diffusion gives us the desired rates of convergence. In what follows, we state the main results; the detailed proofs can be found in our recent work [13].

Define

$$\begin{cases} t_n = \sum_{i=1}^{n-1} \varepsilon_i, \quad m(t) = \max\{n : t_n \le t\}, \\ N_n = \min\{i : t_{n+i} - t_n \ge T\}, \quad \text{for an arbitrary } \; T > 0, \\ z^0(t) = z_n \quad \text{for} \quad t \in [t_n, t_{n+1}), \\ z^n(t) = z^0(t + t_n). \end{cases} \tag{3.1}$$

Note that $z^0(\cdot)$ is a piecewise constant process and $z^n(\cdot)$ is its shift whose purpose is to bring the asymptotics to the foreground. It follows that the interpolated process $z^n(\cdot)$ can be written as

$$\begin{aligned} z^n(t) = z_n &+ \sum_{j=n}^{m(t_n+t)-1} \varepsilon_j \varphi_z(z_j) + \sum_{j=n}^{m(t_n+t)-1} \varepsilon_j \frac{\rho_j}{2\delta_j} \\ &+ \sum_{j=n}^{m(t_n+t)-1} \varepsilon_j b_j + \sum_{j=n}^{m(t_n+t)-1} \frac{\varepsilon_j}{2\delta_j} \chi_j(z_j, \xi_j). \end{aligned} \tag{3.2}$$

Note that $z^n(\cdot) \in D^2[0,\infty)$ the space of $\mathbb{R}^2$-valued functions that are right continuous and have left-hand limits, endowed with the Skorohod topology [7]. We need the following assumptions.

(A0) The step-size sequence $\{\varepsilon_n\}$ and the finite difference step-size sequence $\{\delta_n\}$ satisfy $0 < \varepsilon_n \to 0$, $\sum_n \varepsilon_n = \infty$, $0 < \delta_n \to 0$, and $\varepsilon_n/\delta_n^2 \to 0$ as $n \to \infty$. Moreover, $\limsup_n \sup_{0 \le i < N_n}(\varepsilon_{n+i}/\varepsilon_n) < \infty$, $\limsup_n (\delta_{n+i}/\delta_n) < \infty$, $\limsup_n \left[(\varepsilon_{n+i}/\delta_{n+i}^2)/(\varepsilon_n/\delta_n^2)\right] < \infty$.

(A1) The second derivative $\varphi_{zz}(\cdot)$ is continuous.

(A2) For each compact set G,

(a) $\sup_n E|Y_n^{\pm} I_{\{z_n \in G\}}|^2 < \infty$.

(b) For each z belonging to a bounded set,

$$\begin{aligned} &\sup_n \sum_{j=n}^{n+N_n-1} E^{1/2} \left|E_n \chi_j(z,\xi_j)\right|^2 < \infty, \\ &\lim_n \sup_{0 \le i < N_n} E|\widetilde{\gamma}_i^n| = 0, \end{aligned} \tag{3.3}$$

where

$$\widetilde{\gamma}_i^n = \frac{1}{\varepsilon_{n+i}} \sum_{j=n+i}^{n+N_n-1} \frac{\varepsilon_j}{2\delta_j} E_{n+i}[\chi_j(z_{n+i+1},\xi_j) - \chi_j(z_{n+i},\xi_j)], \quad i < N_n.$$

Theorem 3.1. *Assume* (A0)–(A2) *and* $\{z_n\}$ *is tight in* $(0,\infty) \times (0,\infty)$. *Suppose the differential equation*

$$\dot{z} = \varphi_z(z) \tag{3.4}$$

has a unique solution for each initial condition. Then $z^n(\cdot)$ *converges weakly to* $z(\cdot)$, *the solution of* (3.4).

Suppose that in addition, (3.4) *has a unique stationary point* z_* *that is globally asymptotically stable in the sense of Liapunov. Then* $z_n \to z_*$ *in probability and* $z^n(\cdot)$ *converges weakly to* z_* *as* $n \to \infty$.

Next, the rate of convergence question is studied through a suitably scaled sequence $n^{\kappa_0}(z_n - z_*)$ of the estimation errors, where $\kappa_0 > 0$. Taking $\varepsilon_n = 1/n^{\kappa_1}$ and $\delta_n = \delta/n^{\kappa_2}$ for some $0 < \kappa_2 < \kappa_1 \le 1$ and $\delta > 0$. Then it is known that the optimal choice is given by $\kappa_0 + \kappa_2 = \kappa_1/2$ and $\kappa_0 = 2\kappa_2$. To be more specific, we take $\varepsilon_n = 1/n$. Then $\delta_n = \delta/n^{1/6}$ and $\kappa_0 = 1/3$. Define $u_n = n^{1/3}(z_n - z_*)$ and assume that the following conditions hold.

(A3) Assume (A1) and (A2) hold, $z_n \to z_*$ in probability, and $\varphi_{zzz}(\cdot)$ exists and is continuous in a neighborhood of z_*. In addition,

(a) $\{u_n\}$ is tight;

(b) The matrix $\varphi_{zz}(z_*)+(1/3)I$ is stable, i.e., all of its eigenvalues have negative real parts;

(c) for each z,

$$\begin{aligned}\chi_n(z,\xi) = {} & \chi_n(z_*,\xi) + \chi_{n,z}(z_*,\xi)(z-z_*) \\ & + \left[\int_0^1 [\chi_{n,z}(z_* + (z_n - z_*)s,\xi) - \chi_{n,z}(z_*,\xi)]ds\right](z-z_*);\end{aligned}$$

(d) $\{\chi_n(z_*,\xi_n)\}$ is stationary φ-mixing with $E|\chi_n(z_*,\xi_n)|^{2+\Delta} < \infty$ for some $\Delta > 0$ and $E\chi_n(z_*,\xi_n) = 0$ and that the mixing measure $\varpi(\cdot)$ is given by $\varpi(j) = \sup_{A\in\mathcal{F}^{n+j}} E^{\frac{1+\Delta}{2+\Delta}}|P(A|\mathcal{F}_n) - P(A)|^{\frac{2+\Delta}{1+\Delta}}$, with $\sum_{j=1}^{\infty}(\varpi(j))^{\frac{\Delta}{1+\Delta}} < \infty$.

Using (A3), it can be shown that $\sum_{j=n}^{m(t_n+t)-1}\frac{1}{\sqrt{j}}(\chi_j(z_*,\xi_j)+\rho_j^*)$ converges weakly to a Brownian motion with covariance Σt. Take a piecewise constant interpolation

$$u^0(t) = u_n, \quad t \in [t_n, t_{n+1}), \quad \text{and} \quad u^n(t) = u^0(t_n + t).$$

Theorem 3.2. *Assume that* (A3) *holds. Then $u^n(\cdot)$ converges weakly to a diffusion process $u(\cdot)$ that is a solution of the stochastic differential equation*

$$du = \left\{\left(\varphi_{zz}(z_*) + \frac{I}{3}\right)u + \frac{\delta^2}{3!}\begin{pmatrix}\varphi_{z^1,z^1,z^1}(z_*)\\ \varphi_{z^2,z^2,z^2}(z_*)\end{pmatrix}\right\}dt + \frac{\Sigma^{1/2}}{2\delta}dw, \tag{3.5}$$

where $w(\cdot)$ is a standard Brownian motion and $\Sigma^{1/2}(\Sigma^{1/2})'t = \Sigma t$ is the covariance.

Since (3.5) is linear, it has a unique solution for each initial condition. Note that (3.5) includes a nonzero bias $(\delta^2/3!)\begin{pmatrix}\varphi_{z^1,z^1,z^1}(z_*)\\ \varphi_{z^2,z^2,z^2}(z_*)\end{pmatrix}$. As a direct consequence of Theorem 3.2, $n^{1/3}(z_n - z_*)$ is asymptotically normally distributed with mean $\left(\varphi_{zz}(z_*) + \frac{I}{3}\right)^{-1}\frac{\delta^2}{3!}\begin{pmatrix}\varphi_{z^1,z^1,z^1}(z_*)\\ \varphi_{z^2,z^2,z^2}(z_*)\end{pmatrix}$ and asymptotic covariance $\widetilde{\Sigma} = \int_0^\infty \exp\left(\left(\varphi_{zz}(z_*) + \frac{I}{3}\right)t\right)\Sigma\exp\left(\left(\varphi'_{zz}(z_*) + \frac{I}{3}\right)t\right)dt$. Note that due to (A3) (b), the integral above is well defined.

If in lieu of $\varepsilon_n = 1/n$, we use $\varepsilon_n = 1/n^{\kappa_1}$ with $\kappa_1 < 1$, then the limit differential equation becomes

$$du = \left\{ \varphi_{zz}(z_*)u + \frac{\delta^2}{3!} \begin{pmatrix} \varphi_{z^1,z^1,z^1}(z_*) \\ \varphi_{z^2,z^2,z^2}(z_*) \end{pmatrix} \right\} dt + \frac{\Sigma^{1/2}}{2\delta} dw.$$

In this case, assuming that $\varphi_{zz}(z_*)$ is stable, we have $n^{\kappa_0}(z_n - z_*)$ is asymptotically normal with a mean equal to the bias and asymptotic covariance given by $\widetilde{\Sigma} = \int_0^\infty \exp(\varphi_{zz}(z_*)t)\Sigma \exp(\varphi'_{zz}(z_*)t)dt$.

