KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

ACTUARIAL AND FINANCIAL MATHEMATICS CONFERENCE

Interplay between Finance and Insurance

February 7-8, 2013

Michèle Vanmaele, Griselda Deelstra, Ann De Schepper, Jan Dhaene, Wim Schoutens, Steven Vanduffel & David Vyncke (Eds.)

CONTACTFORUM

KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

ACTUARIAL AND FINANCIAL MATHEMATICS CONFERENCE

Interplay between Finance and Insurance

February 7-8, 2013

**Michèle Vanmaele, Griselda Deelstra, Ann De Schepper,
Jan Dhaene, Wim Schoutens, Steven Vanduffel & David Vyncke (Eds.)**

CONTACTFORUM

Handelingen van het contactforum "Actuarial and Financial Mathematics Conference. Interplay between Finance and Insurance" (7-8 februari 2013, hoofdaanvrager: Prof. M. Vanmaele, UGent) gesteund door de Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten.

Afgezien van het afstemmen van het lettertype en de alinea's op de richtlijnen voor de publicatie van de handelingen heeft de Academie geen andere wijzigingen in de tekst aangebracht. De inhoud, de volgorde en de opbouw van de teksten zijn de verantwoordelijkheid van de hoofdaanvrager (of editors) van het contactforum.

KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

Paleis der Academiën
Hertogsstraat 1
1000 Brussel

D/2013/0455/07
ISBN 978 90 6569 123 1

KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

Actuarial and Financial Mathematics Conference
Interplay between finance and insurance

CONTENTS

Invited talk

Contributed talk

Poster session

KONINKLIJKE VLAAMSE ACADEMIE VAN BELGIE
VOOR WETENSCHAPPEN EN KUNSTEN

Actuarial and Financial Mathematics Conference
Interplay between finance and insurance

PREFACE

In 2013, our two-day international "Actuarial and Financial Mathematics Conference" was organized in Brussels for the sixth time. As for the previous editions, we could use the facilities of the Royal Flemish Academy of Belgium for Science and Arts. The organizing committee consisted of colleagues from 6 Belgian universities, i.e. the University of Antwerp, Ghent University, the KU Leuven and the Vrije Universiteit Brussel on the one hand, and the Université Libre de Bruxelles and the Université Catholique de Louvain on the other hand. Next to 9 invited lectures, there were 7 selected contributions as well as an extensive poster session. Just as in the previous years, we could welcome renowned international speakers, both from academia and from practice, and we could rely on leading international researchers in the scientific committee.

There were 124 registrations in total, with 71 participants from Belgium, and 53 participants from 17 other countries from all continents. The population was mixed, with 71% of the participants associated with a university (PhD students, post doc researchers and professors), and with 29% working in the banking and insurance industry.

On the first day, February 7, we had 8 speakers, among them 5 international and eminent invited speakers, alternated with 3 contributions selected by the scientific committee.
In de morning, the first speaker was *Prof.dr. Fred Espen Benth*, from *the University of Oslo (Norway)*, on "Pricing and hedging average-based options in energy markets"; afterwards *Prof.dr. Klaus Reiner Schenk-Hoppé, University of Leeds (UK)* gave an interesting talk about "Costs and benefits of speculation: On the equilibrium effects of financial regulation". These two lectures were followed by 2 presentations by researchers from Germany and France.
In the afternoon, we heard *Prof.dr. Martino Grasselli, Università degli Studi di Padova (Italy)*, who presented new research results about "Smiles all around: FX joint calibration with and without risk neutral measure", *Prof.dr. Emmanuel Gobet, Ecole Polytechnique (France)*, with a paper "Almost sure optimal hedging strategy", and *Prof.dr. Dilip Madan, Robert H. Smith School of Business, University of Maryland (USA)*, with a clear review lecture entitled "A theory of risk for two price market equilibria". In addition, there was an extra selected contribution by a young Belgian researcher.

During the lunch break, we organized a poster session, preceded by a poster storm session, where the 17 different posters were introduced very briefly by the researchers. The posters attracted a great deal of interest, judging by the lively interaction between the participants and the posters' authors. The posters remained in the central hall during the whole conference, so that they could be consulted and discussed during the coffee breaks.

Also on the second day, February 8, we had 8 lectures, with 4 keynote speakers and 4 selected contributions. The first speaker was *Prof.dr. Antoon Pelsser, Maastricht University & Kleynen Consultants (the Netherlands)*, with a paper on "Convergence results for replicating portfolios". Afterwards, *Prof.dr. Claudia Czado, Technische Universität München (Germany)* presented her newest results on "Vine copulas and their applications to Financial data". In the afternoon, we could listen to *Prof.dr. Antje Mahayni, Universität Duisburg-Essen (Germany)*, about "Evaluation of optimized proportional portfolio insurance strategies". Finally, *Prof.dr. Uwe Schmock, Vienna University of Technology (Austria)* had the floor, with a well-received lecture "Approximation and aggregation of risks by variants of Panjer's recursion". The other 4 presentations were again selected from a large number of submissions by the scientific committee; the speakers came from Canada, the Netherlands, Great Britain and Germany.

The proceedings contain two articles related to invited and contributed talks, and nine extended abstracts of poster presenters of the poster sessions, giving an overview of the topics and activities at the conference.

We are much indebted to the members of the scientific committee, *Hansjoerg Albrecher (University of Lausanne, Switzerland), Freddy Delbaen (ETH Zürich, Switzerland), Michel Denuit (Université Catholique de Louvain, Belgium), Jan Dhaene (Katholieke Universiteit Leuven, Belgium), Ernst Eberlein (University of Freiburg, Germany), Monique Jeanblanc (Université d'Evry Val d'Essonne, France), Ragnar Norberg (SAF, Université Lyon 1, France), Steven Vanduffel (Vrije Universiteit Brussel, Belgium), Michel Vellekoop (University of Amsterdam, The Netherlands), Noel Veraverbeke (University Hasselt, Belgium) and the chair Griselda Deelstra (Université Libre de Bruxelles, Belgium)*. We appreciate their excellent scientific support, their presence at the meeting and their chairing of sessions. We also thank Wouter Dewolf (Ghent University, Belgium), for the administrative work.
We are very grateful to our sponsors, namely the *Royal Flemish Academy of Belgium for Science and Arts, the Research Foundation – Flanders (FWO), the Scientific Research Network (WOG) "Stochastic modelling with applications in finance", le Fonds de la Recherche Scientifique (FNRS), Cambridge Springer, KBC Bank en Verzekeringen, and the BNP Paribas Fortis Chair in Banking at the Vrije Universiteit Brussel and Université Libre de Bruxelles*. Without them it would not have been possible to organize this event in this very enjoyable and inspiring environment.

The continuing success of the meeting encourages us to go on with the organization of this contact-forum, in order to create future opportunities for exchanging ideas and results in this fascinating research field of actuarial and financial mathematics.

The editors:
Griselda Deelstra
Ann De Schepper
Jan Dhaene
Wim Schoutens
Steven Vanduffel
Michèle Vanmaele
David Vyncke

The other members of the organising committee:
Jan Annaert
Pierre Devolder

INVITED TALK

STOCHASTIC MODELLING OF POWER PRICES BY VOLTERRA PROCESSES

Fred Espen Benth

Center of Mathematics for Applications (CMA), University of Oslo, PO Box 1053 Blindern, N-0316 Oslo, Norway
Email: `fredb@math.uio.no`

Abstract

We propose a Volterra process driven by an independent increment process as the basic model for the spot price dynamics of power. This class will encompass most of the existing models, like Lévy-driven Ornstein-Uhlenbeck and continuous-time autoregressive moving average processes. The rich structure of the model will allow for explaining most of the stylized facts of power prices like seasonality, mean reversion and spikes. We derive the forward price dynamics for contracts with delivery of power over a period, using the Esscher transform to construct a pricing measure. Finally, the risk premium is discussed, and we show that the time-inhomogeneity together with a time-varying market price of risk can yield a change in sign in the risk premium.

1. INTRODUCTION

Power markets have been liberalized world-wide in the recent decades, and there is a demand for sophisticated modelling tools for pricing and risk management purposes. The power markets have their distinct characteristics, making modelling and pricing challenging tasks. In this paper we develop further some modelling concepts that have proven fruitful in electricity and related markets like weather and gas.

Stationarity is the key property of prices in power markets, at least after explaining the seasonal trends and long-term effects. Barndorff-Nielsen et al. (2013) suggest the class of Lévy semistationary models for the power spot price dynamics. We develop this class further, and consider Volterra dynamics driven by a so-called independent increment process. An independent increment process can be viewed as a time-inhomogeneous Lévy process, that is, a process where the increments are independent but not necessarily stationary. This opens for modelling seasonally varying spike intensities, for example.

We apply the Esscher transform to construct a pricing measure when analysing the problem of deriving forward prices. The spot price of power is not tradeable in the usual sense, as one cannot store this commodity. Hence, any pricing measure is only required to be equivalent to the market probability, and not turning the discounted spot into a martingale. The forward price is defined as the conditional expectation of the spot under this pricing measure, yielding arbitrage-free martingale dynamics.

We analyse the risk premium defined as the difference between the forward price and the expected spot at delivery in this context. In classical commodity markets, the risk premium is negative since producers have hedging needs to insure their future revenues. However, in power markets there are ample reasons for the occurrence of a positive premium, stemming from the fact that the opposite side of the producers, the retailers, also have hedging needs, in particular when there is a high chance for excessive prices resulting from spikes. We show that it is possible to accommodate for this in our set-up, due to the time-inhomogeneity of the driving noise process and the possibility to have a time-varying market price of risk being the parameter in the Esscher transform.

2. THE SPOT PRICE DYNAMICS

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T^*]}, P)$ be a complete filtered probability space, where $T^* < \infty$ is a finite time horizon for the market we model. We introduce the class of *independent increment* (II) processes as follows:

Definition 2.1 *An adapted RCLL*[1] *process $I(t)$ starting in zero is called an II-process if it satisfies the following two conditions:*

1. *The increments $I(t_0), I(t_1) - I(t_0), \dots, I(t_n) - I(t_{n-1})$ are independent random variables for any partition $0 \leq t_0 < t_1 < \dots < t_n \leq T^*$, and $n \geq 1$ a natural number.*

2. *It is continuous in probability.*

If an II-process $I(t)$ satisfies stationarity of the increments, that is, if $I(t) - I(s)$ has the same distribution as $I(s)$ for any $0 \leq s < t \leq T^*$, then I is a Lévy process (see Cont and Tankov (2004)).

The characteristic function of an increment of the II-process $I(t)$ can be expressed as

$$\psi(s,t;\theta) = \ln \mathbb{E}\left[\exp\left(\mathrm{i}\theta(I(t) - I(s))\right)\right] , \tag{1}$$

for $0 \leq s < t \leq T^*$, $\theta \in \mathbb{R}$, and

$$\psi(s,t;\theta) = \mathrm{i}\theta(\gamma(t) - \gamma(s)) - \frac{1}{2}\theta^2(C(t) - C(s)) + \int_s^t \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}\theta z} - 1 - \mathrm{i}\theta z 1_{|z|\leq 1}\right) \ell(dz, du) . \tag{2}$$

Here,

1. $\gamma : [0,T^*] \mapsto [0,T^*]$ is a continuous function with $\gamma(0) = 0$,

2. $C : [0,T^*] \mapsto [0,T^*]$ is a non-decreasing and continuous function with $C(0) = 0$,

[1]RCLL is short-hand for right-continuous with left-limits.

3. ℓ is a σ-finite measure on the Borel σ-algebra of $[0,T^*]\times\mathbb{R}$, with the properties

$$\ell(A\times\{0\})=0\,,\quad \ell(\{t\}\times\mathbb{R})=0\,,\quad \text{for } t\in[0,T^*] \text{ and } A\in\mathcal{B}([0,T^*])$$

and

$$\int_0^t\int_{\mathbb{R}}\min(1,z^2)\,\ell(dz,ds)<\infty\,.$$

We call ψ the *cumulant function* of I, and (γ,C,ℓ) the *generating triplet*, where γ is the drift, C is the variance process and ℓ is the compensator measure. If $I(t)$ is a Lévy process, we have that $\psi(s,t;\theta)=\widetilde{\psi}(\theta)(t-s)$, $\gamma(t)=\gamma t$, $C(t)=ct, c\geq 0$ and $\ell(dz,ds)=ds\widetilde{\ell}(dz)$. In this case, $\widetilde{\ell}(dz)$ is called the Lévy measure and $\widetilde{\psi}(\theta)$ is known as the Lévy exponent. We restrict our attention to the case when I is a square-integrable semimartingale, that is, when $\int_0^t\int_{\mathbb{R}}z^2\,\ell(dz,ds)<\infty$ for all $t\in[0,T^*]$ and γ is of bounded variation.

For a given II-process I we define the spot price dynamics as

$$S(t)=\Lambda(t)+X(t)\,, \tag{3}$$

where $\Lambda:[0,T^*]\mapsto\mathbb{R}_+$ is a bounded deterministic function modelling the seasonal mean, while X is the Volterra process

$$X(t)=\int_0^t g(s,t)\,dI(s)\,. \tag{4}$$

Here we have g given as a deterministic function defined on the half-space $\{(s,t)\in[0,T^*]^2\,:\,s\leq t\}$ and being square integrable with respect to C and ℓ, i.e., for every $t\in[0,T^*]$,

$$\int_0^t g^2(s,t)\,dC(s)<\infty\quad\int_0^t\int_{\mathbb{R}}g^2(s,t)z^2\ell(dz,ds)<\infty\,.$$

This integrability condition ensures that X is well-defined as a stochastic integral with respect to the square-integrable semimartingale I (see Protter (1990)).

The spot dynamics S in (3) with X as in (4) covers many of the classical models. First of all, choosing a so-called arithmetic structure is reasonable in energy markets as argued empirically by Bernhardt et al. (2008). For example, in the German power market EEX one observes frequently negative prices in the spot market explained by unexpectedly high production of unregulated wind power. Geometric models will not manage to explain such price behaviour, while an arithmetic structure can account for this if X can turn negative. The seasonality function Λ models the mean level of prices, which typically vary over the year with high prices in the winter due to added demand for heating, and lower prices in the warmer seasons. Also, there are weekend and intraday effects in the power market, with higher prices during the day than in the night time, and in the working week compared to the weekend. Such deterministic mean effects are explained by the function Λ.

Coming back to the factor X describing the stochastic evolution of the prices S, the simplest case is $I=B$, a Brownian motion, and $g(s,t)=\exp(-\alpha(t-s))$. In this case, X is a classical Ornstein-Uhlenbeck process which is the standard choice of modelling the dynamics of a commodity (see Benth et al. (2008)). As this choice leads to prices which are normally distributed, we will not be able to explain the observed heavy tails in power prices (see Benth et al. (2008)), which calls for Lévy processes as the modelling device for the stochastic drivers. However, one of the reasons for non-Gaussian spot price dynamics is the price spikes commonly observed in power markets. These spikes occur in periods with imbalances between supply and demand for power, for example arising when there is a sudden drop in temperature leading to an unexpected high demand for heating. A spike is characterized by a huge price increase of several magnitudes over a short period of time (within a day, say), followed by a rapid reversion back to ”normal levels”. The reversion is due to the market's immediate reaction to high prices by reducing the demand. In the Nordic power market NordPool, such spikes are mostly occurring during the cold season, and almost never in the summer period. Hence, it is reasonable to imagine a stochastic driver I which can jump depending on the season, with a high probability of a price spike in the winter, and low in the summer. This can be achieved by choosing I to be an II-process. The function g takes care of the reversion, with α in the Ornstein-Uhlenbeck case being directly interpretable as the *speed of mean reversion*. Combined with such a choice of g, we could for example choose I to be a compound Poisson process with time-dependent jump intensity, that is

$$I(t) = \sum_{i=1}^{N(t)} J_i \,,$$

where J_i are *iid* random variables and N is a time-inhomogeneous Poisson process with jump intensity described by $\lambda(t)$. Here, λ is a positive continuous function on $[0, T^*]$. In Geman and Roncoroni (2006) such a jump process is applied to a jump-diffusion spot price model estimated by data from several power markets.

In Barndorff-Nielsen et al. (2013) so-called Lévy (semi-)stationary processes are suggested for the spot price dynamics. By selecting I to be a Lévy process and $g(s,t) = g(t-s)$, we obtain a class of processes X which are stationary (under some mild additional technical assumption on the Lévy measure of I, see Sato (1999)). Note that in the Ornstein-Uhlenbeck case, g has the required structure to ensure stationarity. Empirical analysis of spot prices at the German EEX market reveals that other choices of g are more reasonable. For example, in the papers Benth et al. (2011) and Bernhardt et al. (2008) it is argued that choosing g from the class of CARMA processes is preferable from a statistical point of view. CARMA is short-hand for *continuous time autoregressive moving average*, and yields a $g(s,t) = g(t-s)$ where

$$g(x) = \mathbf{b}^{\mathrm{T}} e^{Ax} \mathbf{e}_p$$

for $x \geq 0$ and $\mathbf{e}_p$ is the pth unit vector in $\mathbb{R}^p$, $p \in \mathbb{N}$. Here, A is a $p \times p$ matrix of the form

$$A = \begin{bmatrix} \mathbf{0}_{p-1} & I_{p-1} \\ -\alpha_p \ldots & -\alpha_1 \end{bmatrix}$$

with $\alpha_p, \alpha_{p_1}, \ldots, \alpha_1$ being positive constants, $\mathbf{0}_{p-1}$ the $p-1$-dimensional vector of zeros and I_{p-1} the $(p-1) \times (p-1)$ identity matrix. Finally, $\mathbf{b} = (b_0, b_1, \ldots, b_{p-1})^{\mathrm{T}} \in \mathbb{R}^p$ is the vector with coordinates such that $b_q = 1$ and $b_j = 0$ for $q < j \leq p$. We say that p is the autoregressive order while q is the moving average order. As it turns out, $p = 2, q = 1$ is a good choice in the case of EEX daily spot data. The empirical analysis in Barndorff-Nielsen et al. (2013) suggests other choices of g as well, including the model of Bjerksund et al. (2010) which corresponds to choosing

$$g(x) = \frac{a}{b+x}$$

for positive constants a and b.