Acknowledgement

The research of G. Yin was supported in part by the National Science Foundation under grants DMS-9877090; the research of Q. Zhang and R.H. Liu was supported in part by USAF Grant F30602-99-2-0548.

References

1. A. Cadenillas and S. R. Pliska, Optimal trading of a security when there are taxes and transaction costs, *Finance & Stochastics*, vol. 3, pp. 137-165, (1999).
2. R. J. Elliott and P. E. Kopp, *Mathematics of Financial Markets*, Springer-Verlag, New York, 1998.
3. J.P. Fouque, G. Papanicolaou, and S.K. Ronnie, *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press, 2000.
4. J.C. Hull, *Options, Futures, and Other Derivatives*, 3rd Ed., Prentice-Hall, Upper Saddle River, NJ, 1997.
5. I. Karatzas, *Lectures on the Mathematics of Finance*, American Mathematical Society, Providence, RI, 1996.
6. I. Karatzas and S. E. Shreve, *Methods of Mathematical Finance*, Springer, New York, 1998.
7. H.J. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*, Springer-Verlag, New York, 1997.
8. P. L'Ecuyer and G. Yin, Budget-dependent convergence rate of stochastic approximation, *SIAM J. Optim.* **8** (1998), 217-247.
9. R. C. Merton, Lifetime portfolio selection under uncertainty: The continuous-time case, *Rev. Econom. Statist.*, vol. 51, pp. 247-257, (1969).
10. M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modelling*, Springer-Verlag, New York, 1997.

11. J. C. Spall, Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, *IEEE Trans. Automat. Control* **AC-37** (1992), 331-341.
12. G. Yin and Q. Zhang, *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach*, Springer-Verlag, New York, 1998.
13. G. Yin, Q. Zhang, and R.H. Liu, Recursive algorithms for stock liquidation: A Monte Carlo approach, preprint, 2001.
14. Q. Zhang, Stock trading: An optimal selling rule, *SIAM J. Control. Optim.*, **40** (2001), 64-87.

Contingent Claims in an Illiquid Market [a]

Hong Liu
Olin School of Business, Washington University
St. Louis, MO 63130, USA
email: liuh@olin.wustl.edu

Jiongmin Yong [b]
Department of Mathematics, Fudan University
Shanghai 200433, China
email: jyong@fudan.edu.cn

Abstract

Consistent with trading in an illiquid market, we assume that a large trader drives up the stock price as she buys and pushes it down as she sells. The effect of this price impact on the replication of a European contingent claim for the large trader is considered and a generalized non-linear Black-Scholes pricing partial differential equation for computing this cost is obtained. The pricing PDE indicates that one of the main effects of the price impact is the resulting endogenous stochastic volatility for the stock return. The existence and uniqueness of a classical solution to such an equation under certain conditions is established. This implies that the large trader can still perfectly replicate the contingent claim (but with a higher cost). The replicating strategy involves an initial discrete trade followed by continuous trading. It turns out that unlike in the presence of transaction costs, super-replication in the presence of price impact incurs larger costs than replication. Compared to the case without price impact, the large trader generally buys more the stock and borrows more (shorts and lends more) to replicate an out-of-the-money call (put), but buys less the stock and borrows less (shorts and lends less) to replicate an in-the-money call (put).

[a] We thank Kerry Back for extremely helpful suggestions.
[b] This author is also affiliated with Laboratory of Mathematics for Nonlinear Sciences and Institute of Mathematical Finance. He is supported in part by the NSFC, under grant 79790130, the National Distinguished Youth Science Foundation of China under grant 19725106 and the Chinese Education Ministry Science Foundation under grant 2000024605, and the Cheung Kong Scholars Programme.

1. Introduction

In an ideal liquid market, trades by investors do not affect asset prices. This explains why most of the asset pricing models (see, e.g., Black and Scholes (1973)) assume that an investor cannot affect stock prices by trading regardless of the size or the direction of her trade. However, in an illiquid market, the price impact of investors' trading has been widely observed (see Chan and Lakonishok (1995), Keim and Madhavan (1996), Sharpe, Alexander and Bailey (1999), Jorion (2000), for example). Even for a very liquid market, large trades also directly affect asset prices. For example, examination of the limit order book of any specialist in NYSE reveals that the depth of the market at each price is not infinite. Trading beyond the quoted depth usually results in a worse price for at least part of the trade.

Since the 1987 stock market crash, there has been large literature on the possible causes of the crash. A widely accepted reason for the crash is program trading for portfolio insurance. It was argued that these trades greatly depressed stock prices and accelerated the downturn of the market. This indicates that even a very liquid market can turn into an illiquid market rather quickly in a crisis. In 1998, the collapse of the empire of the Long-Term Capital Management demonstrated again for large traders no market is liquid when investors panic.

Consistent with the above discussion, in this paper we take this existence of price impact of a large trader as given and examine how this price impact affects the replication of European contingent claims for the large trader. In particular, we assume that as the large trader buys, the stock price is driven up and as she sells, the stock price is pushed down.

In this paper we would like to address the following questions in such an illiquid market. First, is it still possible to perfectly replicate a European contingent claim with this adverse price impact? Second, even if it is still replicable, is it cheaper to super-replicate instead like in the presence of transaction cost? Third, if it is cheaper to replicate, what is the difference in the replication, including costs and strategies, from the Black-Scholes case? Finally, does the existence of the price impact help explain the well known "volatility smile"?

To answer these questions, we use the idea of the Four-Step-Scheme for forward and backward stochastic differential equations (FBSDEs) (see Ma, Protter and Yong (1994)) to derive a generalized nonlinear Black-Scholes partial differential equation for computing the replicating cost of a European contingent claim. We provide sufficient conditions under which the contingent claim is perfectly replicable. The replicating strategy involves an initial discrete trade followed by continuous trading. To replicate contingent claims whose payoffs are linear in the stock price (e.g., forwards, futures or shares), the large trader adopts the same strategy as in the case with no price impact. However, the cost is higher due to the adverse impact from the initial trade.

We then show that unlike the case with transaction cost, super-replicating is more costly than replicating in the presence of adverse price impact. Moreover, the effect of the price impact is summarized in the impact of the large trader's trading on the stock return volatility. This implies that even when the volatility without the trading of large traders is constant, the implied volatility computed using the Black-Scholes formula would be stochastic when there is trading from large traders. This suggests that the stochastic volatility models where the volatility process is exogenously specified such as Heston (1993) might be of a reduced form.

We find that compared to the case without price impact, at any time $t > 0$ the large trader generally buys more the stock and borrows more (shorts and lends more) to replicate an out-of-the-money call (put), but buys less the stock and borrows less (shorts and lends less) to replicate an in-the-money call (put). However, at time $t = 0$ after taking into account the price impact of the initial trade, the large trader generally buys more the stock and borrows more (shorts and lends more) to replicate a call (a put), but still buys less the stock and borrows less to replicate a deep-in-the-money call. In addition, the presence of price impact can potentially help explaining the well known "volatility smile" for calls (see Dumas, Fleming and Whaley (1998), for example). Intuitively, as a trader trades the price moves against him, therefore increasing the replicating cost. When an option is in-the-money, one needs to trade more in the stock. So this additional replicating cost (above Black-Scholes) is greater. On the other hand, option price is not sensitive to volatility in the Black-Scholes world for an in-the-money option. This implies that a large implied volatility is required to generate the higher price resulted from the price impact. When an option is out-of-the-money, however, one needs to trade less in the stock to replicate the option. So the excess cost is smaller, making the implied volatility smaller. However, the implied volatility is still greater than that in the case with no price impact.

In addition to the price impact model considered in this paper, a number of option valuation models are capable of explaining the "volatility smile" to some extent. The stochastic volatility models of Heston (1993) and Hull and White (1987), for example, can potentially explain the smile when the asset price and the volatility are negatively correlated. Similarly, the jump model of Bates (1996) is also consistent with the smile when the mean jump is negative. The deterministic volatility model examined by Dumas, Fleming and Whaley (1998) can also generate the smile. However, all these models assume exogenously the volatility process, while in our model the stock return volatility is endogenous, coming from the trading of the large trader.

In the presence of price impact, the no arbitrage price of a contingent claim for a large trader is no longer unique for trading a certain units of contingent claims. Rather, it consists of a continuum of prices within an interval. We find that this no-arbitrage interval expands as the price impact increases. The price impact also introduces nonlinearity into the dynamics of the value of a replicating portfolio . This in turn implies in particular that the replicating cost of two units of of an option is not necessarily twice the cost for one unit of the option. Indeed, we show that the replicating cost is approximately quadratic in the number of units of the option. We also show that the excess replicating cost of one option over Black-Scholes is significant and converges monotonically to the excess cost of trading one share as the option becomes more and more in-the-money.

There is an extensive literature on large investors. In the presence of asymmetric information, Kyle (1985) and Back (1992) used an equilibrium approach to investigate how informed traders reveal information and affect the market price through trading. As shown by Kyle (1985) and Back (1992), equilibrium asset prices are directly affected by the informed trader's trades. These models provide theoretic justifications on the existence, form and direction of the price impact a large trader can have on stock prices. In particular, the price impact form used in the numerical analysis section of this paper, which is linear in the trading size, is consistent with the equilibrium price impact forms derived in these models. As a matter of fact, Hubermman and Stanzl (2000) showed that when the price impact is time stationary, only the linear price impact form rules out arbitrage.

Some of the literature on large investors focuses on how large traders can manipulate stock prices. Examples include Jarrow (1992), Allen and Gale (1992), Vila (1989) and Bagnoli and Lipman (1990). Allen and Gale (1992) categorized these manipulations into three categories: The first is action-based manipulation, that is, manipulation based on actions that change the actual or perceived value of the stocks; The second is information-based manipulation, which is based on releasing false information; The third one is trade-based manipulation, by which the uninformed large trader can make a profit by just trading. Schönbucher and Wilmott (2000) showed that with exogenously given demand function from small investors, large investor can manipulate the share price to any level as desired. The paper then claimed that "Typically the market for the option will collapse because of the possibility of manipulation in the share price by the large trader". However, this argument ignores the cost of the manipulation.

Cvitanić and Ma (1996) and Ma and Yong (1999) also examined the hedging costs of options for a large investor. Cuoco and Cvitanić (1998) considered the effect of the price impact on the optimal consumption and investment policy. In these papers, it was assumed that price impact depends only on the total wealth and/or the position of the large investor but not how she trades.

Our model includes Frey (2000) as a special case. Frey (2000) considers the hedging cost of a large investor under the assumption that the interest rate is zero and the marginal price impact is proportional to the stock price. With this price impact function, given the same trade size, the investor has a smaller price impact on a stock with lower price and a greater impact on a stock with a higher price. He does not show the existence or the uniqueness of a classical solution to the pricing PDE derived in his paper. The initial price impact of a replicating strategy was also ignored. Sircar and Papanicolaou (1998) assumed exogenous demand function for the reference traders and derived a different nonlinear pricing PDE which depends on the exogenous income process of the reference traders and the relative size of the programme traders. Some asymptotic results were obtained.

The rest of the paper is organized as follows. In Section 2, we introduce our model. In Section 3, we derive the generalized Black-Scholes pricing PDE in the presence of price impact and provide the conditions under which a European contingent claim can be replicated. We also show super-replicating is more expensive. Section 4 contains the concluding remarks.

2. The Model

Throughout this paper we fix a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ on which a standard one dimensional Brownian motion $B(t)$ is defined with $\{\mathcal{F}_t\}_{t\geq 0}$ being its natural filtration augmented by all the $\mathbf{P}$-null sets. All the stochastic processes in the sequel are assumed to be $\{\mathcal{F}_t\}_{t\geq 0}$-adapted.

There are two assets being continuously traded in the market. The first asset is a money market account. The second is a risky asset, which we will call a stock. Let $S(t)$ be the ex-dividend stock price and $\delta(t, S(t))$ be the dividend yield of the stock. We assume that the risk free asset price $S_0(t)$ satisfies

$$\frac{dS_0(t)}{S_0(t)} = r(t, S(t))dt, \qquad t \geq 0, \tag{1}$$

where $r(t, S(t))$ is the interest rate and we allow it to depend on the (current) stock price directly. Next, different from the standard framework, and consistent with the situation in an illiquid market , we assume that the trading of the large trader has a direct impact on the stock price. In particular, when the trader buys the stock price goes up and when she sells the stock price goes down. If we let $N(t)$ be the number of shares that the trader has in the stock at time t, then the stock price $S(t)$ is assumed to evolve as follows,

$$\frac{dS(t)}{S(t)} = \mu(t, S(t))dt + \sigma(t, S(t))dB(t) + \lambda(t, S(t))dN(t), \quad t \geq 0, \tag{2}$$

where $\lambda(t, S(t)) \geq 0$ is the price impact function of the trader, $\mu(t, S(t))$ and $\sigma(t, S(t))$ are the expected return and the volatility of the return respectively in the absence of any trading by the large trader. The term $\lambda(t, S(t))dN(t)$ represents the price impact of the large investor's trading. We note that the classical Black-Scholes model is nothing but the special case with $\lambda(t, S(t)) \equiv 0$.

The wealth process $W(\cdot)$ for the trader then satisfies the following budget equation

$$\begin{aligned} dW(t) = r(t, S(t))W(t)dt + N(t)S(t)[\mu(t, S(t)) + \delta(t, S(t)) - r(t, S(t))]dt \\ +N(t)S(t)\sigma(t, S(t))dB(t) + N(t)S(t)\lambda(t, S(t))dN(t), \quad t \geq 0. \end{aligned} \tag{3}$$

The price impact term in (2) leads to the last quadratic term in the budget equation. This quadratic term is the only difference from the wealth equation for a small trader. The presence of this term implies that unlike the standard case, the wealth dynamics for a large trader is no longer linear in her trading strategy N. Implications of this nonlinearity will be explored in later parts of this paper. We assume that $N(\cdot)$ is an Itô process, i.e., $N(\cdot)$ satisfies the following :

$$\begin{cases} dN(t) = \eta(t)dt + \zeta(t)dB(t), \quad t \geq 0, \\ N(0) = N_0, \end{cases} \tag{4}$$

for some processes $\eta(\cdot)$ and $\zeta(\cdot)$ (to be endogenously determined), where N_0 is the initial number of shares in the stock. Thus, by (4) and (2), we have

$$\frac{dS(t)}{S(t)} = [\mu(t, S(t)) + \lambda(t, S(t))\eta(t)]dt + [\sigma(t, S(t)) + \lambda(t, S(t))\zeta(t)]dB(t), \quad t \geq 0. \tag{5}$$

Consequently, the wealth process $W(\cdot)$ satisfies the following SDE:

$$\begin{aligned} dW(t) = &\Big\{ r(t, S(t))W(t) + [\mu(t, S(t)) + \delta(t, S(t)) - r(t, S(t)) \\ &+\lambda(t, S(t))\eta(t)]N(t)S(t) \Big\} dt \\ &+[\sigma(t, S(t)) + \lambda(t, S(t))\zeta(t)]N(t)S(t)dB(t), \quad t \geq 0. \end{aligned} \tag{6}$$

3. Replication of a European Contingent Claim

Let $h(S(T))$ be the payoff of a European contingent claim maturing at time T, where $h : \mathbb{R}_+ \to \mathbb{R}$ and $S(T)$ is the price of the stock at time T. Hereafter, for convenience, we simply call $h(S(T))$ a contingent claim. Then replicating such a contingent claim amounts to solving (4), (5) and (6) subject to the terminal condition $W(T) = h(S(T))$. For clarity, we collect them together in the following system of stochastic differential equations:

$$\left\{\begin{array}{l} dN(t) = \eta(t)dt + \zeta(t)dB(t), \\ \frac{dS(t)}{S(t)} = [\mu(t,S(t)) + \lambda(t,S(t))\eta(t)]dt \\ \qquad\qquad +[\sigma(t,S(t)) + \lambda(t,S(t))\zeta(t)]dB(t), \\ dW(t) = \Big\{r(t,S(t))W(t) + [\mu(t,S(t)) + \delta(t,S(t)) - r(t,S(t)) \\ \qquad\qquad +\lambda(t,S(t))\eta(t)]N(t)S(t)\Big\}dt \\ \qquad\qquad +[\sigma(t,S(t)) + \lambda(t,S(t))\zeta(t)]N(t)S(t)dB(t), \\ N(0) = N_0, \quad S(0) = S_0, \quad W(T) = h(S(T)). \end{array}\right. \tag{7}$$

This system of SDEs is called a forward-backward stochastic differential equation (FBSDE, for short) system, since it involves solving forward for $N(t)$ and $S(t)$ and backward for $W(t)$.

In the presence of price impact, as the large trader trades, the stock price is directly affected and thus the potential payoff of a contingent claim is also changed. Therefore, one of the interesting questions is whether the large trader can still replicate the contingent claim with this price impact. The following theorem implies that the answer is positive and provides a generalized nonlinear Black-Scholes pricing PDE required to compute the replicating cost.

Theorem 1 *Under some regularity conditions, there exists a unique classical solution $f(\cdot\,,\cdot)$ of the following equation:*

$$\left\{\begin{array}{l} f_t + \dfrac{\sigma(t,S)^2S^2f_{SS}}{2[1-\lambda(t,S)Sf_{SS}]^2} + (r(t,S)-\delta(t,S))Sf_S - r(t,S)f = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad (S,t) \in [0,T) \times (0,\infty), \\ f(T,S) = h(S). \end{array}\right. \tag{8}$$

In addition, FBSDE (7) admits a unique adapted solution $(S(\cdot), W(\cdot), N(\cdot))$ such that

$$\left\{\begin{array}{l} W(t) = f(t,S), \\ N(t) = f_S(t,S), \end{array}\right. \tag{9}$$

and $S(\cdot)$ satisfies:

$$\frac{dS(t)}{S(t)} = \widehat{\mu}(t,S(t))dt + \widehat{\sigma}(t,S(t))dB(t), \qquad t \geq 0, \tag{10}$$

where

$$\begin{cases} \widehat{\mu}(t,S) \equiv \dfrac{\mu(t,S) + \lambda(t,S) f_{St}}{1 - \lambda(t,S) S f_{SS}} + \dfrac{\lambda(t,S)\sigma^2(t,S) S^2 f_{SSS}}{2\left(1 - \lambda(t,S) S f_{SS}\right)^3}, \\ \widehat{\sigma}(t,S) \equiv \dfrac{\sigma(t,S)}{1 - \lambda(t,S) S f_{SS}}. \end{cases} \tag{11}$$

PROOF. See Appendix.