As suggested in Benth et al. (2011), it might be reasonable to consider a model with several factors and not only one modelling the stochasticity of the spot price. For example, one may have a non-stationary term $X_1(t)$ with a constant $g(s,t)$ and a stationary $X_2(t)$ with a kernel function g being of CARMA-form. Letting the driving II processes be of Lévy type, we are in the situation studied by Benth et al. (2011).

In this paper we propose a general one-factor model with the flexibility to account for time-inhomogeneous stochastic drivers I and general specifications of g. We want to derive forward prices for this general model, and analyse the implied risk premium.

3. FORWARD PRICING

In power markets, the spot is a physical commodity in the sense that if you have entered a long or short position, there will be a physical transmission of electricity over a given hour. Hence, by the very nature of electricity, it cannot be stored and therefore not traded as a classical commodity or financial asset. Thus, when pricing forwards on power, we cannot resort to the classical buy-and-hold strategy, which prescribes the forward price to be the cost of carrying the spot forward (see Geman (2005)). On the other hand, the forward contracts are typically financial, that is, one pays or receives the money-equivalent of the spot over a given period. If, for example, one has bought a forward on power with delivery in the time period $[T_1, T_2]$, being a specific month, say, then one receives

$$\int_{T_1}^{T_2} S(T)\, dT$$

in return of paying the agreed forward price at time T_2. From this example we see another characteristic feature of power markets, namely that forward contracts do not deliver at fixed times, but over a specific period. In the market, these delivery periods are typically specific weeks, months, quarters of years. As the forwards are financial contracts, one may use them for speculation, and they can be traded in a portfolio. Hence, the arbitrage theory of mathematical finance says that the forward price must be a (local) martingale with respect to some pricing measure Q being equivalent to P. Observe that Q is *not* an equivalent martingale measure, in the sense that the discounted spot price becomes a (local) Q martingale. If we denote the forward price at time t for a contract with delivery over the period $[T_1, T_2]$, $T_1 < T_2, t \leq T_2$ by $F(t, T_1, T_2)$, then we have

$$F(t, T_1, T_2) = \mathbb{E}_Q \left[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} S(T)\, dT \mid \mathcal{F}_t \right] . \tag{5}$$

The reason for taking the conditional expectation of the *average* spot price is that the forward price is denominated in terms of currency per mega Watt hours (Euro/MWh, say).

Entering forward agreements can be viewed as an insurance on the underlying commodity price. A producer, say, locks in the price of her production by selling it in the forward market. When the underlying commodity is power, it cannot as already discussed be traded. Hence, we are in a situation where we want to assign a premium on an ”insurance” on the price of power, which cannot be hedged. It is customary to choose a class of pricing probabilities based on the so-called Esscher transform, a technique adopted from insurance mathematics (see Gerber and Shiu (1994)).

To this end, let $\theta : [0,T^*] \mapsto \mathbb{R}$ be a function which is integrable with respect to I, and suppose that

$$\phi(s,t;\theta) = \ln \mathbb{E}\left[\exp\left(\int_s^t \theta(u)\,dI(u)\right)\right], \tag{6}$$

is well-defined for all $0 \leq s < t \leq T^*$. Define the martingale process $Z(t)$ for $0 \leq t \leq T^*$ by

$$Z(t) = \exp\left(\int_0^t \theta(u)\,dI(u) - \phi(0,t,\theta)\right). \tag{7}$$

Introduce a probability measure Q where the Radon-Nikodym derivative has a density given by $Z(t)$, that is,

$$\frac{dQ}{dP}\Big|_{\mathcal{F}_t} = Z(t)\,. \tag{8}$$

The probability Q is parametric in θ, which is often referred to as the *market price of risk*. This construction of a probability Q is called the Esscher transform of the process I.

From Prop. 3.1 in Benth and Sgarra (2012) it holds that I is an II-process under Q, with characteristic triplet $(\gamma_\theta, C, \ell_\theta)$, where

$$\gamma_\theta(t) = \gamma(t) + \int_0^t \int_{|z|<1} z\left(e^{\theta(u)z} - 1\right)\,\ell(dz,du) + \int_0^t \theta(u)\,dC(u)\,,$$

and

$$\ell_\theta(dz,dt) = \exp(\theta(t)z)\ell(dz,dt)\,.$$

Thus, the drift γ is shifted and the compensator measure ℓ is exponentially tilted.[2]

As ϕ in (6) is the log-moment generating function of $I(t) - I(s)$ when $\theta = x$ is chosen as a constant, we find that

$$\mathbb{E}[I(t)] = \frac{\partial}{\partial x}\phi(0,t;0)\,.$$

But from the cumulant function, for which we have the relationship $\phi(s,t;x) = \psi(s,t;-\mathrm{i}x)$, we find,

$$\frac{\partial}{\partial x}\phi(s,t;x) = \gamma(t) - \gamma(s) + x(C(t) - C(s)) + \int_s^t \int_{\mathbb{R}} z\left(e^{xz} - 1_{|z|<1}\right)\,\ell(dz,du)\,,$$

and therefore

$$\mathbb{E}[I(t)] = \gamma(t) + \int_0^t \int_{|z|\geq 1} z\,\ell(dz,ds)\,. \tag{9}$$

Thus, we can restate the spot price dynamics as

$$S(t) = \Lambda(t) + \int_0^t g(s,t)\,d\mathbb{E}[I(s)] + \int_0^t g(s,t)\,d\widetilde{I}(s) \tag{10}$$

[2] Note that there is a typo in Benth and Sgarra (2012) concerning the drift.

where $\widetilde{I}(t) = I(t) - \mathbb{E}[I(t)]$ is a martingale. Analogously, we find by referring to the characteristic triplet of $I(t)$ with respect to Q

$$\begin{aligned}\mathbb{E}_\theta[I(t)] &= \gamma_\theta(t) + \int_0^t \int_{|z|\geq 1} z\,\ell_\theta(dz,ds)\\ &= \gamma(t) + \int_0^t \theta(u)\,dC(u) + \int_0^t \int_{\mathbb{R}} z\left(e^{\theta(u)z} - 1_{|z|\leq 1}\right)\ell(dz,du)\,. \end{aligned} \tag{11}$$

Here, $\mathbb{E}_\theta[\cdot]$ denotes the expectation operator with respect to the probability Q. But then after defining the Q-martingale $\widetilde{I}_\theta(t) = I(t) - \mathbb{E}_\theta[I(t)]$ we find

$$S(t) = \Lambda(t) + \int_0^t g(s,t)\,d\,\mathbb{E}_\theta[I(s)] + \int_0^t g(s,t)\,d\widetilde{I}_\theta(s)\,. \tag{12}$$

In both the P and Q dynamics of S we see that we have the representation of a deterministic term and a stochastic factor driven by a martingale. The seasonality function can be modified to take into account the trend imposed by I in both cases.

The computation of forward prices implied by the choice of Q by the Esscher transform now becomes particularly simple. We present the result in the following proposition:

Proposition 3.1 *The forward price $F(t,T_1,T_2)$ of a contract with delivery period $[T_1,T_2]$, where $t \leq T_2$ and $T_1 < T_2 \leq T^*$, is*

$$\begin{aligned}F(t,T_1,T_2) = &\frac{1}{T_2-T_1}\int_{T_1}^{\max(t,T_1)} S(T)\,dT + \frac{1}{T_2-T_1}\int_{\max(t,T_1)}^{T_2} \Lambda(T)\,dT\\ &+ \frac{1}{T_2-T_1}\int_{\max(t,T_1)}^{T_2}\int_0^T g(s,T)\,d\mathbb{E}_\theta[I(s)]\,dT\\ &+ \int_0^t \left\{\frac{1}{T_2-T_1}\int_{\max(t,T_1)}^{T_2} g(s,T)\,dT\right\} d\widetilde{I}_\theta(s)\,.\end{aligned}$$

Proof. We must calculate

$$F(t,T_1,T_2) = \mathbb{E}_\theta\left[\frac{1}{T_2-T_1}\int_{T_1}^{T_2} S(T)\,dT\,|\,\mathcal{F}_t\right]\,.$$

By measurability, we find for $t \leq T_2$ that

$$\begin{aligned}F(t,T_1,T_2) &= \frac{1}{T_2-T_1}\int_{T_1}^{\max(t,T_1)} S(T)\,dT + \mathbb{E}_\theta\left[\frac{1}{T_2-T_1}\int_{\max(t,T_1)}^{T_2} S(T)\,dT\,|\,\mathcal{F}_t\right]\\ &= \frac{1}{T_2-T_1}\int_{T_1}^{\max(t,T_1)} S(T)\,dT + \frac{1}{T_2-T_1}\int_{\max(t,T_1)}^{T_2} \mathbb{E}_\theta\left[S(T)\,|\,\mathcal{F}_t\right]\,dT\,.\end{aligned}$$

In the last equality we applied the Fubini-Tonelli Theorem. From the Q dynamics of S in (12), we find for $t \leq T$

$$\begin{aligned}\mathbb{E}_\theta[S(T)\,|\,\mathcal{F}_t] &= \Lambda(T) + \int_0^T g(s,T)\,d\mathbb{E}_\theta[I(s)] + \mathbb{E}_\theta[\int_0^T g(s,T)\,d\widetilde{I}_\theta(s)\,|\,\mathcal{F}_t] \\ &= \Lambda(T) + \int_0^T g(s,T)\,d\mathbb{E}_\theta[I(s)] + \int_0^t g(s,T)\,d\widetilde{I}_\theta(s) + \mathbb{E}_\theta[\int_t^T g(s,T)\,d\widetilde{I}_\theta(s)] \\ &= \Lambda(T) + \int_0^T g(s,T)\,d\mathbb{E}_\theta[I(s)] + \int_0^t g(s,T)\,d\widetilde{I}_\theta(s)\,.\end{aligned}$$

Here we applied adaptedness and independent increment property of the II-process $\widetilde{I}_\theta$, as well as its martingale property with respect to Q. After applying the Fubini-Tonelli Theorem once more, this concludes the proof. ■

We observe that in general the forward price is not expressible in terms of the spot. We must in fact recover the path of $I(t)$ from the spot in order to obtain the forward price. In the case of the Ornstein-Uhlenbeck process $g(s,t) = \exp(-\alpha(t-s))$ for $\alpha > 0$, we find

$$\frac{1}{T_2 - T_1}\int_{\max(t,T_1)}^{T_2} g(s,T)\,dT = \frac{1}{\alpha(T_2 - T_1)}\left(e^{-\alpha(\max(t,T_1)-t)} - e^{-\alpha(T_2-t)}\right)e^{-\alpha(t-s)}\,,$$

and hence we can represent the forward price in terms of the current spot price $S(t)$. In general, we find that the forward price is a Volterra process as the spot, except with a different function g.

In some energy markets one can trade forwards within the delivery period, that is, the exchange organizes the trade for $t \in [T_1, T_2]$. This is not always the situation, where the trade is closed after time T_1, and from there on the parties with positions in the forwards have the wait until time T_2 for settlement. Of course, one can do trade over-the-counter.

4. RISK PREMIUM

The risk premium is defined as the difference between the forward price and the predicted average spot price over the delivery period:

$$R(t,T_1,T_2) = F(t,T_2,T_2) - \mathbb{E}\left[\frac{1}{T_2 - T_1}\int_{T_1}^{T_2} S(T)\,dT\,|\,\mathcal{F}_t\right]\,. \tag{13}$$

If we multiply this by the length of the delivery period, $(T_2 - T_1) \times R(t,T_1,T_2)$ gives the amount of money that the producer will loose in entering the forward compared to the expected revenue by selling the production on the spot market. Interpreted as such, $(T_2 - T_1) \times R(t,T_1,T_2)$ is a negative number, and can be assigned as the premium paid by the producer to insure (hedge) her production revenue by locking in the price by entering a forward contract.

The amazing fact in power markets is that the risk premium might be positive. This comes from the fact that the retailer operating on the other side of the table of the producers might also be interested in hedging. A retailer sits in a position where prices are fixed against the end-users, and they would like to hedge the possibility of excessively high prices. They can do this by buying power in the forward market, thus creating a price pressure leading to a potentially positive risk premium. This is then the premium paid by the retailers for insuring the price risk. In this case the producers (or speculators) are the ones that collect the premium and act as the insurers.

We observe from the definition of the forward price and the Fubini-Tonelli Theorem that

$$R(t,T_1,T_2)=\frac{1}{T_2-T_1}\int_{\max(t,T_1)}^{T_2}\left(\mathbb{E}_\theta[S(T)\,|\,\mathcal{F}_t]-\mathbb{E}[S(T)\,|\,\mathcal{F}_t]\right)\,dT\,. \tag{14}$$

To analyse the risk premium, it is sufficient to study the difference

$$r(t,T)=\mathbb{E}_\theta[S(T)\,|\,\mathcal{F}_t]-\mathbb{E}[S(T)\,|\,\mathcal{F}_t]\,, \tag{15}$$

for $t\leq T$, which, by inspecting Prop. 3.1, can be computed as

$$\begin{aligned}r(t,T)&=\int_0^T g(s,T)\,d\left(\mathbb{E}_\theta[I(s)]-\mathbb{E}[I(s)]\right)+\int_0^t g(s,T)\,d\left(\widetilde{I}_\theta(s)-\widetilde{I}(s)\right)\\&=\int_t^T g(s,T)\,d\left(\mathbb{E}_\theta[I(s)]-\mathbb{E}[I(s)]\right)\,.\end{aligned}$$

Here we used that $\widetilde{I}_\theta(s)-\widetilde{I}(s)=\mathbb{E}[I(s)]-\mathbb{E}_\theta[I(s)]$. But by a direct computation using the expression of the expectations in (9) and (11),

$$\mathbb{E}_\theta[I(s)]-\mathbb{E}[I(s)]=\int_0^s\theta(u)\,dC(u)+\int_0^s\int_{\mathbb{R}}z\left(e^{\theta(u)z}-1\right)\,\ell(dz,du)\,.$$

Hence, we conclude that

$$r(t,T)=\int_t^T g(s,T)\theta(u)\,dC(u)+\int_t^T\int_{\mathbb{R}}g(s,T)z\left(e^{\theta(u)z}-1\right)\,\ell(dz,du)\,. \tag{16}$$

To obtain the full risk premium for a power contract $R(t,T_1,T_2)$, one must integrate the expression $r(t,T)$ with respect to T over the interval $\max(t,T_1)$ and T_2, and divide by the length of delivery.

We next discuss the sign of $r(t,T)$. In most relevant situations, g is a positive function and we restrict our attention to this case. Let us simplify the situation further, and consider $\theta(u)\equiv\theta$, a constant. If $\theta>0$, then we know that $z(\exp(\theta z)-1)$ is positive for all $z\in\mathbb{R}$, and together with g being positive, we obtain that

$$\int_t^T\int_{\mathbb{R}}g(s,T)z\left(e^{\theta(u)z}-1\right)\,\ell(dz,du)>0\,.$$

Moreover, C is an increasing function, so the first term in $r(t,T)$ is positive as well. Hence, we find that $r(t,T)$ is positive. In conclusion, a positive market price of risk θ leads to a positive risk premium. If $\theta<0$, we find similarly

$$\int_t^T\int_{\mathbb{R}}g(s,T)z\left(e^{\theta(u)z}-1\right)\,\ell(dz,du)<0\,.$$

Moreover, the first term becomes negative as well, and thus a negative market price of risk leads to a negative risk premium. For constant market prices of risk we see therefore that the risk premium becomes either positive or negative for all delivery periods $[T_1,T_2]$.

We may accommodate a change in sign of the risk premium in the following stylized situation: If we consider an II process $I(t)$ which does not have any jump component but a covariance given by $dC(t) = \sigma^2(t)\,dt$ for some positive-valued function $\sigma : [0, T^*] \mapsto \mathbb{R}_+$. This function is scaling the noise driving the factor X, and we can imagine a situation where this is seasonal. For example, we could mimic volatile prices in the winter, and more stable prices in the summer which is the situation in the Nordic power market, say, by letting σ be big in the winter, and small in the summer. Since the seasonal function Λ is low for summer as well, the prices will have a relatively small variation around the mean, and one could imagine that the producers in this case would impact the market with a hedging pressure as the retailers are relatively certain about their prices. Hence, choosing θ as a function with negative values in the summer seems reasonable. On the other hand, during winter one may choose θ to be positive as high volatility may yield excessively high prices, that the retailers want to avoid by hedging in forwards. As in this situation we have chosen $\ell(dz, ds) = 0$, the risk premium $r(t, T)$ becomes

$$r(t,T) = \int_t^T g(s,T)\theta(s)\sigma^2(s)\,ds\,.$$

We find this to be negative when t and T are times during summer, while t, T in the winter would yield positive values of r. However, if t is in the summer, and T goes into the winter period, we might get a situation where the premium r is changing sign from negative to positive.

Another example of a similar situation, which might be more relevant, is when $C = 0$ and we have a pure-jump II process $I(t)$. Imagine that $I(t)$ is a compound Poisson process with a time-inhomogeneous jump intensity. We find that

$$\ell(dz, ds) = \lambda(s)F_J(dz)\,ds\,,$$

where F_J is the distribution of the jump size J, and $\lambda : [0, T^*] \mapsto \mathbb{R}_+$ is the jump intensity. Recall that in the NordPool market, it is more likely to have big price spikes during winter time than in the summer time. Hence, we could have λ small in the summer and big during winter. By choosing θ as a function being negative during summer and positive during winter, we can obtain the same situation as for the case of no jumps above. For this example, we mimic a market where retailers take into account the excessive jump risk during winter. It is to be noted that this model probably would require more factors to describe the price dynamics during summer more accurately, since low λ implies few jumps, and therefore essentially a deterministic price path. We refer to Benth and Sgarra (2012) for more on the change of sign of the risk premium in power markets. Note that we manage to achieve such a sign change due to the time-inhomogeneity of I.

5. CONCLUSIONS AND OUTLOOK

We have discussed the basic models for the spot price dynamics in power markets. Considering the stylized facts of power spot prices, Volterra processes driven by independent increment processes provide a natural modelling class. Based on such a class, which encompasses many of the classical models like Lévy-driven Ornstein-Uhlenbeck and continuous-time autoregressive moving average processes, we derive the forward price dynamics for contracts delivering over a period. This is the situation for forward contracts written on electricity, which naturally cannot be settled at a fixed delivery time. As the situation in power is similar to an insurance context, since the underlying spot cannot be traded in a financial sense, we apply the Esscher transform to construct a pricing measure when analysing the forwards. Finally, we showed that the spot model can accommodate a change in sign of the risk premium in the forward market, a result achieved by appealing to the time-inhomogeneity of the driving noise and a time-varying market price of risk. We explained such a change from the opposite hedging needs of retailers and producers in the power market.

We show that the forward price dynamics are expressible in terms of a Volterra process driven by the same independent increment process as the spot, however, with a different integrand function. In general it is not possible to express the forward in terms of the current spot. However, in some situations one may recover the forward as a function of the path of the spot, see Benth and Solanilla Blanco (2012).

European call and put options are traded in the power exchanges in the Nordic NordPool market and the German EEX market. These are written on the forwards as underlying. Furthermore, spread options between different power markets, and also between different commodities like gas and power, coal and power are traded over-the-counter. By appealing to transform-based pricing methods, using the explicit knowledge of the cumulant function of I, one can derive pricing formulas which can be calculated efficiently on a computer (see Benth and Zdanowicz (2013)). Other relevant derivatives include Asian-style options on the spot, which actually were traded at NordPool around the year 2000. Benth et al. (2013) have developped an efficient algorithm for pathwise simulation of Lévy semistationary processes, an interesting subclass of the Volterra processes studied here. Such simulation algorithms have clear applications to Monte Carlo pricing of path-dependent options in power markets.

Finally, the issue of hedging these derivatives is of course relevant. In a forthcoming paper by Benth and Dethering (2013) quadratic hedging has been analysed in situations where you cannot trade the underlying all the way up to the exercise date. This is the relevant situation in power markets when hedging an Asian option on the spot, using electricity forwards to hedge.

Acknowledgements The Norwegian Research Council is gratefully acknowledged for the financial support of the two projects ”Managing Weather Risk in Electricity Markets (MAWREM)” and ”Energy Markets: Modelling, Optimization and Simulation (EMMOS)”, grants RENERGI484264 and EVITA205328, resp.

References

O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Modelling energy spot prices by volatility modulated volterra processes. *Bernoulli*, 2013. To appear.

C. Benth, F.E:.and Klüppelberg, G. Müller, and L. Vos. Futures pricing in electricity markets based on stable carma spot models. *Submitted*, 2011.

F.E. Benth and N. Dethering. Quadratic hedging of power derivatives. *In preparation*, 2013.

F.E. Benth and C. Sgarra. The risk premium and the Esscher transform in power markets. *Stochastic Analysis and Applications*, 30:20–43, 2012.

F.E. Benth and S.A. Solanilla Blanco. Forward prices in markets driven by continuous-time autoregressive processes. *Submitted*, 2012.

F.E. Benth and H. Zdanowicz. Pricing spread options for stationary models in power markets. *In preparation*, 2013.

F.E. Benth, J. Šaltytė Benth, and S. Koekebakker. *Stochastic Modelling of Electricity and Related Markets*. Advanced Series on Statistical Science and Applied Probability. World Scientific, Singapore, 2008.

F.E. Benth, H. Eyjolfsson, and A.E.D. Veraart. Approximating Levy semistationary processes via Fourier methods in the context of power markets. *Submitted*, 2013.

C. Bernhardt, C. Klüppelberg, and T. Meyer-Brandis. Estimating high quantiles for electricity prices by stable linear models. *Journal of Energy Markets*, 1:3–19, 2008.

P. Bjerksund, H. Rasmussen, and G. Stensland. Valuation and risk management in the Nordic electricity market. In P. Bjørndal, M. Bjørndal, and M. Ronnqvist, editors, *Energy, Natural Resources and Environmental Economics*, pages 167–185. Springer Verlag, Heidelberg, 2010.

R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. CRC Financial Mathematics Series. Chapman & Hall, Boca Raton, 2004.

H. Geman. *Commodities and Commodity Derivatives: Modelling and Pricing for Agriculturals, Metals and Energy*. John Wiley & Sons, Chichester, 2005.

H. Geman and A. Roncoroni. Understanding the fine structure of electricity prices. *Journal of Business*, 79:1225–1261, 2006.

H.U. Gerber and E.S.W. Shiu. Option pricing by Esscher transforms. *Transactions of the Society of Actuaries*, 46:99–191, 1994.

Ph. Protter. *Stochastic Integration and Differential Equations*. Springer Verlag, New York, 1990.

K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999.

CONTRIBUTED TALK

ROBUSTNESS OF LOCALLY RISK-MINIMIZING HEDGING STRATEGIES IN FINANCE VIA BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS WITH JUMPS

Giulia Di Nunno[§], Asma Khedher† and Michèle Vanmaele[‡]

[§]*Center of Mathematics for Applications, University of Oslo, PO Box 1053 Blindern, N-0316 Oslo, Norway*
†*Chair of Mathematical Finance, Technische Universität München, Parkring 11, D-85748 Garching-Hochbruck, Germany*
‡*Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281 S9, 9000 Gent, Belgium*
Email: giulian@math.uio.no, asma.khedher@tum.de, michele.vanmaele@UGent.be

1. INTRODUCTION

Since Bismut (1973) introduced the theory of backward stochastic differential equations (BSDEs), there has been a wide range of literature about this topic. Researchers have kept on developing results on these equations and recently, many papers have studied BSDEs driven by Lévy processes (see, e.g., Carbone et al. (2008) and Øksendal and Zhang (2009)).

In Di Nunno et al. (2013) we consider a BSDE which is driven by a Brownian motion and a Poisson random measure (BSDEJ). We present two candidate-approximations to this BSDEJ and we prove that the solution of each candidate-approximation converges to the solution of the BSDEJ in a space which we specify. Here we will discuss one of these two approximations.

Our aim from considering such approximations is to investigate the effect of the small jumps of the Lévy process in quadratic hedging strategies in incomplete markets in finance (see, e.g., Föllmer and Schweizer (1991) and Vandaele and Vanmaele (2008) for more about quadratic hedging strategies in incomplete markets). These strategies are related to the study of the Föllmer-Schweizer decomposition (FS) or/and the Galtchouk-Kunita-Watanabe (GKW) decomposition which are both backward stochastic differential equations (see Choulli et al. (2010) for more about these decompositions).

The two most popular types of quadratic hedging strategies are the locally risk-minimizing strategies and the mean-variance hedging strategies. Let us consider a market in which the risky asset is modelled by a jump-diffusion process $S(t)_{t\geq 0}$. Let ξ be a contingent claim. A locally risk-minimizing strategy is a non self-financing strategy that allows a small cost process $C(t)_{t\geq 0}$ and insists on the fact that the terminal condition of the value of the portfolio is equal to the contingent claim (see Schweizer (2001)). Translating this into conditions on the contingent claim ξ shows that there exists a locally risk-minimizing strategy for ξ if ¡textcolorredand only if ξ admits a decomposition of the form

$$\xi = \xi^{(0)} + \int_0^T \chi^{FS}(s)dS(s) + \phi^{FS}(T), \tag{1}$$

where $\chi^{FS}(t)_{t\geq 0}$ is a process such that the integral in (1) exists and $\phi^{FS}(t)_{t\geq 0}$ is a martingale which has to satisfy certain conditions that we will show in the next sections of the paper. The decomposition (1) is called the FS decomposition. Its financial importance lies in the fact that it directly provides the locally risk-minimizing strategy for ξ. In fact at each time t the number of risky assets is given by $\chi^{FS}(t)$ and the cost $C(t)$ is given by $\phi^{FS}(t) + \xi^{(0)}$.
The mean-variance hedging strategy is a self-financing strategy which minimizes the hedging error in mean square sense (see Föllmer and Sondermann (1986)). In Di Nunno et al. (2013) we study the robustness of these two latter hedging strategies toward the model choice. Here, we report about the locally risk-minimizing strategy.

Hereto we assume that the process $S(t)_{t\geq 0}$ is a jump-diffusion driven by a pure jump term with infinite activity and a Brownian motion $W(t)_{t\geq 0}$. We consider an approximation $S_\varepsilon(t)_{t\geq 0}$ to $S(t)_{t\geq 0}$ in which we truncate the small jumps and replace them by a Brownian motion $B(t)_{t\geq 0}$ independent of $W(t)_{t\geq 0}$ and scaled with the standard deviation of the small jumps.
This idea of shifting from a model with small jumps to another where those variations are modeled by some appropriately scaled continuous component goes back to Asmussen and Rosinski (2001) who proved that the second model approximates the first one. This result is interesting from modelling point of view since the underlying model and the approximating models have the same distribution for ε very small. It is also interesting from a simulation point of view. In fact no easy algorithms are available for simulating general Lévy processes. However the approximating processes we obtain contain a compound Poisson process and a Brownian motion which are both easy to simulate (see Cont and Tankov (2004)).

In this paper we show that the value of the portfolio, the amount of wealth, and the cost process in a locally risk-minimizing strategy are robust to the choice of the model. In Di Nunno et al. (2013) we also show the robustness of the mean-variance hedging strategy. To prove these results we use the existence of the FS decomposition (1) and the convergence results on BSDEJs.

2. SOME MATHEMATICAL PRELIMINARIES

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. We fix $T > 0$. Let $W = W(t)$ and $B = B(t)$, $t \in [0, T]$, be two independent standard Wiener processes and $\widetilde{N} = \widetilde{N}(dt, dz)$, $t, z \in [0, T] \times \mathbb{R}_0$ ($\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$) be a centered Poisson random measure, i.e. $\widetilde{N}(dt, dz) = N(dt, dz) - \ell(dz)dt$, where $\ell(dz)$ is the jump measure and $N(dt, dz)$ is the Poisson random measure independent of the Brownian motions W and B and such that $\mathbb{E}[N(dt, dz)] = \ell(dz)dt$. Define $\mathcal{B}(\mathbb{R}_0)$ as the σ-algebra generated by the Borel sets $\bar{U} \subset \mathbb{R}_0$.

We assume that the jump measure has a finite second moment. Namely $\int_{\mathbb{R}_0} z^2 \ell(dz) < \infty$. We introduce the $\mathbb{P}$-augmented filtrations $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$, $\mathbb{G} = (\mathcal{G}_t)_{0\leq t\leq T}$, $\mathbb{G}^\varepsilon = (\mathcal{G}_t^\varepsilon)_{0\leq t\leq T}$, respectively by

$$\mathcal{F}_t = \sigma\Big\{W(s), \int_0^s \int_A \widetilde{N}(du,dz), \quad s \le t, \quad A \in \mathcal{B}(\mathbb{R}_0)\Big\} \vee \mathcal{N},$$

$$\mathcal{G}_t = \sigma\Big\{W(s), B(s), \int_0^s \int_A \widetilde{N}(du,dz), \quad s \le t, \quad A \in \mathcal{B}(\mathbb{R}_0)\Big\} \vee \mathcal{N},$$

$$\mathcal{G}_t^\varepsilon = \sigma\Big\{W(s), B(s), \int_0^s \int_A \widetilde{N}(du,dz), \quad s \le t, \quad A \in \mathcal{B}(\{|z| > \varepsilon\})\Big\} \vee \mathcal{N},$$

where $\mathcal{N}$ represents the set of $\mathbb{P}$-null events in $\mathcal{F}$. We introduce the notation $\mathbb{H} = (\mathcal{H}_t)_{0\le t\le T}$, such that $\mathcal{H}_t$ will be given either by the σ-algebra $\mathcal{F}_t$, $\mathcal{G}_t$, or $\mathcal{G}_t^\varepsilon$ depending on our analysis later. Define the following spaces for all $\beta \ge 0$;

- $L^2_{T,\beta}$: the space of all $\mathcal{H}_T$-measurable random variables $X : \Omega \to \mathbb{R}$ such that

$$\|X\|_\beta^2 = \mathbb{E}[\mathrm{e}^{\beta T} X^2] < \infty.$$

- $H^2_{T,\beta}$: the space of all $\mathbb{H}$-predictable processes $\phi : \Omega \times [0,T] \to \mathbb{R}$, such that

$$\|\phi\|^2_{H^2_{T,\beta}} = \mathbb{E}\Big[\int_0^T \mathrm{e}^{\beta t} |\phi(t)|^2 dt\Big] < \infty.$$

- $\widetilde{H}^2_{T,\beta}$: the space of all $\mathbb{H}$-adapted, càdlàg processes $\psi : \Omega \times [0,T] \to \mathbb{R}$ such that

$$\|\psi\|^2_{\widetilde{H}^2_{T,\beta}} = \mathbb{E}\Big[\int_0^T \mathrm{e}^{\beta t} |\psi^2(t) dt|\Big] < \infty.$$

- $\widehat{H}^2_{T,\beta}$: the space of all $\mathbb{H}$-predictable mappings $\theta : \Omega \times [0,T] \times \mathbb{R}_0 \to \mathbb{R}$, such that

$$\|\theta\|^2_{\widehat{H}^2_{T,\beta}} = \mathbb{E}\Big[\int_0^T \int_{\mathbb{R}_0} \mathrm{e}^{\beta t} |\theta(t,z)|^2 \ell(dz) dt\Big] < \infty.$$

- $S^2_{T,\beta}$: the space of all $\mathbb{H}$-adapted, càdlàg processes $\gamma : \Omega \times [0,T] \to \mathbb{R}$ such that

$$\|\gamma\|^2_{S^2_{T,\beta}} = \mathbb{E}[\mathrm{e}^{\beta T} \sup_{0\le t\le T} |\gamma^2(t)|] < \infty.$$

- $\nu_\beta = S^2_{T,\beta} \times H^2_{T,\beta} \times \widehat{H}^2_{T,\beta}$.
- $\widetilde{\nu}_\beta = S^2_{T,\beta} \times H^2_{T,\beta} \times \widehat{H}^2_{T,\beta} \times H^2_{T,\beta}$.
- $\widehat{L}^2_T(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \ell)$: the space of all $\mathcal{B}(\mathbb{R}_0)$-measurable mappings $\psi : \mathbb{R}_0 \to \mathbb{R}$ such that

$$\|\psi\|^2_{\widehat{L}^2_T(\mathbb{R}_0,\mathcal{B}(\mathbb{R}_0),\ell)} = \int_{\mathbb{R}_0} |\psi(z)|^2 \ell(dz) < \infty.$$

For notational simplicity, when $\beta = 0$, we skip the β in the notation.

The following result is an application of the Kunita-Watanabe decomposition of a random variable $\xi \in L_T^2$ with respect to orthogonal martingale random fields as integrators. See Kunita and Watanabe (1967) for the essential ideas.