This theorem suggests that to replicate a contingent claim, one has to first trade a discrete $f_S(0, S(0))$ shares of the stock and then follow a continuous trading strategy prescribed by $N(t)$. Given the stock price $S(0_-)$ before this discrete trade and (2), the stock price $S(0)$ the initial trade will drive to can be calculated as follows. Let $N(0_-) = 0$ and $N(0) = f_S(0, S(0))$. Assuming the trader can still work the initial order even when her trading speed is large[c]. By (2), we have,

$$dS = \lambda(t,S) S dN(t),$$

for a discrete trade. This implies that $S(0)$ solves

$$\int_{S(0_-)}^{S(0)} \frac{dS}{\lambda(0,S)S} = f_S(0, S(0)). \tag{12}$$

and the initial cost of acquiring $f_S(0, S(0))$ shares is thus

$$c = \int_{S(0_-)}^{S(0)} S(t) dN(t) = \int_{S(0_-)}^{S(0)} \frac{dS}{\lambda(0,S)}. \tag{13}$$

The initial cost of replicating the contingent claim $h(S(T))$ is therefore

$$f^h = f(0, S(0)) - S(0) f_S(0, S(0)) + c. \tag{14}$$

The pricing PDE (8)implies that the effect of the price impact on the replication after the initial trade is only through its impact on the stock return volatility. This suggests, in particular, that the replicating cost for the large investor in Cuoco and Cvitanic (1998) is the same as for a small investor, since in their model, the position of the large investor only affects the expected return but not the volatility.

As is well known, in the presence of transaction costs, super-replicating an option is cheaper than replicating. This is because exact replication involves

[c] Alternatively, one can assume that all $f_S(0, S(0))$ shares are traded at $S(0)$. This would only increase the replicating cost, magnify the effect of price impact and thus strengthen the main results in this paper.

continuous trading and thus incurs infinite costs. Similar to the transaction cost case, the excess costs the large investor with price impact incurs increases with the frequency of her trading. This seems to suggest that super-replicating (buying a share and never trading again to super-replicate a call, for example) might also be less expensive than replicating in the presence of price impact. However, the following theorem implies that this is not the case.

Theorem 2 *Let $h(\cdot)$ and $\overline{h}(\cdot)$ be such that*

$$h(S) \leq \overline{h}(S), \qquad S \in (0, \infty). \tag{15}$$

Suppose $f(\cdot,\cdot)$ and $\overline{f}(\cdot,\cdot)$ satisfy the following:

$$\begin{cases} f_t + \dfrac{\sigma(t,S)^2 S^2 f_{SS}}{2[1-\lambda(t,S)Sf_{SS}]^2} + (r(t,S)-\delta(t,S))Sf_S - r(t,S)f \geq 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad (t,S) \in [0,T) \times (0,\infty), \\ f(T,S) \leq h(S), \end{cases} \tag{16}$$

and

$$\begin{cases} \overline{f}_t + \dfrac{\sigma(t,S)^2 S^2 \overline{f}_{SS}}{2[1-\lambda(t,S)S\overline{f}_{SS}]^2} + (r(t,S)-\delta(t,S))S\overline{f}_S - r(t,S)\overline{f} \leq 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad (t,S) \in [0,T) \times (0,\infty), \\ \overline{f}(T,S) \geq \overline{h}(S), \end{cases} \tag{17}$$

and some regularity conditions are satisfied. Then

$$f(t,S) \leq \overline{f}(t,S), \qquad (t,S) \in [0,T] \times (0,\infty). \tag{18}$$

If a strict equality holds in (15), so does in (18). The above is the case if, in particular, f and $\overline{f}$ are solutions to (8) corresponding to h and $\overline{h}$.

PROOF. See Appendix. □

This theorem shows that the replicating strategy described above is indeed the cheapest way to hedge a European contingent claim for the large trader. This result suggests that the excess cost incurred from the adverse price impact is of a lower order than the transaction cost.

4. Concluding Remarks

In this paper, we investigate the effect of the price impact of a large trader on the replication of a European contingent claim. Consistent with trading in an illiquid market, we assume that a large trader drives up the stock price as she buys and pushes it down as she sells. We obtain a generalized nonlinear Black-Scholes pricing partial differential equation. We derive sufficient conditions for the existence and uniqueness of a classical solution and the replicability of the contingent claim by the large trader. We also show that in contrast to the case with transaction costs, super-replicating an option with adverse price impact is more costly than the unique replicating strategy proposed in the paper.

We find that compared to the Black-Scholes strategy, at any time $t > 0$ the large trader generally buys more the stock and borrows more (shorts and lends more) to replicate an out-of-the-money call (put), but buys less the stock and borrows less (shorts and lends less) to replicate an in-the-money call (put). However, at time $t = 0$ after taking into account the price impact of the initial trade, the large trader generally buys more the stock and borrows more (shorts and lends more) to replicate a call (a put), but still buys less the stock and borrows less to replicate a deep-in-the-money call. The replicating cost is approximately quadratic in the number of units of options to be replicated. The excess cost the large investor incurs is found to be significant with the adverse price impact for options that are not far out of the money. We also show that the presence of price impact can potentially help explain the well known "volatility smile" for calls.

Different from most of the existing literature on large traders, this model allows a direct price impact. This provides a reasonable model also for the pricing of a large block order. An interesting problem would be to estimate the price impact functions for illiquid assets, such as some small stocks. This way one can then empirically test the implications of this model, such as comparing the model implied excess costs over the Black-Scholes prices to the observed excess costs. An equilibrium model with informed traders trading in both options and stocks would shed light on the form and magnitude of the price impact function. Another interesting issue might be the optimal liquidation strategy for a fund facing adverse price impact.

References

ALLEN, F. AND D. GALE, 1992, " Stock-Price Manipulation", *Review of Financial Studies*, 5, 503-529.

BACK, K., 1992, " Insider Trading in Continuous Time", *Review of Financial Studies*, 5, 387-409.

BATES, D. 1996, "Jumps and stochastic volatility: Exchange rate process implicit in Deutsche mark options", , *Review of Financial Studies*, 9, 69-107.

BAGNOLI, M. AND B. L. LIPMAN, 1990, "Stock Price Manipulation through Takeover Bids," *The Rand Journal of Economics*, 27, 124-147.

BERMITIS, D., AND A. W. LO, 1998, "Optimal control of execution costs," *Journal of Financial Markets*, 1, 1-50.

BLACK, F. AND M. SCHOLES, 1973, " The pricing of options and corporate liability," *Journal of Political Economy*, 81, 637-659.

CHAN, L., AND J. LAKONISHOK, 1995, " The behavior of stock prices around institutional trades," *Journal of Finance*, 50, 1147-1174.

CUOCO, D. AND J. CVITANIĆ, 1998, "Optimal Consumption Choices for a 'Large' Investor", *Journal of Economic Dynamics and Control*, 22, 401-436.

CVITANIĆ, J. AND J. MA, 1996, "Hedging Options for a Large Investor and Forward-backward SDEs", *Annals of Applied Probability*, 6, 370-398.

DUMAS, B., J. FLEMING, AND R. E. WHALEY, 1998, "Implied Volatility Functions: Empirical Tests" , *Journal of Finance*, Vol. LIII, no. 6, 2059-2106.

EL KAROUI, N., S. PENG, AND M.-C. QUENEZ, 1997, "Backward Stochastic Differential Equation in Finance", *Mathematical Finance*, 7, 1-71.

FREY, R., 2000, "Market Illiquidity as a Source of Model Risk in Dynamic Hedging", Working paper.

HESTON, S., 1993, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *Review of Financial Studies*, v6, n2, 327-343.

HUBERMAN, G., AND W. STANZL, 2001, "Arbitrage-free Price-Update and Price-Impact Functions ", working paper, Yale University.

HULL, J., AND A. WHITE, 1987, "The pricing of options on assets with stochastic volatility", *Journal of Finance*, 42, 281-300.

KEIM, D., AND A. MADHAVAN, 1996, "The upstairs market for large -block transactions: analysis and measurement of price effects," *Review of Financial Studies*, v9, n1, 1-36.

JARROW, R.A., 1992, "Market Manipulation, Bubbles Corners and Short Squeezes", *Journal of Financial and Quantitative Analysis*, 27, 311-336.

JORION, P, 2000, *Value at Risk*, McGraw-Hill, New York.

KAUL, A., V. MEHROTRA AND R. MORCK, 2000, "Demand Curves for Stocks Do Slope Down: New Evidence from an Index Weights Adjustment," *Journal of Finance*, 55, 893-912.

KAVAJECZ, K., 1999, "A Specialist's Quote Depth and the Limit Order Book," *Journal of Finance*, 54, 747-771.

KYLE, A., 1985, "Continuous Auctions and Insider Trading", *Econometrics*, 53, 1315-1335.

LADYŽENSKAJA, O. A., V. A. SOLONNIKOV, AND N. N. URAL'CEVA, 1968, *Linear and Quasi-linear Equations of Parabolic Type*, AMS, Providence, RI.

MA, J., P. PROTTER, AND J. YONG, 1994,"Solving forward-backward stochastic differential equations explicitly - a four step scheme", *Prob. Th. & Rel. Fields*, 98, 339-359.

MA, J. AND J. YONG, 1999,*Forward Backward Stochastic Differential Equations and Their Applications*, Springer-Verlag, New York.

SHARPE, W. F., G. J. ALEXANDER, AND J. V. BAILEY, 1999, *Investments* , Prentice Hall, New Jersey.