Theorem 2.1 *Let $\mathbb{H} = \mathbb{G}$. Every $\mathcal{G}_T$-measurable random variable $\xi \in L_T^2$ has a unique representation of the form*

$$\xi = \xi^{(0)} + \sum_{k=1}^{3} \int_0^T \int_{\mathbb{R}} \varphi_k(t,z)\mu_k(dt,dz), \tag{2}$$

where the stochastic integrators

$$\mu_1(dt,dz) = W(dt) \times \delta_0(dz), \quad \mu_2(dt,dz) = B(dt) \times \delta_0(dz),$$
$$\mu_3(dt,dz) = \widetilde{N}(dt,dz)\mathbf{1}_{[0,T]\times\mathbb{R}_0}(t,z),$$

are orthogonal martingale random fields on $[0,T] \times \mathbb{R}_0$ and the stochastic integrands are φ_1, $\varphi_2 \in H_T^2$ and $\varphi_3 \in \widehat{H}_T^2$. Moreover $\xi^{(0)} = \mathbb{E}[\xi]$.
Let $\mathbb{H} = \mathbb{G}^\varepsilon$. Then for every $\mathcal{G}_T^\varepsilon$-measurable random variable $\xi \in L_T^2$, (2) holds with $\mu_3(dt,dz) = \widetilde{N}(dt,dz)\mathbf{1}_{[0,T]\times\{|z|>\varepsilon\}}(t,z)$.
Let $\mathbb{H} = \mathbb{F}$. Then for every $\mathcal{F}_T$-measurable random variable $\xi \in L_T^2$, (2) holds with $\mu_2(dt,dz) = 0$.

The above result plays a central role in the analysis. Let us now consider a pair (ξ, f), where ξ is called the terminal condition and f the driver such that

Assumptions 2.2
(A) $\xi \in L_T^2$ is $\mathcal{H}_T$-measurable
(B) $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that

- *$f(\cdot,x,y,z)$ is $\mathbb{H}$-progressively measurable for all x,y,z,*
- *$f(\cdot,0,0,0) \in H_T^2$,*
- *$f(\cdot,x,y,z)$ satisfies a uniform Lipschitz condition in (x,y,z), i.e. there exists a constant C such that for all $(x_i,y_i,z_i) \in \mathbb{R} \times \mathbb{R} \times \widehat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \ell)$, $i = 1,2$, we have*

$$|f(t,x_1,y_1,z_1) - f(t,x_2,y_2,z_2)| \le C\Big(|x_1 - x_2| + |y_1 - y_2| + \|z_1 - z_2\|\Big), \quad \textit{for all } t.$$

We consider the following backward stochastic differential equation with jumps (in short BSDEJ)

$$\begin{cases} -dX(t) &= f(t,X(t),Y(t),Z(t,\cdot))dt - Y(t)dW(t) - \displaystyle\int_{\mathbb{R}_0} Z(t,z)\widetilde{N}(dt,dz), \\ \quad X(T) &= \xi. \end{cases} \tag{3}$$

Definition 2.1 *A solution to the BSDEJ* (3) *is a triplet of* $\mathbb{H}$*-adapted or predictable processes* $(X, Y, Z) \in \nu$ *satisfying*

$$X(t) = \xi + \int_t^T f(s, X(s), Y(s), Z(s, \cdot))ds - \int_t^T Y(s)dW(s) \\ - \int_t^T \int_{\mathbb{R}_0} Z(s, z)\widetilde{N}(ds, dz), \qquad 0 \le t \le T.$$

The existence and uniqueness result for the solution of the BSDEJ (3) is guaranteed by the following result proved in Tang and Li (1994).

Theorem 2.3 *Given a pair* (ξ, f) *satisfying Assumptions 2.2(A) and (B), there exists a unique solution* $(X, Y, Z) \in \nu$ *to the BSDEJ* (3)*.*

3. A CANDIDATE-APPROXIMATING BSDEJ AND ROBUSTNESS

In this section we present a candidate approximation to the BSDEJ (3). Let $\mathbb{H} = \mathbb{G}$. We introduce a sequence of random variables $\mathcal{G}_T$-measurable $\xi_\varepsilon \in L_T^2$ such that

$$\lim_{\varepsilon \to 0} \xi_\varepsilon = \xi$$

and a function f^1 satisfying

Assumptions 3.1 $f^1 : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *is such that*

- $f^1(\cdot, x, y, z, \zeta)$ *is* $\mathbb{H}$*-progressively measurable for all* x, y, z, ζ*,*
- $f^1(\cdot, 0, 0, 0, 0) \in H_T^2$*,*
- $f^1(\cdot, x, y, z, \zeta)$ *satisfies a uniform Lipschitz condition in* (x, y, z, ζ)*.*

Besides Assumptions 3.1 which we impose on f^1, we need moreover to assume the following condition in the robustness analysis later on. For all $(x_i, y_i, z_i, \zeta) \in \mathbb{R} \times \mathbb{R} \times \widehat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \ell) \times \mathbb{R}$, $i = 1, 2$, and for a positive constant C we have

$$|f(t, x_1, y_1, z_1) - f^1(t, x_2, y_2, z_2, \zeta)| \le C\Big(|x_1 - x_2| + |y_1 - y_2| + \|z_1 - z_2\| + |\zeta|\Big), \quad \text{for all } t. \tag{4}$$

We introduce the candidate BSDEJ approximation to (3) as follows

$$\begin{cases} -dX_\varepsilon(t) &= f^1(t, X_\varepsilon(t), Y_\varepsilon(t), Z_\varepsilon(t, \cdot), \zeta_\varepsilon(t))dt - Y_\varepsilon(t)dW(t) - \displaystyle\int_{\mathbb{R}_0} Z_\varepsilon(t, z)\widetilde{N}(dt, dz) \\ & \qquad -\zeta_\varepsilon(t)dB(t), \\ X_\varepsilon(T) &= \xi_\varepsilon, \end{cases} \tag{5}$$

where B is a Brownian motion independent of W. Because of the presence of the additional noise B the solution processes are expected to be $\mathbb{G}$-adapted (or predictable). Notice that the solution of such equation is given by $(X_\varepsilon, Y_\varepsilon, Z_\varepsilon, \zeta_\varepsilon) \in \widetilde{\nu}$. In the next theorem we state the existence and uniqueness of the solution of the equation (5).

Theorem 3.2 *Let $\mathbb{H} = \mathbb{G}$. Given a pair (ξ_ε, f^1) such that $\xi_\varepsilon \in L_T^2$ is $\mathcal{G}_T$-measurable and f^1 satisfies Assumptions 3.1, then there exists a unique solution $(X_\varepsilon, Y_\varepsilon, Z_\varepsilon, \zeta_\varepsilon) \in \widetilde{\nu}$ to the BSDEJ* (5).

In the following theorem we state the convergence of the BSDEJ (5) to the BSDEJ (3).

Theorem 3.3 *Assume that f^1 satisfies* (4). *Let (X, Y, Z) be the solution of* (3) *and $(X_\varepsilon, Y_\varepsilon, Z_\varepsilon, \zeta_\varepsilon)$ be the solution of* (5). *Then we have for $t \in [0, T]$,*

$$\begin{aligned}
&\mathbb{E}\Big[\int_t^T |X(s) - X_\varepsilon(s)|^2 ds\Big] + \mathbb{E}\Big[\int_t^T |Y(s) - Y_\varepsilon(s)|^2 ds\Big] \\
&\quad + \mathbb{E}\Big[\int_t^T \int_{\mathbb{R}_0} |Z(s,z) - Z_\varepsilon(s,z)|^2 \ell(dz) ds\Big] + \mathbb{E}\Big[\int_t^T |\zeta_\varepsilon(s)|^2 ds\Big] \\
&\quad \le K\mathbb{E}[|\xi - \xi_\varepsilon|^2],
\end{aligned}$$

where K is a positive constant. It also holds that for some constant $C > 0$

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X(t) - X_\varepsilon(t)|^2\Big] \le C\mathbb{E}[|\xi - \xi_\varepsilon|^2].$$

The proofs can be found in Di Nunno et al. (2013).

4. ROBUSTNESS OF THE FÖLLMER-SCHWEIZER DECOMPOSITION WITH APPLICATIONS TO PARTIAL-HEDGING IN FINANCE

We assume we have two assets. One of them is a riskless asset with price $S^{(0)}$ given by

$$dS^{(0)}(t) = S^{(0)}(t) r(t) dt,$$

where $r(t) = r(t, \omega) \in \mathbb{R}$ is the short rate. The dynamics of the risky asset are given by

$$\begin{cases} dS^{(1)}(t) &= S^{(1)}(t)\Big\{a(t)dt + b(t)dW(t) + \int_{\mathbb{R}_0} \gamma(t,z)\widetilde{N}(dt,dz)\Big\}, \\ S^{(1)}(0) &= x \in \mathbb{R}_+ , \end{cases}$$

where $a(t) = a(t, \omega) \in \mathbb{R}$, $b(t) = b(t, \omega) \in \mathbb{R}$, and $\gamma(t, z) = \gamma(t, z, \omega) \in \mathbb{R}$ for $t \ge 0$, $z \in \mathbb{R}_0$ are adapted processes. We assume that $\gamma(t, z, \omega) = g(z)\widetilde{\gamma}(t, \omega)$, such that

$$G^2(\varepsilon) := \int_{|z|\le\varepsilon} g^2(z)\ell(dz) < \infty. \tag{6}$$

The dynamics of the discounted price process $\widetilde{S} = \frac{S^{(1)}}{S^{(0)}}$ are given by

$$d\widetilde{S}(t) = \widetilde{S}(t)\Big[(a(t) - r(t))dt + b(t)dW(t) + \int_{\mathbb{R}_0} \gamma(t,z)\widetilde{N}(dt,dz)\Big]. \tag{7}$$

Since $\widetilde{S}$ is a semimartingale, we can decompose it into a local martingale M starting at zero in zero and a finite variation process A, with $A(0) = 0$, where M and A have the following expressions

$$M(t) = \int_0^t b(s)\widetilde{S}(s)dW(s) + \int_0^t \int_{\mathbb{R}_0} \gamma(s,z)\widetilde{S}(s)\widetilde{N}(ds,dz), \tag{8}$$

$$A(t) = \int_0^t (a(s) - r(s))\widetilde{S}(s)ds.$$

We denote by $\langle X\rangle(t)$ the predictable compensator of the process X, i.e. $X(t) - \langle X\rangle(t)$, $0 \le t \le T$, is a local martingale. Then we can represent the process A as follows

$$A(t) = \int_0^t \frac{a(s) - r(s)}{\widetilde{S}(s)\big(b^2(s) + \int_{\mathbb{R}_0} \gamma^2(s,z)\ell(dz)\big)} d\langle M\rangle(s). \tag{9}$$

Let α be the integrand in equation (9), that is the process given by

$$\alpha(t) := \frac{a(t) - r(t)}{\widetilde{S}(t)\big(b^2(t) + \int_{\mathbb{R}_0} \gamma^2(t,z)\ell(dz)\big)}, \qquad 0 \le t \le T. \tag{10}$$

We define a process K by means of α as follows

$$K(t) = \int_0^t \alpha^2(s) d\langle M\rangle(s) = \int_0^t \frac{(a(s) - r(s))^2}{b^2(s) + \int_{\mathbb{R}_0} \gamma^2(s,z)\ell(dz)} ds. \tag{11}$$

The process K is called the mean-variance-trade-off (MVT) process.

In order to formulate our robustness study for the quadratic hedging strategies, we present the definition of the FS decomposition. We first introduce the following notations. Let S be a semimartingale. Then S can be decomposed as follows $S = S(0) + M + A$, where $S(0)$ is finite-valued and $\mathcal{F}_0$-measurable, M is a local martingale with $M(0) = 0$, and A is a finite variation process with $A(0) = 0$. We denote by $L(S)$ the class of predictable processes for which we can determine the stochastic integral with respect to S. We define the space Θ by

$$\Theta := \Big\{\theta \in L(S) \,|\, \mathbb{E}\Big[\int_0^T \theta^2(s) d\langle M\rangle(s) + \Big(\int_0^T |\theta(s)dA(s)|\Big)^2\Big] < \infty\Big\}.$$

Now we give the definition of the FS decomposition.

Definition 4.1 *Let S be a semimartingale. An $\mathcal{F}_T$-measurable and square integrable random variable H admits a Föllmer-Schweizer decomposition if there exist a constant H_0, an S-integrable process $\chi^{FS} \in \Theta$, and a square integrable martingale ϕ^{FS} such that ϕ^{FS} is orthogonal to M and*

$$H = H_0 + \int_0^T \chi^{FS}(s)dS(s) + \phi^{FS}(T).$$

Monat and Stricker (1995) show that a sufficient condition for the existence of the FS decomposition is to assume that the MVT process K given by (11) is uniformly bounded. The most general result concerning the existence and uniqueness of the FS decomposition is given by Choulli et al. (1998). In our case we assume that the process K is uniformly bounded in t by a constant C. Under this condition we can define the minimal martingale density by

$$\mathcal{E}\Big(\int_0^{\cdot} \alpha(s)dM(s)\Big)_t, \tag{12}$$

where M and α are respectively given by (8) and (10) and $\mathcal{E}(X)$ is the exponential martingale for X (see Theorem II, 37 in Protter (2005) for a general formula for exponential martingales). Notice that (12) defines a signed minimal martingale measure. For this martingale to exist as a probability martingale measure we have to assume that $\mathcal{E}\Big(\int_0^{\cdot} \alpha(s)dM(s)\Big)_t > 0$ (see, e.g., Choulli et al. (2010)). This latter condition is equivalent to (see Proposition 3.1 in Arai (2001))

$$\widetilde{S}(t)\alpha(t)\gamma(t,z) > -1, \qquad \text{a.e. in } (t,z,\omega). \tag{13}$$

In the following we assume that (13) holds. Let ξ be a square integrable contingent claim and $\widetilde{\xi} = \frac{\xi}{S^{(0)}(T)}$ its discounted value. Let $\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{P}}\big|_{\mathcal{F}_t} = \mathcal{E}\Big(\int_0^{\cdot} \alpha(s)dM(s)\Big)_t$ be the minimal martingale measure. Define $\widetilde{V}(t) = \mathbb{E}_{\widetilde{\mathbb{Q}}}[\widetilde{\xi}|\mathcal{F}_t]$. Then from Proposition 4.2 in Choulli et al. (2010), we have the following FS decomposition for $\widetilde{V}$ written under the world measure $\mathbb{P}$

$$\widetilde{V}(t) = \mathbb{E}_{\widetilde{\mathbb{Q}}}[\widetilde{\xi}] + \int_0^t \chi^{FS}(s)d\widetilde{S}(s) + \phi^{FS}(t), \tag{14}$$

where ϕ^{FS} is a $\mathbb{P}$-martingale orthogonal to M and $\chi^{FS} \in \Theta$. Replacing $\widetilde{S}$ by its value (7) in (14) we get

$$\begin{cases} d\widetilde{V}(t) &= \widetilde{\pi}(t)(a(t)-r(t))dt + \widetilde{\pi}(t)b(t)dW(t) + \displaystyle\int_{\mathbb{R}_0} \widetilde{\pi}(t)\gamma(t,z)\widetilde{N}(dt,dz) + d\phi^{FS}(t), \\ \widetilde{V}(T) &= \widetilde{\xi}, \end{cases} \tag{15}$$

where $\widetilde{\pi} = \chi^{FS}\widetilde{S}$.

Since $\phi^{FS}(T)$ is a F_T-measurable square integrable martingale then applying Theorem 2.1 with $\mathbb{H} = \mathbb{F}$ and the martingale property of $\phi^{FS}(T)$ we know that there exist stochastic integrands Y^{FS}, Z^{FS}, such that

$$\phi^{FS}(t) = \mathbb{E}[\phi^{FS}(T)] + \int_0^t Y^{FS}(s)dW(s) + \int_0^t \int_{\mathbb{R}_0} Z^{FS}(s,z)\widetilde{N}(ds,dz). \tag{16}$$

Since ϕ^{FS} is a martingale then we have $\mathbb{E}[\phi^{FS}(T)] = \mathbb{E}[\phi^{FS}(0)] = 0$. In that case, the set of equations (15) are equivalent to

$$\begin{cases} d\widetilde{V}(t) &= \widetilde{\pi}(t)(a(t)-r(t))dt + \big(\widetilde{\pi}(t)b(t) + Y^{FS}(t)\big)dW(t) \\ &\qquad + \displaystyle\int_{\mathbb{R}_0} \big(\widetilde{\pi}(t)\gamma(t,z) + Z^{FS}(t,z)\big)\widetilde{N}(dt,dz), \\ \widetilde{V}(T) &= \widetilde{\xi}. \end{cases} \tag{17}$$

Now we assume we have another model for the price of the risky asset. In this model we approximate the small jumps by a Brownian motion B which is independent of W and which we scale with the standard deviation of the small jumps, see (6). That is

$$\begin{cases} dS_\varepsilon^{(1)}(t) &= S_\varepsilon^{(1)}(t)\Big\{a(t)dt + b(t)dW(t) + \int_{|z|>\varepsilon} \gamma(t,z)\widetilde{N}(dt,dz) + G(\varepsilon)\widetilde{\gamma}(t)dB(t)\Big\}, \\ S_\varepsilon^{(1)}(0) &= S^{(1)}(0) = x\,. \end{cases}$$

The discounted price process is given by

$$d\widetilde{S}_\varepsilon(t) = \widetilde{S}_\varepsilon(t)\Big\{(a(t)-r(t))dt + b(t)dW(t) + \int_{|z|>\varepsilon} \gamma(t,z)\widetilde{N}(dt,dz) + G(\varepsilon)\widetilde{\gamma}(t)dB(t)\Big\}.$$

It was proven in Benth et al. (2013), that the process $\widetilde{S}_\varepsilon(t)_{t\geq 0}$ converges to $\widetilde{S}(t)_{t\geq 0}$ in L^2 when ε goes to 0 with rate of convergence $G(\varepsilon)$.