SCHÖNBUCHER, P. J., AND P. WILMOTT, 2000, "The Feedback Effect of Hedging in Illiquid Markets", *SIAM Journal of Applied Mathematics*, 61, 232-272.

SIRCAR, K. R., AND G. PAPANICOLAOU, 1998, "General Black-Scholes Models Accounting for Increased Market Volatility From Hedging Strategies", *Applied Mathematical Finance*, 5, 45-82.

VILA, J.L., 1989, "Simple Games of Market Manipulation",*Economic Letters*, 29, 21-26.

YONG, J., 1999, "European Type Contingent Claims in an Incomplete Market with Constrained Wealth and Portfolio", *Mathematical Finance*, 9, 387-412.

YONG, J. AND X. Y. ZHOU, 1999, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York.

Appendix

In this appendix, we present a heuristic derivation of the generalized Black-Scholes PDE (8).

Suppose $(S(\cdot), W(\cdot), N(\cdot))$ is an adapted solution of FBSDE (7) and

$$W(t) = f(t, S(t)), \qquad t \in [0,T], \text{ a.s. }, \tag{19}$$

for some smooth function $f(\cdot\,,\cdot)$. Applying Itô's formula to (19), and using (7), we obtain (suppressing the arguments (S,t) for simplicity)

$$\begin{aligned} &[rW + (\mu + \delta - r + \lambda\eta)NS]dt + [\sigma + \lambda\zeta]NSdB = dW \\ &= \left\{f_t + Sf_S(\mu + \lambda\eta) + \tfrac{1}{2}S^2 f_{SS}(\sigma + \lambda\zeta)^2\right\}dt + Sf_S(\sigma + \lambda\zeta)dB. \end{aligned} \tag{20}$$

Comparing the diffusion terms in the above, we see that one should choose

$$N(t) = f_S(S(t), t), \quad t \in [0,T] \text{ a.s.} \tag{21}$$

Then comparing the drift terms in (20) and using (19) and (21), one has

$$\begin{aligned} 0 &= f_t + Sf_S(\mu + \lambda\eta) + \tfrac{1}{2}S^2 f_{SS}(\sigma + \lambda\zeta)^2 - [rf + (\mu + \delta - r + \lambda\eta)Sf_S] \\ &= f_t + \tfrac{1}{2}S^2 f_{SS}(\sigma + \lambda\zeta)^2 + (r - \delta)Sf_S - rf. \end{aligned} \tag{22}$$

We hope to obtain an equation in $f(\cdot\,,\cdot)$. Thus, we need to eliminate ζ in the above. To this end, let us first note that

$$\begin{aligned} \eta dt + \zeta dB = dN = d[f_S] = \left[f_{St} + Sf_{SS}(\mu + \lambda\eta) + \tfrac{1}{2}S^2 f_{SSS}(\sigma + \lambda\zeta)^2\right]dt \\ + Sf_{SS}(\sigma + \lambda\zeta)dB. \end{aligned} \tag{23}$$

Hence, comparing the diffusion terms, we obtain

$$\zeta = (\sigma + \lambda\zeta)Sf_{SS}, \tag{24}$$

which implies (assuming that $\lambda Sf_{SS} \neq 1$)

$$\zeta = \frac{\sigma Sf_{SS}}{1 - \lambda Sf_{SS}}. \tag{25}$$

Thus, the "volatility" of the stock (noting (7)) is given by

$$\widehat{\sigma} \equiv \sigma + \lambda\zeta = \frac{\sigma}{1 - \lambda Sf_{SS}} \tag{26}$$

Combining (22) and (26), we see that one should choose $f(\cdot\,,\cdot)$ to be a solution of the PDE (8) in Theorem 1.

Finally, comparing the drift terms in (23) we have (suppressing arguments (t, S))

$$\eta = f_{St} + S f_{SS}(\mu + \lambda\eta) + \tfrac{1}{2}S^2 f_{SSS}(\sigma + \lambda\zeta)^2, \tag{27}$$

which implies (still assuming $\lambda S f_{SS} \neq 1$)

$$\eta = \frac{1}{1 - \lambda S f_{SS}}\left\{f_{St} + \mu S f_{SS} + \frac{\sigma^2 S^2 f_{SSS}}{2(1 - \lambda S f_{SS})^2}\right\}. \tag{28}$$

Hence, the "appreciation rate" of $S(\cdot)$ is given by (noting (7) and (28))

$$\widehat{\mu} \equiv \mu + \lambda\eta = \frac{\mu + \lambda f_{St}}{1 - \lambda S f_{SS}} + \frac{\lambda\sigma^2 S^2 f_{SSS}}{2(1 - \lambda S f_{SS})^3}. \tag{29}$$

Arbitrage Pricing Systems in a Market Driven by an Itô Process

Shunlong Luo and Jia-an Yan[a]
Academy of Mathematics and System Sciences
Chinese Academy of Sciences, Beijing 100080, P. R. China

Qiang Zhang
Department of Economics and Finance
City University of Hong Kong, Kowloon, Hong Kong

Abstract. A pair of numeraire and equivalent martingale measure is called an arbitrage pricing system. In a security market driven by an Itô process, if we take the wealth process of an admissible self-financing strategy as a numeraire, then there is a natural family of equivalent martingale measures associated with market prices of risk. A subclass of arbitrage pricing systems is identified explicitly by a maximum entropy rationale in the spirit of Föllmer, Schweizer and Sondermann.

1. Introduction

The arbitrage argument for contingent claims valuation, initiated by Ross (1976), standardized by Harrison and Kreps (1979), Harrison and Pliska (1981), is a great achievement of modern asset pricing theory. In the literature on mathematical finance, the fundamental theorem of asset pricing states roughly that the absence of arbitrage opportunities is essentially equivalent to the existence of a probability measure (called equivalent local martingale measure) such that all security prices, denominated by a given numeraire, are local martingales under this new probability measure (see Delbaen and Schachermayer (1994, 1998a). The existence or no-existence of equivalent local martingale measures depends crucially on the choice of numeraire (see Delbaen and Shachermayer (1995)). To avoid this drawback, the fundamental theorem of asset pricing was reformulated in the equivalent martingale measure setting by Yan (1998), in which such a market is called a *fair market.* Recently, Xia and Yan (2001) further clarify some basic concepts and results in arbitrage pricing theory for fair markets.

We consider a fair market. An admissible self-financing strategy is said to be *regular*, if its wealth process denominated by a trading asset price process is a martingale under a martingale measure corresponding to this numeraire as-

[a]Luo's work was supported by the National Natural Science Foundation. Yan's work was supported by the Ministry of Science and Technology, the 973 project on mathematics. Their works are also supported by the knowledge innovation program of the Chinese Academy of Sciences.

set. Thus, in a fair market, if the wealth process of an admissible self-financing and regular strategy is strictly positive, we can take it as a numeraire. A pair of numeraire and a corresponding martingale measure is called an *arbitrage pricing system.* In an arbitrage pricing system, the denominated fair price of a contingent claim is defined as the mathematical expectation of the denominated contingent claim with respect to the martingale measure. Two arbitrage pricing systems are said to be equivalent, if for any contingent claims they determine the same fair price. When market is complete, for a given numeraire the martingale measure is unique. In this case any contingent claim can be perfectly hedged. The fair price of a contingent claim is nothing but the cost of the hedge. However, if the market is incomplete, there exist infinitely many martingale measures, the valuation and hedging issues are not definite. Many authors studied these issues from various perspectives. Föllmer and Sondermann (1986), Föllmer and Schweizer (1991) proposed the notion of minimal martingale measure and mean-variance criterion, and related it to an entropy rationale. Hofmann, Platen and Schweizer (1992) applied the idea to incomplete markets with stochastic volatility and presented an elegant orthogonal decomposition interpretation of the minimal martingale measure. Delbaen and Schachermayer (1998b) studied the variance-optimal martingale measures.

The relationship between numeraires and martingale measures is well understood in financial literature, see e.g. Geman, El Karoui and Rochet (1995), Bajeux-Besnainou and Portait (1997). Their main concerns are around two cases: (1) take the bond price as the numeraire, the corresponding martingale measure is a simple Girsanov transform of the objective probability; (2) keep the objective probability as the martingale measure, the numeraire is the reciprocal of state price density process, also known as the state price deflater, or the pricing kernel, or the Arrow-Debreu price (Duffie(1994)), or the numeraire portfolio, see e.g. Long (1990), Bajeux-Besnainou and Portait (1997), Artzner (1997), Yan, Zhang and Zhang (2000).

In this paper, we shall interpolate continuously between the above two extreme cases. Specifically, we show that in an incomplete security market driven by an Itô process, if we take the wealth process of an admissible self-financing and regular strategy as numeraire, then there is a natural family of equivalent martingale measures associated with *market prices of risk.* Moreover, all arbitrage pricing systems with different numeraires and martingale measures associated with the same market price of risk are equivalent. We also identify a unique martingale measure by a maximum entropy argument in the spirit of Föllmer and Schweizer (1991). In particular, if we take the bond price as the numeraire, then this unique one coincides with the minimal martingale measure of Föllmer and Schweizer (1991). If we take the growth-optimal wealth process as the numeraire, then the objective probability itself becomes a martingale measure.