In the following we study the robustness of the locally risk-minimizing hedging strategy toward the model choice where the price processes are modeled by $\widetilde{S}$ and $\widetilde{S}_\varepsilon$.

The local martingale M_ε in the semimartingale decomposition of $\widetilde{S}_\varepsilon$ is given by

$$\begin{aligned} M_\varepsilon(t) = \int_0^t b(s)\widetilde{S}_\varepsilon(s)dW(s) + \int_0^t \int_{|z|>\varepsilon} \gamma(t,z)\widetilde{S}_\varepsilon(s)\widetilde{N}(dt,dz) \\ + G(\varepsilon)\int_0^t \widetilde{\gamma}(s)\widetilde{S}_\varepsilon(s)dB(s) \end{aligned} \tag{18}$$

and the finite variation process A_ε is given by

$$A_\varepsilon(t) = \int_0^t \frac{a(s)-r(s)}{\widetilde{S}_\varepsilon(s)\big(b^2(s) + \int_{|z|\geq\varepsilon} \gamma^2(s,z)\ell(dz)\big)} d\langle M_\varepsilon\rangle(s).$$

We define the process α_ε by

$$\alpha_\varepsilon(t) := \frac{a(t)-r(t)}{\widetilde{S}_\varepsilon(t)\big(b^2(t) + G^2(\varepsilon)\widetilde{\gamma}^2(t) + \int_{|z|>\varepsilon} \gamma^2(t,z)\ell(dz)\big)}, \qquad 0\leq t\leq T. \tag{19}$$

Thus the mean-variance trade-off process K_ε is given by

$$\begin{aligned} K_\varepsilon(t) &= \int_0^t \alpha_\varepsilon^2(s)d\langle M_\varepsilon\rangle(s) = \int_0^t \frac{(a(s)-r(s))^2}{b^2(s) + G^2(\varepsilon)\widetilde{\gamma}^2(s) + \int_{|z|>\varepsilon} \gamma^2(s,z)\ell(dz)} ds \\ &= K(t), \end{aligned} \tag{20}$$

in view of the definition of $G(\varepsilon)$, equation (6). Hence the boundedness of K ensures the existence of the FS decomposition with respect to $\widetilde{S}_\varepsilon$ for any square integrable $\mathcal{G}_T$-measurable random variable.

Let ξ_ε be a square integrable contingent claim. We denote by $\widetilde{\xi}_\varepsilon = \frac{\xi_\varepsilon}{S^{(0)}(T)}$ the discounted pay-off of the contingent claim with $\widetilde{S}_\varepsilon$ as underlying. As we have seen before, for the minimal measure to be a probability martingale measure, we have to assume that

$$\mathcal{E}\Big(\int_0^\cdot \alpha_\varepsilon(s)dM_\varepsilon(s)\Big)_t > 0,$$

which is equivalent to

$$\widetilde{S}_\varepsilon(t)\alpha_\varepsilon(t)\gamma(t,z) > -1, \qquad \text{a.e. in } (t,z,\omega). \tag{21}$$

Define $\frac{d\widetilde{\mathbb{Q}}^\varepsilon}{d\mathbb{P}}\big|_{\mathcal{G}_t} := \mathcal{E}\Big(\int_0^\cdot \alpha_\varepsilon(s)dM_\varepsilon(s)\Big)_t$ and $\widetilde{V}_\varepsilon(t) := \mathbb{E}_{\widetilde{\mathbb{Q}}^\varepsilon}[\widetilde{\xi}_\varepsilon|\mathcal{G}_t]$. Then from Proposition 4.2 in Choulli et al. (2010), we have the following FS decomposition for $\widetilde{V}_\varepsilon$ written under the world measure $\mathbb{P}$

$$\widetilde{V}_\varepsilon(t) = \mathbb{E}_{\widetilde{\mathbb{Q}}^\varepsilon}[\widetilde{\xi}_\varepsilon] + \int_0^t \chi_\varepsilon^{FS}(s)d\widetilde{S}_\varepsilon(s) + \phi_\varepsilon^{FS}(t), \tag{22}$$

where ϕ_ε^{FS} is a $\mathbb{P}$-martingale orthogonal to M_ε and $\chi_\varepsilon^{FS} \in \Theta$. Replacing $\widetilde{S}_\varepsilon$ by its expression in (22), we get

$$\begin{cases} d\widetilde{V}_\varepsilon(t) &= \widetilde{\pi}_\varepsilon(t)(a(t)-r(t))dt + \widetilde{\pi}_\varepsilon(t)b(t)dW(t) + \widetilde{\pi}_\varepsilon(t)G(\varepsilon)\widetilde{\gamma}(t)dB(t) \\ & \qquad + \int_{|z|>\varepsilon} \widetilde{\pi}_\varepsilon(t)\gamma(t,z)\widetilde{N}(dt,dz) + d\phi_\varepsilon^{FS}(t), \\ \widetilde{V}_\varepsilon(T) &= \widetilde{\xi}_\varepsilon, \end{cases}$$

where $\widetilde{\pi}_\varepsilon = \chi_\varepsilon^{FS}\widetilde{S}_\varepsilon$. Notice that $\phi_\varepsilon^{FS}(T)$ is a $\mathcal{G}_T^\varepsilon$-measurable square integrable $\mathbb{P}$-martingale. thus applying Theorem 2.1 with $\mathbb{H} = \mathbb{G}^\varepsilon$ and using the martingale property of $\phi_\varepsilon^{FS}(T)$ we know that there exist stochastic integrands $Y_{1,\varepsilon}^{FS}$, $Y_{2,\varepsilon}^{FS}$, and Z_ε^{FS}, such that

$$\begin{aligned} \phi_\varepsilon^{FS}(t) = \mathbb{E}[\phi_\varepsilon^{FS}(T)] &+ \int_0^t Y_{1,\varepsilon}^{FS}(s)dW(s) + \int_0^t Y_{2,\varepsilon}^{FS}(s)dB(s) \\ &+ \int_0^t \int_{|z|>\varepsilon} Z_\varepsilon^{FS}(s,z)\widetilde{N}(ds,dz). \end{aligned} \tag{23}$$

Using the martingale property of ϕ_ε^{FS} and equation (22), we get $\mathbb{E}[\phi_\varepsilon^{FS}(T)] = \mathbb{E}[\phi_\varepsilon^{FS}(0)] = 0$. The equation we obtain for the approximating problem is thus given by

$$\begin{cases} d\widetilde{V}_\varepsilon(t) &= \widetilde{\pi}_\varepsilon(t)(a(t)-r(t))dt + (\widetilde{\pi}_\varepsilon(t)b(t) + Y_{1,\varepsilon}^{FS}(t))dW(t) \\ & \qquad + (\widetilde{\pi}_\varepsilon(t)G(\varepsilon)\widetilde{\gamma}(t) + Y_{2,\varepsilon}^{FS}(t))dB(t) \\ & \qquad + \int_{|z|>\varepsilon} \big(\widetilde{\pi}_\varepsilon(t)\gamma(t,z) + Z_\varepsilon^{FS}(t,z)\big)\widetilde{N}(dt,dz), \\ \widetilde{V}_\varepsilon(T) &= \widetilde{\xi}_\varepsilon. \end{cases} \tag{24}$$

In order to apply the robustness results studied in Section 3, we have to prove that $\widetilde{V}$ and $\widetilde{V}_\varepsilon$ are respectively equations of type (3) and (5). That's the purpose of the next lemma. Notice that here above $\widetilde{V}_\varepsilon$, $\widetilde{\pi}_\varepsilon$, and ϕ_ε^{FS} are all $\mathcal{G}_t^\varepsilon$-measurable. However since $\mathcal{G}_t^\varepsilon \subset \mathcal{G}_t$, then $\widetilde{V}_\varepsilon$, $\widetilde{\pi}_\varepsilon$, and ϕ_ε^{FS} are also $\mathcal{G}_t$-measurable.

Lemma 4.1 *Let $\kappa(t) = b^2(t) + \int_{\mathbb{R}_0} \gamma^2(t,z)\ell(dz)$. Assume that for all $t \in [0,T]$,*

$$\frac{|a(t)-r(t)|}{\sqrt{\kappa(t)}} \le C, \qquad \mathbb{P}\text{-a.s.}, \tag{25}$$

for a positive constant C. Let $\widetilde{V}$, $\widetilde{V}_\varepsilon$ be given by (17), (24)*, respectively. Then $\widetilde{V}$ satisfies a BSDEJ of type* (3) *and $\widetilde{V}_\varepsilon$ satisfies a BSDEJ of type* (5).

Now we present the following main result in which we prove the robustness of the value of the portfolio, the robustness result for the amount of wealth to invest in the stock in a locally risk-minimizing strategy, and the robustness of the process ϕ^{FS} defined in (16)

Theorem 4.2 *Assume that* (25) *holds. Let* $\widetilde{V}$, $\widetilde{V}_\varepsilon$ *be given by* (17), (24)*, respectively. Then it holds that*

$$\mathbb{E}\Big[\sup_{0\le t\le T}|\widetilde{V}(t)-\widetilde{V}_\varepsilon(t)|^2\Big]\le C\mathbb{E}[|\widetilde{\xi}-\widetilde{\xi}_\varepsilon|^2].$$

Assume that (25) *holds and that for all* $t\in[0,T]$

$$\inf_{t\le s\le T}\kappa(s)\ge K,\qquad \mathbb{P}\text{-a.s.},\tag{26}$$

where K *is a strictly positive constant. Let* $\widetilde{\pi}=\chi^{FS}\widetilde{S}$ *and* $\widetilde{\pi}_\varepsilon=\chi^{FS}_\varepsilon\widetilde{S}_\varepsilon$*. Then for all* $t\in[0,T]$*,*

$$\mathbb{E}\Big[\int_t^T|\widetilde{\pi}(s)-\widetilde{\pi}_\varepsilon(s)|^2ds\Big]\le C\mathbb{E}[|\widetilde{\xi}-\widetilde{\xi}_\varepsilon|^2],$$

where C *is a positive constant. Assume that* (25) *and* (26) *hold and for all* $t\in[0,T]$

$$\sup_{t\le s\le T}\widetilde{\gamma}^2(s)\le\widetilde{K},\qquad\sup_{t\le s\le T}\kappa(s)\le\widehat{K}<\infty,\qquad\mathbb{P}\text{-a.s.}$$

Let ϕ^{FS}, ϕ^{FS}_ε *be given by* (16), (23)*, respectively. Then for all* $t\in[0,T]$*, we have*

$$\mathbb{E}\Big[|\phi^{FS}(t)-\phi^{FS}_\varepsilon(t)|^2\Big]\le C\mathbb{E}[|\widetilde{\xi}-\widetilde{\xi}_\varepsilon|^2]+C'G(\varepsilon),$$

where C *and* C' *are positive constants.*

The processes C and C_ε with $C(t)=\phi^{FS}(t)+\widetilde{V}(0)$ and $C_\varepsilon(t)=\phi^{FS}_\varepsilon(t)+\widetilde{V}_\varepsilon(0)$, are the cost processes in a locally risk-minimizing strategy for $\widetilde{\xi}$ and $\widetilde{\xi}_\varepsilon$. Using the last theorem it is easy to show that for all $t\in[0,T]$, we have

$$\mathbb{E}[|C(t)-C_\varepsilon(t)|^2]\le\widetilde{K}\mathbb{E}[|\widetilde{\xi}-\widetilde{\xi}_\varepsilon|^2]+K'G(\varepsilon),$$

where $\widetilde{K}$ and K' are two positive constants.

References

T. Arai. On the equivalent martingale measures for Poisson jump type model. Technical Report KSTS/RR-01/004, Keio University, Japan, July 2001.

S. Asmussen and J. Rosinski. Approximations of small jump Lévy processes with a view towards simulation. *Journal of Applied Probability*, 38:482–493, 2001.

F.E. Benth, G. Di Nunno, and A. Khedher. A note on convergence of option prices and their Greeks for Lévy models. *Stochastics: An International Journal of Probability and Stochastic Processes*, 0(0):1–25, 2013. to appear.

J.M. Bismut. Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.*, 44: 384–404, 1973.

R. Carbone, B. Ferrario, and M. Santacroce. Backward stochastic differential equations driven by cádlág martingales. *Theory Probab. Appl.*, 52(2):304–314, 2008.

T. Choulli, L. Krawczyk, and C. Stricker. $\mathcal{E}$-martingales and their applications in mathematical finance. *Annals of Probability*, 26(2):853–876, 1998.

T. Choulli, N. Vandaele, and M. Vanmaele. The Föllmer-Schweizer decomposition: Comparison and description. *Stochastic Processes and their Applications*, 120(6):853–872, 2010.

R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall, 2004.

G. Di Nunno, A. Khedher, and M. Vanmaele. Robustness of quadratic hedging strategies in finance via backward stochastic differential equations with jumps. Technical report, 2013.

H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. In M.H.A. Davis and R.J. Elliot, editors, *Applied Stochastic Analysis*, volume 5 of *Stochastic Monographs*, pages 389–414. Gordon and Breach, 1991.

H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. In W. Hildenbrand and A. Mas-Colell, editors, *Contributions to Mathematical Economics*, pages 205–223. North-Holland, Elsevier, 1986.

H. Kunita and S. Watanabe. On square integrable martingales. *Nagoya Mathematical Journal*, 30: 209–245, 1967.

P. Monat and C. Stricker. Föllmer-Schweizer decomposition and mean-variance hedging for general claims. *Annals of Probability*, 23:605–628, 1995.

B. Øksendal and T. Zhang. Backward stochastic differential equations with respect to general filtrations and applications to insider finance. *Preprint No. 19, September, Department of Mathematics, University of Oslo, Norway*, 2009.

P. Protter. *Stochastic integration and differential equations, Second Edition, Version* 2.1. Number 21 in Stochastic Modelling and Applied Probability. Springer, Berlin, 2005.

M. Schweizer. A guided tour through quadratic hedging approaches. In E. Jouini, J. Cvitanić, and M. Musiela, editors, *Option Pricing, Interest Rates and Risk Management*, pages 538–574. Cambridge University Press, 2001.

S. Tang and X. Li. Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J. Control Optim*, 32(5):1447–1475, 1994.

N. Vandaele and M. Vanmaele. A locally risk-minimizing hedging strategy for unit-linked life insurance contracts in a Lévy process financial market. *Insurance: Mathematics and Economics*, 42(3):1128–1137, 2008.

POSTER SESSION

FACTORS AFFECTING THE SMILE AND IMPLIED VOLATILITY IN THE CONTEXT OF OPTION PRICING MODELS

Akhlaque Ahmad

Financial Engineering and Risk Management Program, National Institute of Securities Markets (Established by SEBI), New Mumbai- 400703, India
Department of Economics, University of Mumbai, Mumbai- 400098, India
Email: `akhlaque.ahmad@nism.ac.in`

1. INTRODUCTION

The application of Fourier transform gives us a convenient method for modelling with general stochastic processes and distributions. Here, we discuss a comprehensive treatment of the Fourier transform in option valuation covering most of the stochastic factors such as stochastic volatilities, stochastic interest rates and Poisson jumps. These are considered risk factors which influence option prices. We start with the general framework of asset pricing and characteristic functions and Fourier transforms which play an important role to incorporate risk factors in the option pricing framework.

We discuss the Heston (1993) model and the Schöbel and Zhu (1999) model and the characteristic functions for these processes. Also, we review two short term stochastic interest rate models, Cox et al. (1985) model and Vasicek (1977) model. We describe the model properties and corresponding characteristic functions. The Fourier transform is an elegant and efficient technique which incorporates discontinuous jump events in an asset process. Here, we deal with traditional jump models where the jump mechanism is governed by a compound Poisson process. We use simple jumps and jumps governed by a lognormal distribution. Finally, we describe our findings and suggest future directions of research. Among the models, we find that the Heston (1993) model and the Schöbel and Zhu (1999) model better explain the smile effects.

2. GENERAL FRAMEWORK FOR ASSET PRICING

If the returns of an asset $S(t)$ follow a fixed quantity plus a Brownian motion, then the asset $S(t)$ follows a Geometric Brownian motion

$$dS(t) = \mu(t)dt + \sigma(t)dW(t)$$

In the Black and Scholes (1973) framework the parameters $\mu(t)$ and $\sigma(t)$ are constant. In this paper, we relax this assumption and consider that they follow some stochastic process.

Black-Scholes Framework

The pay-off of a European Call option $C(T)$ at maturity date T with strike price K is given by

$$C(T) = \max[S(T) - K, 0]$$

which depends on the underlying $S(T)$. The fundamental question is, what is the fair price of the call option? The question is answered by Black and Scholes (1973) as the fair price of an option must be an arbitrage free price in the sense that a risk-less portfolio comprising of options and underlying stocks must reward a risk-free return. Using this idea, they derived the following formula

$$C_0 = S_0 N(d_1) - Ke^{-rT} N(d_2)$$

The above formula is known as the famous Black and Scholes (1973) formula and it is derived using partial differential equations. Another approach to find the fair call price is risk-neutral valuation. In this method, the fair value of a call is the discounted present value of its expectation at maturity. Instead of solving a partial differential equation, we can calculate the expected value $E[C(T)]$ and discount with the risk free interest rate r to obtain a call price. Therefore, $C_0 = e^{-rT}E[C(T)]$. The Feynman-Kac Theorem plays an important role to establish the equivalence between these two approaches.

Pricing Via Fourier Transform

Under the risk-neutral valuation, the process for $X(t)$ is given by

$$dX(t) = \left(r(t) - \frac{1}{2}b^2(v(t), t)\right) dt + b(v(t), t)dW_1(t)$$

In a general setting, the dynamics of the stock prices $S(t)$ are driven by a pure diffusion $dW(t)$ as in the simple Black and Scholes (1973) model. The essential extensions are a stochastic interest rate $r(t)$ and a stochastic volatility term $b(., t)$, that will be specified in the appropriate modelling. Also, we extend this setting by introducing different types of jumps which will improve the model accuracy.

By considering the above discussion, in this extended framework the exercise probabilities are no longer strictly normally distributed. However, they can be expressed by Fourier inversion of the associated characteristic functions which may often have closed form solutions with different specifications of stochastic factors. We can express the option pricing formula in the following form

$$C(T, K) = S_0 F_1^{Q_1}[X(T) > \ln K] - B(0, T)KF_2^{Q_2}[X(T) > \ln K]$$

where Q_1 denotes the delta measure and Q_2 denotes the forward T-measure at time T. The probabilities $F_1^{Q_1}$ and $F_2^{Q_2}$ are two standard normal distributions. We can express these probabilities by Fourier Transform. The characteristic function of $F_1^{Q_1}$ and $F_2^{Q_2}$ is given by

$$\begin{aligned}\phi_1(u) &= E^{Q_1}[\exp(iuX(T)] = E^{Q}[g_1(T)\exp(iuX(T)]\\ \phi_2(u) &= E^{Q_2}[\exp(iuX(T)] = E^{Q}[g_2(T)\exp(iuX(T)]\end{aligned}$$

where $g_1(T)$ and $g_2(T)$ are two risk-neutral densities at time T. The closed form formula for the probabilities F_j, $j = 1, 2$, is given by (Iacus 2011)

$$F_j = \frac{1}{2} + \frac{1}{\pi}\int_0^T Re(\phi_j(u))\frac{\exp(-iu\ln K)}{iu}du, \quad j = 1, 2. \tag{1}$$

Writing probability through the characteristic function is equivalent to writing through its density function. The one to one correspondence between a characteristic function and its distribution guarantees a unique form of the option pricing formula (Zhu 2010).

3. STOCHASTIC VOLATILITY MODELS

Stochastic volatility models provide a natural way to capture the volatility smile by assuming that volatility follows a stochastic process. The stochastic process to model volatility should be stationary with some possible features such as mean reverting, correlation and stock dynamics.

Heston Model

Model Description

The Heston (1993) model is the first stochastic volatility model with the utilization of characteristic functions. It models stochastic variance rather than stochastic volatility. The risk-neutral dynamics is given by following the stochastic differential equations

$$\begin{aligned}\frac{dS(t)}{S(t)} &= rdt + \sqrt{V(t)}dW_1(t)\\ dV(t) &= \kappa(\theta - V(t))dt + \sigma\sqrt{V(t)}dW_2(t)\\ dW_1(t)dW_2(t) &= \rho dt\end{aligned}$$

The parameter θ is the long-term level variance which gradually converges to $V(t)$. The parameter κ is the speed of variance reverting to θ. The parameter σ is referred to as the volatility of variance. If κ, θ and σ satisfy the following condition $2\kappa\theta > \sigma^2$, where $V_0 > 0$, then the variance $V(t)$ is always positive and the variance process is well defined. This condition is referred to as the Feller condition for a square root process.

Characteristic Functions

$$\begin{aligned}\phi_1(u) &= \exp[iu(X_0 + rT) - s_{21}(V_0 + \kappa\theta T) + A_1(T)V_0 + A_2(T)]\\ \phi_2(u) &= \exp[iu(X_0 + rT) - s_{22}(V_0 + \kappa\theta T) + A_3(T)V_0 + A_4(T)]\end{aligned}$$

Schöbel and Zhu Model

Model Description

Schöbel and Zhu (1999) extended the Stein-Stein stochastic volatility model in more general case using a mean reverting Ornstein-Uhlenbeck process. The advantage of a OU-process is that it is a Gaussian process and has nice analytical tractability. They provide analytical option prices using the following stochastic differential equations

$$\begin{aligned}\frac{dS(t)}{S(t)} &= rdt + \nu(t)dW_1(t)\\ d\nu(t) &= \kappa(\theta - \nu(t))dt + \sigma dW_2(t)\\ dW_1(t)dW_2(t) &= \rho dt\end{aligned}$$

Here, we model the volatility, not the variance. The volatility process is mean-reverting with mean level θ and reverting parameter κ. The parameter σ is the volatility of volatility and controls the variation $\nu(t)$.

Characteristic Functions

$$\phi_1(u) = \exp[iu(X_0 + rT) - \frac{\rho}{2\sigma}(1+iu)\nu_0^2 - \frac{1}{2}(1+iu)\rho\sigma T + \frac{1}{2}A_5(T)\nu_{0^2} + A_6(T)\nu_0 + A_7(T)]$$
$$\phi_2(u) = \exp[iu(X_0 + rT) - \frac{\rho}{2\sigma}(1+iu)\nu_0^2 - \frac{1}{2}(1+iu)\rho\sigma T + \frac{1}{2}A_8(T)\nu_{0^2} + A_9(T)\nu_0 + A_{10}(T)]$$

4. STOCHASTIC INTEREST RATE MODELS

In this section, we discuss two stochastic interest rate models that follow the same stochastic process as the stochastic volatility models in the previous section. These stochastic interest rate models are directly incorporated in the option pricing framework using characteristic functions. Here, we focus on only single factor short rate models, the Cox et al. (1985) model and the Vasicek (1977) model which are again specified by a mean reverting square root process and the mean reverting Ornstein-Uhlenbeck process respectively.

The Cox-Ingersoll-Ross Model

Model Description

The Cox et al. (1985) model first time modelled interest rate using a square root process. The model is described by the following stochastic differential equations

$$\begin{aligned}\frac{dS(t)}{S(t)} &= r(t)\left(1 - \frac{1}{2}\nu^2\right)dt + \nu\sqrt{r(t)}dW_1(t)\\ dr(t) &= \kappa[\theta - r(t)] + \sigma\sqrt{r(t)}dW_3(t)\\ dW_1(t)dW_3(t) &= \rho dt\end{aligned}$$

Characteristic Functions

$$\phi_1(u) = E\left[\exp\left(iuX_0 - s_{11}\int_0^T r(t)dt + s_{12}r(T) - s_{12}(r_0 + \kappa\theta T)\right)\right]$$

$$\phi_2(u) = E\left[\exp\left(X_0 - \ln B(0,T,r_0) - s_{21}\int_0^T r(t)dt + s_{22}r(T) - s_{22}(r_0 + \kappa\theta T)\right)\right]$$

Vasicek Model

Model Description

The drawback of modelling interest rate as a square root process in option pricing is that we need an alternative stock price process if the correlation between stock prices and interest rate is available. The Vasicek (1977) model where short rates are governed by mean reverting Ornstein-Uhlenbeck process overcomes this drawback. The pricing dynamics is governed by the following stochastic differential equations

$$\begin{aligned} \frac{dS(t)}{S(t)} &= r(t)dt + \nu\sqrt{r(t)}dW_1(t) \\ dr(t) &= \kappa[\theta - r(t)] + \sigma dW_3(t) \\ dW_1(t)dW_3(t) &= \rho dt \end{aligned}$$

Characteristic Functions

$$\begin{aligned} \phi_1(u) &= \exp\left(iuX_0 - \frac{iu+1}{2}\nu^2T - \frac{(iu+1)\nu\rho}{\sigma}(r_0 + \kappa\theta T) + \frac{1}{2}(iu+1)^2\nu^2(1-\rho^2)T\right) \\ &\quad \times E\left[\exp\left(-s_{11}\int_0^T r(t)dt + s_{12}r(t)\right)\right] \\ \phi_2(u) &= \exp\left(iuX_0 - \frac{iu}{2}\nu^2T - s_{22}(r_0 + \kappa\theta T) - \frac{1}{2}u^2\nu^2(1-\rho^2)T - \ln B(0,T)\right) \\ &\quad \times E\left[\exp\left(-s_{21}\int_0^T r(t)dt + s_{22}r(T)\right)\right] \end{aligned}$$

5. POISSON PROCESS JUMP MODELS

To model jump events in the market, we need two quantities: jump frequency and jump size. The former specifies how many times jumps happen in a given time period and the latter determines how large a jump is if it occurs. Here we discuss option pricing models with simple jumps and lognormal.

Simple Jump Model

Model Description

The dynamics of a stock price with pure jumps is given by

$$\frac{dS(t)}{S(t)} = [r(t) - \lambda J]dt + \nu(t)dW_1(t) + JdY(t)$$

where λ denotes the jump intensity and J is the jump size. Our concern is to show what the jump contributes to the characteristic function. Here, we assume that volatility and interest rate are constant. If $X(t) = \ln(S(t))$ then,

$$dX(t) = \left(r - \frac{1}{2}\nu^2 - \lambda J\right) dt + \nu dW_1 + \ln(1+J)dY(t)$$

Characteristic Functions

$$\phi_1(u) = \exp(iu(X_0 + rT) - (1+iu)(\frac{1}{2}\nu^2 + \lambda J)T + \frac{1}{2}(1+iu)^2\nu^2 T + \lambda T e^{(1+iu)\ln(1+J)} - \lambda T)$$

$$\phi_2(u) = \exp(iurT + iuX_0 - iu(\frac{1}{2}\nu^2 + \lambda J)T + \frac{1}{2}(1+iu)^2\nu^2 T + \lambda T e^{iu\ln(1+J)} - \lambda T)$$

Lognormal Jump Model
Model Description
The stock price dynamics is given by

$$dX(t) = \left[r(t) - \lambda\mu_J - \frac{1}{2}\nu^2(t)\right] + \nu(t)dW_1 + \ln(1+J)dY(t)$$

If jump size J is lognormally distributed and Brownian motion W_1, the Poisson process Y and jump size J are mutually stochastically independent, then

$$\ln(1+J) \sim N\left[\ln(1+\mu_J) - \frac{1}{2}\sigma_J^2, \sigma_J^2\right], \qquad \mu_J \geq -1$$

where μ_J is the mean of J and σ_J^2 is variance of $\ln(1+J)$.

Characteristic Functions

$$\phi_1(u) = \exp(iu(rt + X_0) - (1+iu)\lambda T\mu_J + \frac{1}{2}iu(1+iu)\nu^2 T + \lambda T[(1+\mu_J)^{(1+iu)} e^{\frac{1}{2}iu(iu+1)\sigma_J^2} - 1])$$

$$\phi_2(u) = \exp(iu(rt + X_0) - iu\lambda T\mu_J + \frac{1}{2}iu(1+iu)\nu^2 T + \lambda T[(1+\mu_J)^{iu} e^{\frac{1}{2}iu(iu-1)\sigma_J^2} - 1])$$

The probabilities F_j can be calculated using formula (1) and we calculate option prices for each affine model by using the following formula (Iacus 2011):

$$C(K,T) = SF_1 - Ke^{-rT}F_2 \tag{2}$$

6. DATA DESCRIPTION AND PARAMETER ESTIMATION

We use data from the India VIX, MIBOR and NIFTY Index to estimate model parameters. The source of the data is the National Stock Exchange (NSE), India and the data period is from March 1, 2009 to March 31, 2012. The method of parameter estimation is taken from Iacus (2008) and the model parameters are estimated using R-Packages SDE (Iacus 2009) and Yuima (Iacus 2010). Table 1 describes the estimated model parameters. Theoretical option prices are calculated using the method described in Carr and Madan (1999) and numerical values are obtained using the modified and extended Matlab codes given by Kienitz and Wetterau (2012). Table 2 explains the calculated option prices for different strike prices.

Models	$S = 4000$	$T = 1$	$r_0 = 0.051$	$\nu_0 = 0.43$
Heston	$\kappa = 3.12$	$\theta = 0.041$	$\sigma = 0.11$	$\rho = -0.81$
Shobel-Zhu	$\kappa = 3.09$	$\theta = 0.211$	$\sigma = 0.10$	$\rho = -0.81$
CIR	$\kappa = 2.01$	$\theta = 0.051$	$\sigma = 0.10$	$\rho = 0$
Vasicek	$\kappa = 2.04$	$\theta = 0.052$	$\sigma = 0.10$	$\rho = 0$
Simple	$\lambda = 0.25$	$\mu = 0.19$	$\sigma = 0.09$	-
Lognormal	$\lambda = 0.20$	$\mu = 0.19$	$\sigma = 0.08$	-

Table 1: Estimated model parameters

Models	K = 3800	K = 3900	K = 4000	K = 4100	K = 4200
BS	318.18	257.23	205.55	159.81	122.71
Heston	318.51	257.91	205.84	159.19	122.01
Schobel-Zhu	319.17	258.18	206.19	159.17	121.81
CIR	319.07	258.03	205.55	159.81	121.71
Vasicek	319.15	258.11	205.64	159.71	122.11
Simple	318.28	258.36	206.53	158.24	121.17
Lognormal	318.17	258.48	206.71	159.21	122.80

Table 2: Theoretical option prices

7. CONCLUSIONS

We incorporate each stochastic factor like stochastic volatility, stochastic interest rate and jumps in stock prices individually as a risk factors in the traditional Black-Scholes framework. Further, we estimate the model parameters on real data and calculate the theoretical option prices for our analysis. The option prices using the above discussed models are calculated with the estimated parameters. Now, we can conclude which models are able to generate more skewness than the Black and Scholes (1973) model. Both the Heston (1993) model and Schöbel and Zhu (1999) model with negative correlation (stock prices and volatility) produces higher prices for ITM options and lower prices for OTM options. This implies that both stochastic models can generate a down sloping smile. Also, we can see that adding a stochastic factor to the Black and Scholes (1973) model produces higher prices in most of the cases. This can be seen as adding an additional risk factor in the model implies more premium. This approach can be extended to more complex options as long as the characteristic function is known.

References

Fisher Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

Peter Carr and Dilip B. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2:61–73, 1999.

John C. Cox, Jonathan E. Ingersoll, and Stephen A Ross Jr. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–408, 1985.

Steven L. Heston. A Closed-Form Solution Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2):327–343, 1993.

Stefano M. Iacus. *Simulation and Inference for Stochastic Differential Equations*, volume 1 of *Springer Series in Statistics*. Springer New York, New York, NY, 2008. ISBN 978-0-387-75838-1.

Stefano M. Iacus. From the sde package to the Yuima Project, 2009.

Stefano M Iacus. The yuima package: an R framework for simulation and inference of SDEs, 2010.

Stefano M. Iacus. *Option Pricing and Estimation of Financial Models with R*. John Wiley & Sons, Ltd, Chichester, UK, 2011. ISBN 9781119990079.

Joerg Kienitz and Daniel Wetterau. *Financial Modelling: Theory, Implementation and Practice with MATLAB Source*. Wiley Finance Series. Wiley, 2012. ISBN 0-387-29909-2.

R. Schöbel and J. Zhu. Stochastic Volatility With an Ornstein-Uhlenbeck Process: An Extension. *Review of Finance*, 3(1):23–46, 1999.

Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, 1977.

Jianwei Zhu. *Applications of Fourier Transform to Smile Modeling*. Springer, Berlin, Heidelberg, 2010. ISBN 978-3-642-01807-7.

PENSION RULES AND IMPLICIT MARGINAL TAX RATE IN FRANCE

Frédéric Gannon[†] **and Vincent Touzé**[§]

[†]*U. Le Havre & EDEHN, France*
[§]*Sciences Po - OFCE, Paris, France*
Email: `frederic.gannon@univ-lehavre.fr, vincent.touze@sciences-po.fr`

Abstract

The pension rules link the amount of the future pension to the contributions during the working period. So, in case the pension rules are not actuarial, they induce an implicit tax. In this paper, we evaluate the implicit marginal tax resulting from the legislation on pensions in France. We formulate the analytical expressions of this tax and estimate them as a benchmarking example for a single man, born in 1952 with a full career.

1. INTRODUCTION

Contributions to Pay-As-You-Go pension schemes are included in the tax burden along with VAT or income tax. However, the computation rules of pensions rely on contributory principles (Devolder (2005)) that tend to make the benefits received conditional on contributions paid. Hence, considering pension contributions as pure taxes is excessive. Following the study by Feldstein and Samwick (1992) for the United States, we evaluate the fiscal nature of pension contributions for France, by calculating the induced net marginal rate. Explicitly, it consists in using actuarial methods to measure the future amount of additional pension induced by each euro of additional wage.

First, we derive an analytical expression for the implicit marginal tax rate resulting from the specific computation rules of pensions for private employees.