2. Security Market Driven by an Itô Process

Consider a security market consisting of one bond and m risky assets on a finite time horizon $[0,\ T]$. We assume that

$$dB_t = B_t r_t dt, \qquad t \in [0,\ T]$$
$$dS_t^{(i)} = S_t^{(i)}\Big(\mu_t^{(i)} dt + \sum_{j=1}^{n} \sigma_t^{(ij)} dW_t^{(j)}\Big), \quad i = 1, 2, \cdots, m,$$

where B_t is the price process of the bond, S_t^i is the price process of i-th asset, $\{W_t = (W_t^{(1)}, W_t^{(2)}, \cdots, W_t^{(n)})\}_{t\in[0,\ T]}$ is an n-dimensional Brownian motion, defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, with natural filtration $\mathbf{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$. The coefficients $\{r_t\}$, $\{\mu_t^{(i)}\}$ and $\{\sigma_t^{(ij)}\}$ are all $\mathbf{F}$-adapted continuous processes. We assume that $m \le n$. According to a well known folklore theorem, under some purely technical conditions, the market is complete when $m = n$, incomplete when $m < n$. See Karatzas and Shreve (1999) or Yan (1998).

Let

$$\mu_t = (\mu_t^{(1)}, \mu_t^{(2)}, \cdots, \mu_t^{(m)}),\ \widetilde{\mu}_t = (\mu_t^{(1)} - r_t, \mu_t^{(2)} - r_t, \cdots, \mu_t^{(m)} - r_t)$$

be row vectors, $\sigma = (\sigma_t^{(ij)})$ be an $m \times n$ matrix with ith row vector

$$\sigma_t^{(i)} = (\sigma_t^{(i1)}, \sigma_t^{(i2)}, \cdots, \sigma_t^{(in)}),\ i = 1, 2, \cdots, m.$$

We shall use vector product and denote transpose of a vector or a matrix by $*$. We assume throughout this paper that all the coefficients $\{r_t\}_{t\in[0,T]}$, $\{\mu_t^{(i)}\}_{t\in[0,T]}$, $\{\sigma_t^{(ij)}\}_{t\in[0,T]}$ and $\{(\sigma_t\sigma_t^*)^{-1}\}_{t\in[0,T]}$ are integrable on $[0,\ T]\times\Omega$, and the matrix σ is of full rank m.

Note the stock price has the explicit expression

$$S_t^{(i)} = S_0^{(i)}\exp\Big\{\int_0^t \Big(\mu_s^{(i)} - \frac{1}{2}\sigma_s^{(i)}\sigma_s^{(i)*}\Big)ds + \int_0^t \sigma_s^{(i)} dW_s^*\Big\}, \tag{1}$$

where $\sigma_s^{(i)} dW_s^* = \sum_{j=1}^{n} \sigma_s^{(ij)} dW_s^{(j)}$ is the vector product.

Let $\pi = \{\pi_t = (\pi_t^{(1)}, \pi_t^{(2)}, \cdots, \pi_t^{(m)})\}_{t\in[0,\ T]}$ be a portfolio process. $\pi_t^{(i)}$ is the proportion of wealth invested in the stock $S^{(i)}$ at time t, $i = 1, 2, \cdots, m$.

The remaining proportion $1-\sum_{i=1}^m \pi_t^{(i)}$ is invested in the bond B at time t. The corresponding self-financing wealth process X^π is determined by

$$\begin{aligned} dX_t^\pi &= \frac{\left(1-\sum_{i=1}^m \pi_t^{(i)}\right)X_t^\pi}{B_t}dB_t + \sum_{i=1}^m \frac{\pi_t^{(i)}X_t^\pi}{S_t^{(i)}}dS_t^{(i)} \\ &= X_t^\pi\Big(1-\sum_{i=1}^m \pi_t^{(i)}\Big)r_t dt + X_t^\pi \sum_{i=1}^m \pi_t^{(i)}\Big(\mu_t^{(i)}dt + \sum_{j=1}^n \sigma_t^{(ij)}dW_t^{(j)}\Big) \\ &= X_t^\pi\Big((r_t+\pi_t\widetilde{\mu}_t^*)dt + \pi_t\sigma_t dW_t^*\Big). \end{aligned}$$

Without loss of generality, we assume that $X_0^\pi = 1$. The explicit solution is

$$X_t^\pi = \exp\Big\{\int_0^t \Big(r_s + \pi_s\widetilde{\mu}_s^* - \frac{1}{2}(\pi_s\sigma_s)(\pi_s\sigma_s)^*\Big)ds + \int_0^t \pi_s\sigma_s dW_s^*\Big\}. \tag{2}$$

3. A Natural family of Arbitrage Pricing Systems

Let $\mathcal{Y}$ denote the collection of all adapted row vector processes $\mathbf{y}_t$ satisfying the equation

$$\sigma_t\mathbf{y}_t^* = \widetilde{\mu}_t^*, \qquad dt\times d\mathrm{P}-\text{a.e., a.s. on } [0,T]\times\Omega, \tag{3}$$

and such that the process $M_t = \exp\Big\{-\frac{1}{2}\int_0^t \mathbf{y}_s\mathbf{y}_s^* ds - \int_0^t \widehat{\mathbf{y}}_s dW_s^*\Big\}, 0\le t\le T$, is a martingale. Then the market is fair if and only if $\mathcal{Y}$ is non-empty. For example, if there is a solution $\mathbf{y}$ of the equation (3) such that $\mathrm{E}\Big(\exp\frac{1}{2}\int_0^T \mathbf{y}_t\mathbf{y}_t^* dt\Big) < \infty$, then by Novikov criterion for the uniform integrability of exponential martingale, $\mathcal{Y}$ is non-empty. In the following we assume that $\mathcal{Y}$ is non-empty.

The following theorem can be regarded as a corollary of a general result in Yan (1998).

Theorem 1. *Let $X^\pi = \{X_t^\pi\}_{t\in[0,\,T]}$ be the self-financing wealth process determined by a portfolio π satisfying the condition*

$$\mathrm{E}\Big(\exp\frac{1}{2}\int_0^T \pi_t\sigma_t(\pi_t\sigma_t)^* dt\Big) < \infty. \tag{4}$$

Let $\mathbf{y}\in\mathcal{Y}$. Put

$$\widehat{\pi}_t := \mathbf{y}_t - \pi_t\sigma_t, \qquad t\in[0,T]. \tag{5}$$

Define a new probability measure $\mathrm{P}_{\pi,,\mathbf{y}}$ on $(\Omega,\mathcal{F}_T)$ by

$$\frac{d\mathrm{P}_{\pi,\mathbf{y}}}{d\mathrm{P}} = \exp\Big\{-\frac{1}{2}\int_0^T \widehat{\pi}_t\widehat{\pi}_t^* dt - \int_0^T \widehat{\pi}_t dW_t^*\Big\}. \tag{6}$$

Then $(X^\pi, \mathrm{P}_{\pi,\mathbf{y}})$ is an arbitrage pricing system.

Moreover, if π' is another portfolio satisfying the same condition as that for π, then two arbitrage pricing systems $(X^\pi, \mathrm{P}_{\pi,\mathbf{y}})$ and $(X^{\pi'}, \mathrm{P}_{\pi',\mathbf{y}})$ are equivalent.

Proof. We define a new probability measure on $(\Omega, \mathcal{F}_T)$ by

$$\frac{d\mathrm{Q}_{\mathbf{y}}}{d\mathrm{P}} = \exp\Big\{ -\frac{1}{2}\int_0^T \mathbf{y}_t\mathbf{y}_t^* dt - \int_0^T \mathbf{y}_t dW_t^* \Big\}. \tag{7}$$

Then (B, Q_y) is an arbitrage pricing system, and $B_t^{-1}X_t^\pi$ is a Q_y-martingale. In particular, π is a regular portfolio. Since it is easy to verify that

$$\frac{X_t^\pi}{B_t} = \frac{d\mathrm{P}_{\pi,\mathbf{y}}}{d\mathrm{P}}\Big|_{\mathcal{F}_t} \Big/ \frac{d\mathrm{Q}_{\mathbf{y}}}{d\mathrm{P}}\Big|_{\mathcal{F}_t} = \frac{d\mathrm{P}_{\pi,\mathbf{y}}}{d\mathrm{Q}_{\mathbf{y}}}\Big|_{\mathcal{F}_t},$$

by Bayes rule on conditional expectation, we know that $(X^\pi, \mathrm{P}_{\pi,\mathbf{y}})$ is an arbitrage pricing system, and it is equivalent to the arbitrage pricing system (B, Q_y) (see Yan (1998) for more detail). Consequently, another conclusion is trivially true.

Remark 1. $\mathbf{y} = \{\mathbf{y}_t\}_{t\in[0,T]}$ is interpreted as the market price of risk. For a given portfolio π, we have associated a natural family of arbitrage pricing systems $(X^\pi, \mathrm{P}_{\pi,\mathbf{y}})$ to market prices of risk in $\mathcal{Y}$. If $m = n$ and the matrix σ_t is invertible, for any $t \in [0,T]$, then Eq. (3) has a unique solution $\mathbf{y}_t^* = \sigma_t^{-1}\widetilde{\mu}_t^*$. In this case the martingale measure is unique, and the market is complete.