Second, we estimate the implicit marginal tax rate for a man, single, born in 1952 with a complete career, who started working at 21 and retires now at 61.

2. ANALYTICAL EXPRESSIONS

We use actuarial methods, likely present value (LPV) and mortality tables, to estimate the tax consequences of a marginal and instantaneous wage increase. The consequences are twofold:

- In the short run, the contribution is the marginal cost.
  At age x, this marginal cost τ_x can be obtained by taking the derivative of the LPV of payroll taxes with respect to the current wage:

$$\frac{\Delta LPV_x\,(\text{payroll taxes})}{\Delta w_x} = \frac{\Delta}{\Delta w_x}\left(\sum_{y=x}^{R-1} \frac{q_{y,x}}{R_{y,x}} \cdot \tau_y \cdot w_y\right) = \tau_x \ ,$$

  where w_x is the wage at age x, $q_{y,x}$ is the survival probability between age x and y ($y \geq x$), $R_{y,x}$ is the factor of interest between age x and y ($y \geq x$), and R denotes the age of the start of the pension.

- In the long run, the gain is the increase of anticipated pensions.
  At age x, this marginal gain can be obtained by taking the derivative of the LPV of pensions with respect to the current wage:

$$\frac{\Delta LPV_x\,(\text{pensions})}{\Delta w_x} = \frac{\Delta}{\Delta w_x}\left(\sum_{y=R}^{120} \frac{q_{y,x}}{R_{y,x}} \cdot p_y\,(W, I_y)\right) ,$$

  where $p_y(W, I_y)$ is the pension rule with W a vector of the wages, and I_y is a vector of institutional parameters prevailing at age y.

The French Pension System relies on two pillars.

1. The **first pillar** is a defined benefit paid by the CNAV (Caisse Nationale d'Assurance Vieillesse). CNAV's computation formula is given by Legros (2006), Bozio (2006):

$$p_R\,(w, I_R) = \rho\,(R, d, d_{pro.}, d_{cl.}) \cdot \left(\frac{1}{N} \cdot \sum_{w_x \in N \text{best years}} \lambda_{x,R} \cdot \min\,(w_x, SSC_x)\right) , \qquad (1)$$

where

$$\begin{aligned}\rho\,(R, d, d_{pro.}, d_{cl.}) = \ & 0.5 \times \min\left(1, \tfrac{d}{d_{pro.}}\right) \times \big(1 - \alpha_1 \times \max\big(0, \min\big((65 - R) \times 4, d_{b/m} - d\big)\big) \\ & + \alpha_2 \times \max\big(0, \min\big((R - 60) \times 4, d - d_{b/m}\big)\big)\big) .\end{aligned}$$

Here, d is the number of quarters validated, "N best years" denotes the set of the N highest discounted wages, SSC_x is the ceiling basis for social security, $\lambda_{x,R}$ is an updating coefficient of past wages, $d_{pro.}$ and $d_{b/m}$ are the durations used for *pro rata* computation and bonus/malus rates, respectively, $N = 25$ years is the number of best wage-earning years set for the computation of the average wage, α_1 is a penalty (malus) discount factor and α_2 is a reward (bonus) discount factor, equal to 1.25% for each exceeding quarter from January 1st, 2009.

The marginal tax rate for a single worker (no reversion pension) can be obtained as

$$\tau_{\text{marg},x} = \tau_x - \rho\,(.) \cdot \frac{q_{R,x}}{R_{R,x}} \cdot \frac{1_{\text{best years}} \cdot 1_{w_x < SSC_x}}{1 + \tau_x^{emp}} \cdot \frac{\lambda_{x,R}}{N} \cdot \ddot{a}_R \;, \tag{2}$$

with $\tau_x = \tau_x^{SSC} \cdot 1_{w_x < SSC_x} + \tau_x^{totwag}$, where τ_x^{emp} is the payroll tax rate paid by the employer and τ_x^{SSC} and τ_x^{totwag} are the contribution rates applying to the fraction of wage lying below the CNAV ceiling (SSC_x) and the whole wage, with $\ddot{a}_R$ the value of 1 euro pension annuity perceived from age R, indexed by legal factor $I^p_{y,R}$, and with $\rho\,(.)$ the replacement rate. The expression also contains two dummies: the dummy $1_{\text{best years}}$ takes the value 1 if the wage belongs to the 25 "best wage-earning years", and the dummy $1_{w_x < SSC_x}$ takes the value 1 if the wage lies below the ceiling.

2. The **second pillar** is a defined contribution —notional (point) accounts— paid by the ARRCO and/or the AGIRC[1]. All workers of the private sector pay a contribution to ARRCO for the part of their wage below the SSC. The blue collars (resp. white collars) pay a contribution to ARRCO (resp. AGIRC) for the part of their wage beyond the SSC. The amount of pension depends on the number of points accumulated at the date of the liquidation of the pension plan, see Legros (2006):

$$p_R(W, I_R) = \rho\,(.) \cdot \sum_{y=D}^{R-1} \frac{\tau_y \cdot w_y}{v_x^{buy}} \cdot v_x^{ann}. \tag{3}$$

with v_x^{buy} the buying price of one point and v_x^{ann} the annuity value of one point. The coefficient $\rho(.)$ depends on the number of missing quarters compared either to the legal insurance period defined by the CNAV or to the age for which the length of insurance is not taken into account. The marginal tax rate can be written as

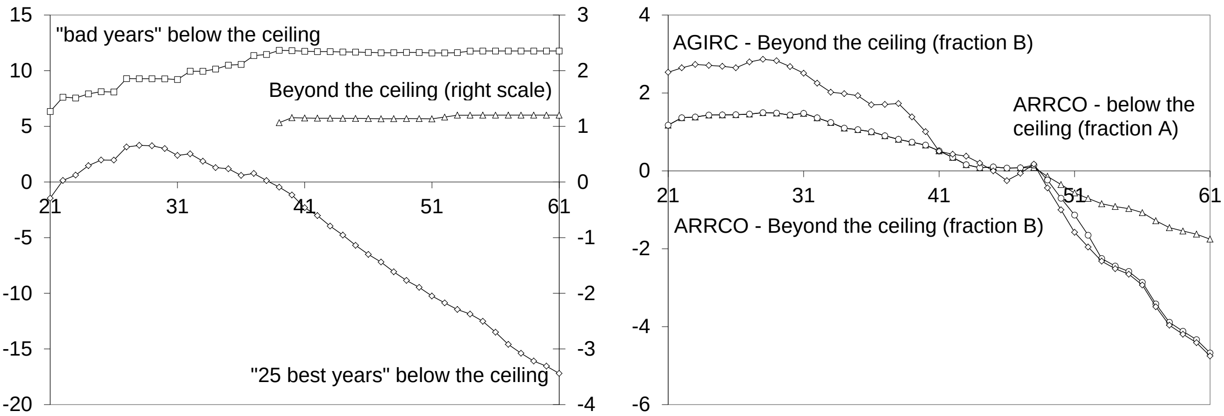
$$\tau_{\text{marg},x} = \tau_x \cdot \left(1 - \rho(.) \cdot \frac{q_{R,x}}{R_{R,x}} \cdot \frac{v_R^{ann}}{v_x^{buy}} \cdot \ddot{a}_R\right). \tag{4}$$

3. COMPUTATION

The benchmark case is a single man born in 1952 with a complete career, who started working at 21 and retires now at 61. Notice that, obviously, no benefits accruing from the reversion pension need to be considered here. People born in 1952 will retire when they are 60 years and 8 months. People born in January 1952 will be allowed to retire from September 2012. Full pension will require 41 years of activity. In our computations, we consider an occupational activity starting at 21 and going on without interruption for 41 years. Retirement age is then reached on the 62nd birthday, which is on January 1st, 2014.

For our prospective analysis, we assume the contribution rates to be constant. To calculate the future values of the points of the supplementary pension plans, we impose that the ratio buying value / liquidation value keeps its trend value. The updating rate of wages and pensions is supposed to be 2% (i.e. long term inflation rate). The discount rate is 4%. We use the TGH/TGF05 mortality tables, which are the prescribed tables for annuities in France.

[1]Association pour le régime de retraite complémentaire des salariés (ARRCO), Association générale des institutions de retraite des cadres (AGIRC).



(a) 1st pillar (CNAV)

(b) 2nd pillar (ARRCO or AGIRC)

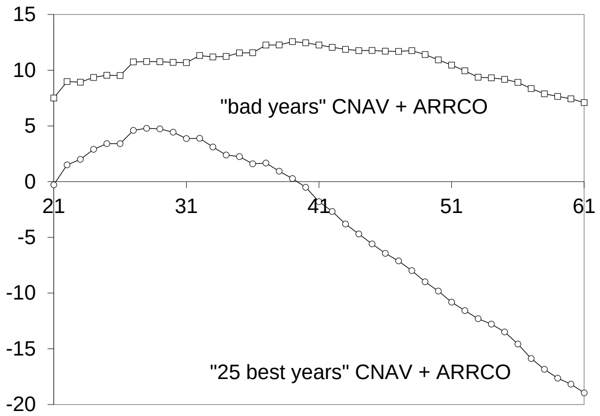
Figure 1: Marginal tax rate of pension contribution for each pillar, with the age on the horizontal axis and the tax rate as a percentage on the vertical axis.

For wages that both lie below the CNAV ceiling (fraction A) and belong to the set of the 25 best wage-earning years, the marginal rate induced by the basic pension regime follows an increasing trajectory with age, from age 21 (-1.45%) to 28 (3.3%). As shown in Figure 1a, the marginal rate is nil about age 38 and rapidly decreases afterwards to reach -17.2% at 61. For wages that lie below the CNAV ceiling without belonging to the 25 best wage-earning years, the marginal rate is exactly equal to the contribution rate. It keeps on increasing until age 39, reaching 11.8%, whereas it is 6.3% at 21. The setting, in 1990, of a CNAV contribution rate on the overall gross wage has a very moderate effect because the contribution rate for the fraction under the CNAV ceiling was lowered. For the basic pension regime, the range of the marginal rate is wide, since the latter can reach 11.8% for the wages of the "bad years" and drop as low as -17.2% for the wages of the "25 best years". As to the wages that are above the CNAV ceiling, the marginal rate is zero until age 38. It is slightly above 1% at age 39 and reaches 1.2% about age 53.

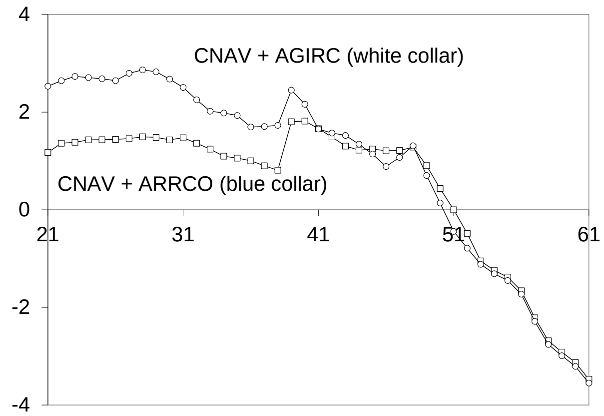
Regarding the supplementary pension plan ARRCO (fraction A of wage), the profile of the marginal rate is slightly modified (Figure 1b). The additional marginal rate is stable and positive at the beginning of the career, where it fluctuates around 1.3% until age 31. This stability is due to the increase of the contribution rate. Afterwards, the marginal rate decreases to stabilize again around 0.1% from age 43 on. This new period of stability is a direct effect of the "repurchase rate" on the contributions, which considerably reduces the purchasing power of points through the contribution. From age 49 on, the marginal rate decreases to reach -1.8% at age 61.

To simplify the presentation of the results, the graph does not show the profile for AGIRC's fraction C (between 4 and 8 times the SSC_x), because it is very similar to that of fraction B (between SSC_x and 4 times SSC_x). Beyond the CNAV ceiling (fraction B of the supplementary pension plans ARRCO and AGIRC), the contribution profiles are rather stable until age 31, because of the historical increase of the ARRCO and AGIRC's contribution rates. Afterwards, the marginal rate decreases until 38. As for the ARRCO's fraction A of the wage, the effect of the repurchase rate applies and the marginal rate stabilizes around 1.5% for ARRCO and 2.6% for AGIRC. This stabilization results in a decrease of the marginal rate with age such that it becomes negative after age 49. The range of fluctuation is less than for the ARRCO's fraction A: between -2% and 2% for the AGIRC and between -2% and 1.5% for the ARRCO. The two plans progressively align with the fraction B with time, which explains why the profiles of the marginal rates are very similar from age 50 on.

The computation rule applying to the supplementary pensions results in a narrower variation interval for the marginal rates:[-1.8%, 1.5%] for ARRCO (fraction A), [-4.8%, 1.5%] for ARRCO (fraction B), [-4.8%, 2.9%] for AGIRC (fraction B).



(a) below the ceiling



(b) beyond the ceiling

Figure 2: Marginal tax rate of pension contribution - summary for each wage fraction, with the age on the horizontal axis and the tax rate as a percentage on the vertical axis.

To summarize the implicit marginal rates for each fraction of wage, we must add all the marginal rates (Figure 2): rates for the fractions of wage below and beyond the CNAV ceiling, rates for the ARRCO's and AGIRC's fractions A and B. For fraction A (Figure 2a), the ranges are: [-19%, 4.8%] for the 25 best wage-earning years and [7.1%, 12.6%] otherwise. For fraction B (Figure 2b), the amplitudes are [-3.5%, 1.8%] for the ARRCO contributors and [-3.6%, 2.9%] for the AGIRC contributors. A significant increase of all the rates of the B fraction occurs at age 39 due to the setting of a CNAV contribution rate (about 1%) applied to the overall wage without any right to retirement attached to it.

4. CONCLUSION

Our computations show that pension contributions in France induce distortions, expressed by an implicit marginal positive or negative taxation of labor, whose amplitude and profile depend on the pension's rules parameters and individual characteristics. Unsurprisingly, the implicit marginal tax rate depends on the computation rules of pensions. The defined benefit system (CNAV) is affected by a greater distortion than the defined contribution system (ARRCO and AGIRC), because the former does not take into account all the wages in the computation of the pension.

Among many possible extensions, we suggest the following four:

1. Our sensitivity analysis would be more accurate if we could use other mortality tables than the TGH/TGF05, which, being too prudential, underestimates future mortality rates.

2. Another way to assess the heterogeneity among individual careers is to rely on samples of historical (Koubi (2002)) or prospective (dynamic microsimulation) career histories. The marginal tax rates could be evaluated according to age and generation, by means of a distribution.

3. Our study focuses on single workers, which restricts the analysis, since the reversion pensions are not taken into account.

4. It could be useful to estimate the likely present value of the costs and benefits induced by an earlier or later retirement, see Hairault et al. (2005).

References

A. Bozio. *Réformes des retraites : estimations sur données françaises*. PhD thesis, EHESS, 2006. URL `http://piketty.pse.ens.fr/fichiers/enseig/memothes/TheseBozio2006.pdf`.

P. Devolder. *Le financement des régimes de retraite*. Economica, Paris, 2005.

M. Feldstein and A. Samwick. Social Security Rules and Marginal Tax Rates. *National Tax Journal*, 45(1):1–22, 1992.

F. Gannon and V. Touzé. Pension Rules and Implicit Marginal Tax Rate in France. Cerp Conference, Turin, 2012. URL `http://www.carloalberto.org/assets/events/2012-cerp-conference/gannonpaper.pdf`.

J.O. Hairault, F. Langot, and T. Sopraseuth. Inciter à différer le départ en retraite : une analyse en termes de courbe de Laffer. *Revue d'Économie Politique*, 115(2):241–263, 2005.

M. Koubi. Eléments de caractérisation des carrières salariales des générations nées entre 1908 et 1980. INSEE Working Paper, F0205, 2002.

F. Legros. NDCs: A Comparison of the French and German Point Systems. In R. Holzmann and E. Palmer, editors, *Pension Reform: Issues and Prospects for Non-Financial Defined Contribution (NDC) Schemes*. The World Bank, 2006.

M.A. Milevsky. *The Calculus of Retirement Income: Financial Models for Pension Annuities and Life Insurance*. Cambridge University Press, 2006.

FAST ORTHOGONAL TRANSFORMS FOR MULTILEVEL QUASI-MONTE CARLO SIMULATION IN COMPUTATIONAL FINANCE

Christian Irrgeher[†§] **and Gunther Leobacher**[†]

[†]*Institute of Financial Mathematics, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria.*

[§]*This author is supported by the Austrian Science Fund (FWF) Project P21943.*

Email: `christian.irrgeher@jku.at, gunther.leobacher@jku.at`

1. INTRODUCTION

Many areas of computational finance, e.g. derivative pricing in a Gaussian model, require the approximation of the expected value of some function f depending on an n-dimensional Gaussian vector X, i.e. $\mathbb{E}(f(X))$. It is a trivial observation that for every orthogonal transform $U : \mathbb{R}^n \to \mathbb{R}^n$ the identity

$$\mathbb{E}\left(f\left(X\right)\right) = \mathbb{E}\left(f\left(UX\right)\right)$$

holds. While this does not change the problem from the probabilistic point of view, it does make a difference for quasi-Monte Carlo methods.