Remark 2. A well known fact is that if we take an appropriate portfolio π and use X^π as the numeraire, it is possible to render the objective probability P to be a martingale measure. In fact, let $\widehat{\pi}_t = (0, 0, \cdots, 0)$, $\forall\, t \in [0,T]$, that is

$$\mathbf{y}_t - \pi_t\sigma_t = (0, 0, \cdots, 0).$$

Multiplying σ_t^* from right, since $\sigma_t\mathbf{y}_t^* = \widetilde{\mu}_t^*$ implies that $\mathbf{y}_t\sigma_t^* = \widetilde{\mu}_t$, we have $\widetilde{\mu}_t - \pi_t\sigma_t\sigma_t^* = 0$. Thus if we take

$$\pi_t = \widetilde{\mu}_t(\sigma_t\sigma_t^*)^{-1}, \quad \mathbf{y}_t = \widetilde{\mu}_t(\sigma_t\sigma_t^*)^{-1}\sigma_t, \qquad t \in [0,T],$$

we have $\mathrm{P}_{\pi,\mathbf{y}} = \mathrm{P}$. The portfolio $\pi = \{\pi_t = \widetilde{\mu}_t(\sigma_t\sigma_t^*)^{-1}\}_{t\in[0,T]}$ is a proportional strategy. As is well known, its corresponding wealth process is exactly the growth-optimal wealth process. This is also the instance of the portfolio numeraire studied by Long (1990), Bajeau-Besnainou and Portait (1997), and surveyed by Artzner (1997).

According to the general theory of generalized inverse of a matrix, there is a *natural* solution of Eq. (3) given by $\tilde{\mathbf{y}}_t = \widetilde{\mu}_t(\sigma_t\sigma_t^*)^{-1}\sigma_t$. Assume that his solution belongs to $\mathcal{Y}$, then it will be of particular interest as can be seen from the following theorem which results from a simple application of Föllmer and Schweizer's idea (1991)(see also Hofmann, Platen and Schweizer, 1992).

Theorem 2. *For any portfolio* π *satisfying condition (4), the following maximization problem*

$$\text{arg.max}_{\mathbf{y}\in\mathcal{Y}} \mathrm{E}^{\mathrm{P}}(\log \frac{d\mathrm{P}_{\pi,\mathbf{y}}}{d\mathrm{P}})$$

is achieved at $\tilde{\mathbf{y}}_t = \widetilde{\mu}_t(\sigma_t\sigma_t^*)^{-1}\sigma_t, \quad t \in [0,T]$, *if* $\tilde{\mathbf{y}} \in \mathcal{Y}$. *In particular, if* $\mathrm{E}\Big(\exp\frac{1}{2}\int_0^T \tilde{\mathbf{y}}_t\tilde{\mathbf{y}}_t^* dt\Big) < \infty$, *then it is the case.*

Proof. From Eqs. (8),(9) and (3), we have

$$\mathrm{E}^{\mathrm{P}}\Big(\log \frac{d\mathrm{P}_{\pi,\mathbf{y}}}{d\mathrm{P}}\Big) = -\frac{1}{2}\mathrm{E}^{\mathrm{P}}\int_0^T \widehat{\pi}_t\widehat{\pi}_t^* dt$$
$$= -\frac{1}{2}\mathrm{E}^{\mathrm{P}}\int_0^T (\mathbf{y}_t - \pi_t\sigma_t)(\mathbf{y}_t - \pi_t\sigma_t)^* dt$$
$$= -\frac{1}{2}\mathrm{E}^{\mathrm{P}}\int_0^T (\mathbf{y}_t\mathbf{y}_t^* - \mathbf{y}_t\sigma_t^*\pi_t^* - \pi_t\sigma_t\mathbf{y}_t^* + \pi_t\sigma_t\sigma_t^*\pi_t^*)dt$$
$$= -\frac{1}{2}\mathrm{E}^{\mathrm{P}}\int_0^T (\mathbf{y}_t\mathbf{y}_t^* - 2\widetilde{\mu}_t\pi_t^* + \pi_t\sigma_t\sigma_t^*\pi_t^*)dt.$$

Thus the original maximization problem is equivalent to the following one:

$$\text{arg.max}_{\mathbf{y}\in\mathcal{Y}}\Big(-\frac{1}{2}\mathrm{E}^{\mathrm{P}}\int_0^T \mathbf{y}_t\mathbf{y}_t^* dt\Big).$$

We shall prove a stronger result: $\forall t \in [0,T]$, $\text{arg.max}_{\mathbf{y}_t}(-\frac{1}{2}\mathbf{y}_t\mathbf{y}_t^*)$ subject to $\sigma_t\mathbf{y}_t^* = \widetilde{\mu}_t^*$ is achieved at $\tilde{\mathbf{y}}_t = \widetilde{\mu}_t(\sigma_t\sigma_t^*)^{-1}\sigma_t$. We form the Lagrangian

$$L(\lambda_t) = -\frac{1}{2}\mathbf{y}_t\mathbf{y}_t^* + \lambda_t(\sigma_t\mathbf{y}_t^* - \widetilde{\mu}_t^*), \qquad t \in [0,T],$$

where $\lambda_t = (\lambda_t^{(1)}, \lambda_t^{(2)}, \cdots, \lambda_t^{(m)})$ is the multiplier. Taking gradients with respect to the vector $\mathbf{y}_t$ and λ_t respectively, and setting them to be zero (first order condition), we have

$$\frac{\partial L(\lambda_t)}{\partial \mathbf{y}_t} = -\mathbf{y}_t + \lambda_t\sigma_t = 0, \tag{8}$$

$$\frac{\partial L(\lambda_t)}{\partial \lambda_t} = (\sigma_t\mathbf{y}_t^* - \widetilde{\mu}_t^*)^* = 0. \tag{9}$$

From Eq. (8), $\mathbf{y}_t = \lambda_t \sigma_t$. Substituting this into Eq. (9), we have $\lambda_t \sigma_t \sigma_t^* - \widetilde{\mu}_t = 0$, which implies $\lambda_t = \widetilde{\mu}_t(\sigma_t \sigma_t^*)^{-1}$. Substituting this back into Eq. (8), we obtain $\mathbf{y}_t = \widetilde{\mu}_t(\sigma_t \sigma_t^*)^{-1}\sigma_t$, or equivalently, $\mathbf{y}_t^* = \sigma_t^*(\sigma_t \sigma_t^*)^{-1}\widetilde{\mu}_t^*$, $t \in [0, T]$. Note that the Hessian of $-\frac{1}{2}\mathbf{y}_t \mathbf{y}_t^*$ is $-I$(I is the $n \times n$ identity matrix) which is negative definite, the maximum is attained at the above specified $\tilde{\mathbf{y}}_t$.

Remark 1. In the above theorem, if we take $\pi_t^{(i)} = 0$, $\forall\ t \in [0, T]$, $i = 1, 2, \cdots, m$, then the numeraire is $X_t^\pi = B_t$ and $\mathrm{P}_{\pi, \tilde{\mathbf{y}}}$ is the minimal martingale measure in the sense of Föllmer and Schweizer (1991); if we take $\pi_t = \widetilde{\mu}_t(\sigma_t \sigma_t^*)^{-1}$, then according to Remark 2 after Theorem 1, $\{X_t^\pi\}$ is the growth-optimal wealth process and $\mathrm{P}_{\pi, \tilde{\mathbf{y}}}$ is the objective probability measure P. According to Theorem 1, all these arbitrage pricing systems associated to this same market price of risk $\tilde{\mathbf{y}}$ are equivalent.

Remark 2. Note that

$$\mathrm{E}^{\mathrm{P}}\left(\log \frac{d\mathrm{P}}{d\mathrm{P}_{\pi,\mathbf{y}}}\right) = -\mathrm{E}^{\mathrm{P}}\left(\log \frac{d\mathrm{P}_{\pi,\mathbf{y}}}{d\mathrm{P}}\right),$$

and $\mathrm{E}^{\mathrm{P}}\left(\log \frac{d\mathrm{P}}{d\mathrm{P}_{\pi,\mathbf{y}}}\right)$ is the relative entropy (Kullback-Leibler information) of P w.r.t. $\mathrm{P}_{\pi,\mathbf{y}}$, which measures a certain distance between P and $\mathrm{P}_{\pi,\mathbf{y}}$. Thus the intuitive interpretation of Theorem 2 is that if we want to keep the equivalent martingale measure as close (in the sense of relative entropy) as possible to the objective probability P, we should choose $\mathrm{P}_{\pi,\mathbf{y}}$ with $\mathbf{y} = \{\mathbf{y}_t = \widetilde{\mu}_t(\sigma_t \sigma_t^*)^{-1}\sigma_t\}_{t \in [0,T]}$.

References

1. Artzner, P.(1997): On the numeraire portfolio, in A. H. Dempster and P. Pliska (eds.), Mathematics of Derivative Securities, pp. 53-57, Cambridge Univ. Press, London.

2. Bajeux-Besnainou I. Portait, R.(1997): The numeraire portfolio, a new perspective on financial theory, European Journal of Finance **3**, 291-309.

3. Delbaen, F. and W. Schachermayer(1994): A general version of the fundamental theorem of asset pricing, Math. Annalen., **300**, 463-520.

4. Delbaen, F. and W. Schachermayer(1995): Arbitrage possibilities in Bessel processes and their relation to local martingales, Probability Theory and Related Fields, **102**, 357-366.

5. Delbaen, F. and W. Schachermayer(1998a): The fundamental theorem of asset pricing for unbounded stochastic processes, Math. Annalen., **312**, 215-250.

6. Delbaen, F., Schachermayer, W.(1998b): The variance-optimal martingale measure for continuous processes, Bernoulli **2**, 81-105 .

7. Duffie, D.(1994): Dynamic Asset Pricing Theory, Second edition, Princeton: Princeton University Press.

8. Föllmer, H., Schweizer, M.(1991): Hedging of contingent claims under incomplete information, in M. H. A. Davis and R. J. Elliot (eds.), Applied Stochastic Analysis, pp. 389-414, Gordon and Breach, London.

9. Föllmer, H., Sondermann, D.(1986): Hedging of non-redundant contingent claims, in W. Hildbrand and A. Mas-Colell (eds), Contributions to Mathematical Economics, pp. 202-223, Amsterdan, North-Holland.

10. Geman, H., El Karoui, N., Rochet, J.(1995): Changes of numeraire, changes of probability and option pricing, Journal of Applied Probability **32**, 443-458.

11. Harrison, M. J., Kreps, D. M.(1979): Martingales and arbitrage in multiperiod securities markets, Journal of Economic Theory **20**, 381-408.

12. Harrison, M. J., Pliska, S. R.(1981): Martingales and stochastic integrals in the theory of continuous trading, Stochastic Processes and their Applications **11**, 215-260.

13. Hofmann, N., Platen, E., Schweizer, M. (1992): Option pricing under incompleteness and stochastic volatility, Mathematical Finance **2**, 153-197.

14. Karatzas, I., Shreve, S.(1999): Methods of Mathematical Finance, Springer-Verlag, Berlin.

15. Long, J. (1990): The numeraire portfolio, J. Fin. Econ. **26**, 29-69.

16. Ross, S. (1976): An arbitrage theory of capital asset pricing, Journal of Economic Theory **13**, 341-360.

17. Xia, J. M., Yan, J. A. (2001): Clarifying some basic concepts and results in arbitrage pricing theory, Preprint.

18. Yan, J. A.(1998): A new look at the fundamental theorem of asset pricing, J. Korean Math. Soc., **35**(3), 659-673.

19. Yan, J. A., Zhang Q., Zhang, S. G.(2000): Growth optimal portfolio in a market griven by a jump-diffusion-like process or a Lévy process, Annals of Economics and Finance, **1**(1), 101-116.

This page is intentionally left blank

Participants

Bagchi, Arunabha
Faculty of Mathematical Sciences University Twente, P.O.Box 217, 7500AE Enschede, The Netherlands
email: A.Bagchi@math.utwente.nl

Bielecki, Tomasz R.
Mathematics Department, The Northeastern Illinois University, 5500 North St. Louis Avenue, Chicago, IL 60625-4699, USA
email: T-Bielecki@neiu.edu

Chen, Zengjing
Department of Mathematics, Shandong University, Jinan 250100, China
email: chzengjing@hotmail.com

Frittelli, Marco
Dipartimento Metodi Quantitativi, Facolte di Economia, Universite di Milano - Bicocca, Piazza dell'Ateneo Nuovo 1, 20126, Milano, Italy
email: marco.frittelli@unimib.it

Guo, Xin
P.O. Box 218, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA
email: xinguo@us.ibm.com

Hu, Yaozhong
Department of Mathematics, University of Kansas, 405 Snow Hall, Lawrence, KS 66045-2142, USA
email: hu@math.ukans.edu

Li, Xunjing
Department of Mathematics, Fundan University, Shanghai 200433, China
email: xjli@fudan.edu.cn

Liu, Hong
Olin School of Business, Washington University, St. Louis, MO 63130, USA
email: LIUH@olin.wustl.edu

Ma, Jin
Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
email: majin@math.purdue.edu

Nagai, Hideo
Department of Mathematical Science, Graduate School of Engineering Science, Osaka University, Toyonaka, 560-8531, Japan
email: nagai@sigmath.es.osaka-u.ac.jp

Peng, Shige
Department of Mathematics, Shandong University, Jinan 250100, China
email: peng@sdu.edu.cn

Platen, Eckhard
University of Technology Sydney, School of Finance & Economics and Department of Mathematical Sciences, PO Box 123, Broadway, NSW, 2007, Australia
email: eplaten@uts.edu.au

Pliska, Stanley
Department of Finance, University of Illinois at Chicago, 601 S. Morgan Street, Chicago, IL 60607-7124, USA
email: srpliska@uic.edu

Runggaldier, Wolfgang
Dipartimento di Matematica Pura ed Applicata, Università di Padova, 7 Via Belzoni, I - 35131Padova, Italy
email: runggal@math.unipd.it

Rutkowski, Marek
Faculty of Mathematics and Information Science, Politechnika Warszawska, pl. Politechniki 1, 00-661 Warszawa, Poland
email: markrut@mini.pw.edu.pl

Savine, Antoine
Asia Group, BNP-Paribas, 18th floor, Tokyo Sankei Building, 1-7-2 Otemachi, Chiyoda-Ku, Tokyo 100-0004, Japan
email: antoine.savine@japan.bnpparibas.com

Sondermann, D.
University of Bonn, Department of Economics, Adenauer Allee 24, 53113 Bonn, Germany
email: sonderma@finasto.uni-bonn.de

Stettner, Lukasz
Institute of Mathematics, Polish Academy of Sciences, Sniadeckich 8, 00-950 Warsaw, Poland
email: stettner@panim.impan.gov.pl

Sun, Yeneng
Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543
email: matsuny@nus.edu.sg

Tang, Shanjian
Department of Mathematics, Fudan University, Shanghai 200433, China
email: sjtang@online.sh.cn

Vorst, Ton
Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
email: vorst@few.eur.nl

Yan, Jia-an
Institute of Applied Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China
email: jayan@amath7.amt.ac.cn

Yang, Hailiang
Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong, China
email: hlyang@hkusua.hku.hk

Yao, David
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong, China
email: yao@se.cuhk.edu.hk

Yin, George
Department of Mathematics, Wayne State University, Detroit, MI 48202, USA
email: gyin@math.wayne.edu

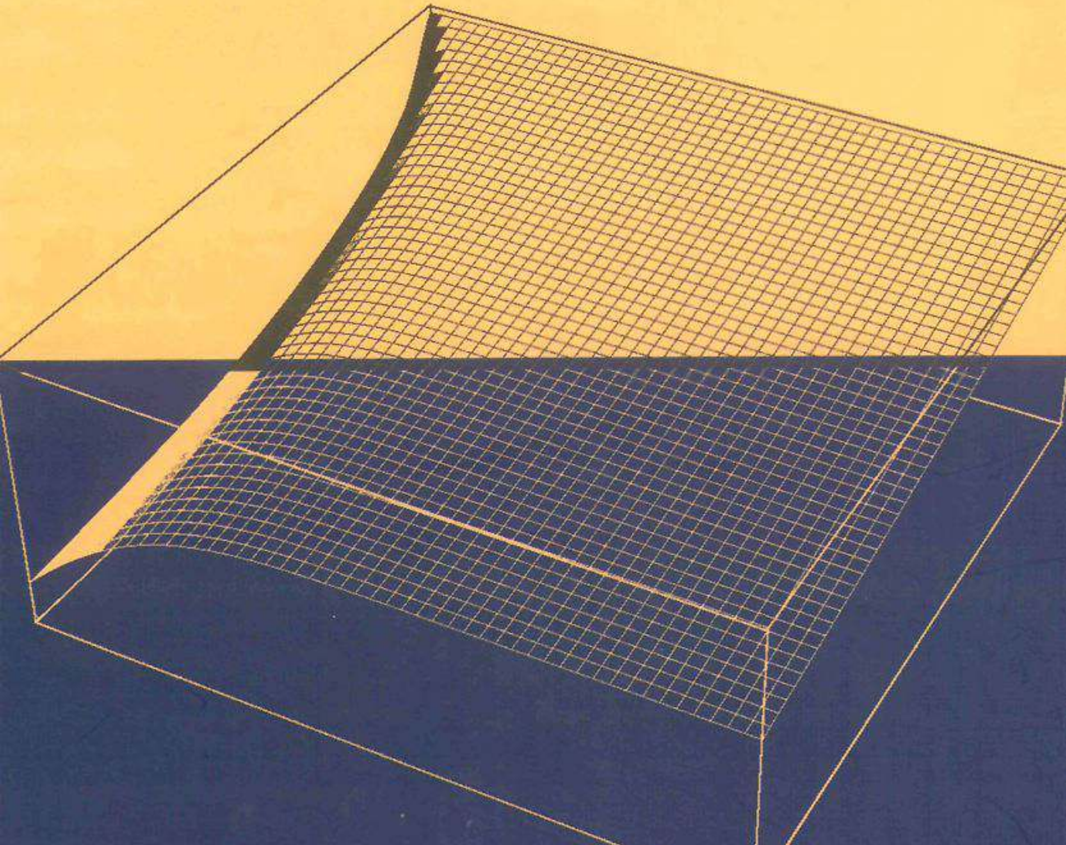
Yong, Jiongmin
Department of Mathematics, Fundan University, Shanghai 200433, China
email: jyong@fudan.edu.cn

Zhang, Qing
Department of Mathematics, University of Georgia, Athens, GA 30602, USA
email: qingz@control.math.uga.edu

Zhang, Qiang
Department of Mathematics, and Department of Economics and Finance City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong, China
email: mazq@math.cityu.edu.hk

Zhang, Shuzhong
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong, China
email: zhang@se.cuhk.edu.hk

Zhou, Xunyu
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong, China
email: xyzhou@se.cuhk.edu.hk



ISBN 981-02-4797-4

9 789810 247973

www.worldscientific.com

4852 hc