Introducing an orthogonal matrix is closely related to the construction of discrete Brownian paths. Assume we are interested in $\mathbb{E}(g(B))$, where g is some function of a standard Brownian motion B with index set $[0, T]$. To apply simulation methods we have to discretize the problem such that

$$\mathbb{E}\left(g\left(B\right)\right) \approx \mathbb{E}\left(\tilde{g}\left(B_\tau, B_{2\tau}, \ldots, B_{n\tau}\right)\right)$$

with $n \in \mathbb{N}$, $\tau = T/n$ and a suitable function $\tilde{g} : \mathbb{R}^n \to \mathbb{R}$. The discrete Brownian path with covariance matrix $\Sigma = \tau\left(\min(j,k)\right)_{j,k=1}^n$ can be constructed from an standard Gaussian vector by using an $n \times n$-matrix A with $AA^t = \Sigma$. Papageorgiou (2002) observed that $AA^t = \Sigma$ if and only if $A = SU$ for some orthogonal matrix U and matrix S given by

$$S = \tau \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 1 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \ldots & 1 \end{pmatrix}.$$

Thus, every linear construction of $(B_\tau, \ldots, B_{n\tau})$ given by A corresponds to an orthogonal transform U.

There are three classical and well-known methods to construct a Brownian path. The forward method is given by $U = I$, the Brownian bridge construction corresponds to the inverse Haar transform, i.e. $U = H^{-1}$ with H denoting the matrix of the Haar transform, and the principal component analysis (PCA) construction is related to the singular value decomposition of the covariance matrix $\Sigma = VD^2V^t$ with $U = S^{-1}VD$. More recent construction methods try to determine an orthogonal transform that is tailored to the underlying integration problem, see Imai and Tan (2007) or Irrgeher and Leobacher (2012).

There are some theories that explain why a suitable orthogonal transform might increase the efficiency of quasi-Monte Carlo methods, see e.g. Caflisch et al. (1997). However, there is also a drawback of this approach. In general, the computation of an orthogonal transform requires $O(n^2)$ floating point operations. Therefore we concentrate on fast orthogonal transforms, which are orthogonal transforms with cost of the order $O(n\log(n))$. In Leobacher (2012) various examples of fast orthogonal transforms are studied including the discrete sine and cosine transform as well as Walsh and Haar transform.

2. MULTILEVEL QUASI-MONTE CARLO INTEGRATION AND THE REGRESSION ALGORITHM

The expected value can be approximated by an equally weighted quadrature rule

$$\mathbb{E}(f(X)) \approx \frac{1}{N}\sum_{i=1}^{N} f(\Phi^{-1}(x_i)) \tag{1}$$

where Φ is the cumulative distribution function of the standard normal distribution and with sample points $\{x_1, \ldots, x_N\} \subset [0,1)^n$. If the sample points are elements of a low-discrepancy sequence (e.g. a Halton sequence or a Sobol sequence), the quadrature rule of (1) is called a quasi-Monte Carlo method. The efficiency of (quasi-)Monte Carlo methods can be improved both in accuracy and in computing time by combining different discretization levels, see Giles (2008) and Giles and Waterhouse (2009).

For $\ell \in \mathbb{N}_0$, X^ℓ denotes an 2^ℓ-dimensional Gaussian vector and C_ℓ is an operator such that $C_\ell X^\ell$ is an $2^{\ell-1}$-dimensional Gaussian vector. Assume that we want to compute the expected value of a function $f^L : \mathbb{R}^{2^L} \to \mathbb{R}$ with $L \in \mathbb{N}_0$. Furthermore, we assume that at each discretization level $\ell = 0, \ldots, L-1$ there exists a function $f^\ell : \mathbb{R}^{2^\ell} \to \mathbb{R}$ with $\mathbb{E}(f^L(X^L)) \approx \mathbb{E}(f^\ell(X^\ell))$. The main idea of multilevel quasi-Monte Carlo methods is to use the identity

$$\mathbb{E}\left(f^L\big(X^L\big)\right) = \mathbb{E}\left(f^0\big(X^0\big)\right) + \sum_{\ell=1}^{L} \mathbb{E}\left(f^\ell\big(X^\ell\big) - f^{\ell-1}\big(C_\ell X^\ell\big)\right) \tag{2}$$

and compute each of these expected values of the right hand side with quasi-Monte Carlo separately. This approach becomes useful if the expected values can be approximated to the required level of accuracy using less function evaluations for bigger ℓ while the costs per function evaluation increases.

In Irrgeher and Leobacher (2012) a method to determine a fast orthogonal transform tailored to the underlying integration problem is presented for square-integrable functions f of the form $f(X) = g(h_1(X), \ldots, h_m(X))$ with $g : \mathbb{R}^m \to \mathbb{R}$, $h_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, \ldots, m$ and $m \ll n$. Every function h_k is approximated by an affine function using a "linear regression" approach. For that we minimize the functional

$$\mathbb{E}\left(\left(h_k(X) - a_k^t X - b_k\right)^2\right),$$

where $a_k \in \mathbb{R}^n$ and $b_k \in \mathbb{R}$. The first order conditions are given by $a_{k,j} = \mathbb{E}\left(X_j h_k(X)\right)$ and $b_k = \mathbb{E}\left(h_k(X)\right)$ with $j = 1, \ldots, n$ and $k = 1, \ldots, m$. So we get the approximation $f(X) \approx \bar{g}(X) := \bar{\bar{g}}\left(a_1^t X, \ldots, a_m^t X\right)$ and Wang and Sloan (2011) noted that there exists an orthogonal transform U such that $\bar{g} \circ U$ is at most m-dimensional. This orthogonal transform U can be determined as a product of at most m Householder reflections and each Householder reflection can be applied in $O(n)$.

Regression Algorithm: Let f be of the form $f(X) = g(h_1(X), \ldots, h_m(X))$ where X is a standard Gaussian vector.

1. Start with $k, \ell = 1$ and $U = I$;
2. $a_{k,j} := \mathbb{E}(X_j h_k(UX))$ for $j = k, \ldots, n$;
3. $a_{k,j} := 0$ for $j = 1, \ldots, k-1$;
4. if $\|a_k\| = 0$ go to 7;
5. else let U_ℓ be a Householder reflection that maps e_ℓ to $a_k / \|a_k\|$;
6. $U = UU_\ell$; $\ell = \ell + 1$;
7. $k = k + 1$;
8. while $k \leq m$, go back to 2;
9. Compute $\mathbb{E}(f(UX))$ using QMC.

Note that the algorithm is practicable if the $a_{k,j}$ can be computed efficiently. Moreover, it can be applied to multilevel quasi-Monte Carlo quite well. Therefore, we assume that $f^\ell(X) = \psi^\ell(h^\ell(X))$ for all $\ell = 1, \ldots, L$. Rewriting identity (2) gives

$$\mathbb{E}\left(f^L(X^L)\right) = \mathbb{E}\left(f^0(X^0)\right) + \sum_{\ell=1}^{L} \mathbb{E}\left(g^\ell\left(h_1^\ell(X^\ell), h_2^\ell(X^\ell)\right)\right)$$

with $h_1^\ell = h^\ell$, $h_2^\ell = h^{\ell-1} \circ C_\ell$ and $g^\ell(y_1, y_2) = \psi^\ell(y_1) - \psi^{\ell-1}(y_2)$. If we apply orthogonal transforms $U^\ell : \mathbb{R}^{2^\ell} \to \mathbb{R}^{2^\ell}$ at each level, we get

$$\mathbb{E}\left(f^L(X^L)\right) = \mathbb{E}\left(f^0(X^0)\right) + \sum_{\ell=1}^{L} \mathbb{E}\left(g^\ell\left(h_1^\ell(U^\ell X^\ell), h_2^\ell(U^\ell X^\ell)\right)\right).$$

Applying the regression algorithm we obtain for each level ℓ a suitable orthogonal transform of the form $U^\ell = U_1^\ell U_2^\ell$ with Householder reflections U_1^ℓ and U_2^ℓ, where U_1^ℓ corresponds to the fine discretization and U_2^ℓ corresponds to the coarse discretization.

3. ASIAN OPTIONS - A NUMERICAL EXAMPLE

We consider an Asian call option with arithmetic average and want to price the option in the Black-Scholes model using multilevel quasi-Monte Carlo simulation. For the time discretization with 2^ℓ points, $\ell \in \mathbb{N}_0$, the payoff function $f^\ell : \mathbb{R}^{2^\ell} \to \mathbb{R}$ is given by

$$f^\ell(X^\ell) = \max\left(\frac{1}{2^\ell} \sum_{k=1}^{2^\ell} S_k(X^\ell) - K, 0 \right)$$

with a fixed strike price K. Under the risk-neutral measure the discrete path of the stock price process S is given by

$$S_k(X^\ell) = S_0 \exp\left(\left(r - \frac{\sigma^2}{2} \right) k \frac{T}{2^\ell} + \sigma\sqrt{\tau} \sum_{i=1}^{k} X_i^\ell \right)$$

with interest rate r and volatility σ. The payoff function can be written as $f^\ell = g^\ell \circ h^\ell$ with $g^\ell(y) = \max(y - K, 0)$ and $h^\ell(X^\ell) = \frac{1}{2^\ell} \sum_{k=1}^{2^\ell} S_k(X^\ell)$. Thus, the regression algorithm can be used to determine the orthogonal transform at each discretization level with 2^ℓ points. It requires the efficient computation of the vectors a_1^ℓ, a_2^ℓ. This can be done analytically and we get

$$a_{1,j}^\ell = \mathbb{E}\left(X_j h_1^{(\ell)} \right) = \sum_{k=j}^{2^\ell} \frac{S_0\,\sigma}{2^\ell} \sqrt{\frac{T}{2^\ell}} \exp\left(\frac{rkT}{2^\ell} \right),$$

$$a_{2,j}^\ell = \mathbb{E}\left(X_j h_2^{(\ell)} \right) = \sum_{k=\left\lfloor \frac{j-1}{2} \right\rfloor + 1}^{2^{\ell-1}} \frac{S_0\,\sigma}{2^{\ell-1}} \sqrt{\frac{T}{2^\ell}} \exp\left(\frac{rkT}{2^{\ell-1}} \right)$$

with $j = 1, \ldots, 2^\ell$.

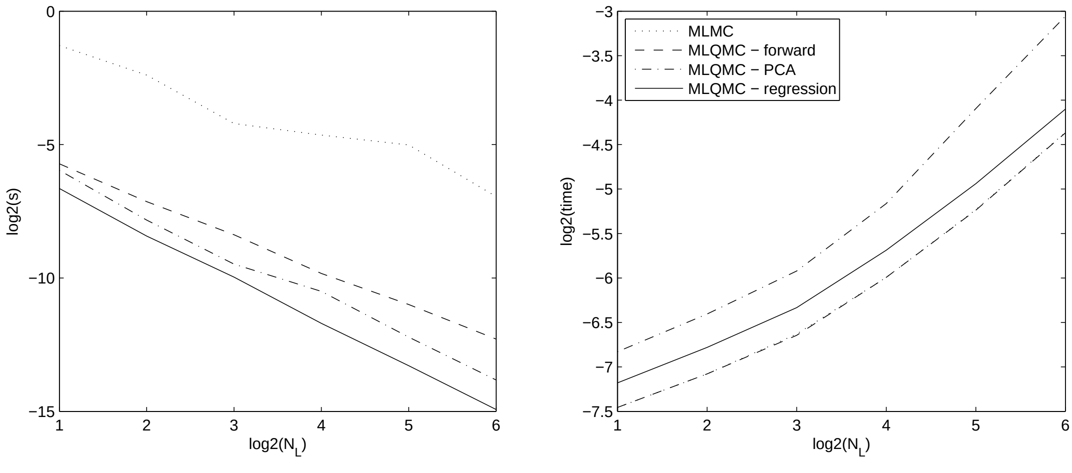


Figure 1: Asian call option. Comparison of the sample standard deviation (left) as well as the computing time (right) based on 1000 runs for different construction methods.

In the numerical analysis we compare the multilevel QMC method combined with the regression algorithm with multilevel Monte Carlo and multilevel quasi-Monte Carlo with forward construction and PCA, respectively. Therefore, we choose the parameters $r = 0.04$, $\sigma = 0.3$, $K = 100$, $S_0 = 100$ and $T = 1$. At the finest level we start with 2^{10} discretization points, i.e. $L = 10$ and the number of sample points are doubled at each level starting with N_L sample points at the finest level L. For the QMC approaches we take a Sobol sequence with a random shift. In Figure 1 we compare for different values of N_L both the sample standard deviation and the computing time of the price of the Asian call option based on 1000 independent runs. As we can see, the regression algorithm yields the lowest standard deviation, but the average computing time of the regression algorithm is slightly worse than the forward method. However, the regression algorithm is better than the PCA construction measured in both standard deviation and computing time.

References

R. Caflisch, W. Morokoff, and A. Owen. Valuation of mortgage-backed securities using Brownian bridges to reduce effectice dimension. *Journal of Computational Finance*, 1(1):27–46, 1997.

M. B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.

M. B. Giles and B. J. Waterhouse. Multilevel quasi-Monte Carlo path simulation. In H. Albrecher, editor, *Advanced financial modelling. Radon Series on Computational and Applied Mathematics*, volume 8, pages 165–181. Walter de Gruyter, Berlin, 2009.

J. Imai and K. S. Tan. A general dimension reduction technique for derivative pricing. *Journal of Computational Finance*, 10:129–155, 2007.

C. Irrgeher and G. Leobacher. Fast orthogonal transforms for pricing derivatives with quasi-Monte Carlo. In C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, editors, *Proceedings of the 2012 Winter Simulation Conference*, 2012.

G. Leobacher. Fast orthogonal transforms and generation of brownian paths. *Journal of Complexity*, 28(2):278–302, 2012.

A. Papageorgiou. The Brownian bridge does not offer a consistent advantage in quasi-Monte Carlo integration. *Journal of Complexity*, 18(1):171–186, 2002.

X. Wang and I. H. Sloan. Quasi-Monte Carlo methods in financial engineering: An equivalence principle and dimension reduction. *Operations Research*, 59(1):80–95, 2011.

MORTALITY SURFACE BY MEANS OF CONTINUOUS-TIME COHORT MODELS

Petar Jevtić[†], Elisa Luciano[‡] and Eligna Vigna[§]

[†] *V. Pareto Doctoral School, Università degli Studi di Torino, corso Unione Sovietica, 218 bis, Italy*
[‡] *Department of Economics and Statistics, Università degli Studi di Torino, corso Unione Sovietica, 218 bis, Italy, Collegio Carlo Alberto*
[§] *Department of Economics and Statistics, Università degli Studi di Torino, corso Unione Sovietica, 218 bis, Italy, Collegio Carlo Alberto, CeRP*
Email: petar.jevtic@studenti.unito.it,
elisa.luciano@unito.it, elena.vigna@unito.it

1. INTRODUCTION

Insurance companies and pension funds are exposed to mortality risk and hope for the development of a liquid and transparent longevity-linked capital market. Active trading of mortality derivatives would help them assessing and hedging the risks they are exposed to, in the same manner as financial markets help them mutualize financial risks. Mortality-risk appraisal consisting in an accurate but easy-to-handle description of human survivorship is fundamental in this respect. In spite of this need, no consensus has been reached yet on the best way to model mortality risk.

In situations where we need to combine the appraisal of mortality and financial risk, the adoption of a continuous–time approach proves useful. In addition, the motivation for adopting a continuous-time description can be found in the search for closed-form evaluation formulas for insurance products and their derivatives.

Continuous-time stochastic mortality models for single generation have been considered by a number of researchers, including Milevsky and Promislow (2001), Dahl (2004), Biffis (2005), Cairns et al. (2006), Schrager (2006) and Luciano and Vigna (2008). A theoretical extension of the continuous-time single-generation model to the mortality surface appears in Biffis and Millossovich (2006) and is followed by Blackburn and Sherris (2012) who make the assumption of perfect correlation across generations and also focus on the calibration aspect.

In order to reconcile the calibration of the whole mortality surface with actuarial practice, which suggests high but not perfect correlation, we fit in Jevtić et al. (2013) the mortality surface by means of a continuous-time cohort model that is able to capture correlations across generations. As a relevant consequence, this model provides the actuary with a calibrated correlation among generations rather than a “best estimate” one. Given the same initial age, the intensities of several generations are written in terms of factors, identified via PCA. The Differential Evolution (DE) algorithm is a robust stochastic search and optimization algorithm which already proved its use across a wide range of engineering applications. In Jevtić et al. (2013), we use it to fit the mortality surface with an extreme precision and discover that the fitted parameters are robust, stable and lead to correlations across generations that are high but less than one. We report here the main results, the interested reader is referred to Jevtić et al. (2013) for more details.

2. SIMPLE APC MODEL

2.1. MODEL

Stochastic mortality of a given generation is described by means of a Cox process, as in Biffis (2005). To give the reader a flavor of the results obtained in the paper Jevtić et al. (2013), we present without derivation our two factor Simple Age-Period-Cohort (APC) model where we define mortality intensity for each cohort i from the set of cohorts under consideration such that

$$\mu^i(t) \stackrel{\text{def}}{=} X_1(t) + X_2^i(t),$$

having

$$\begin{aligned} dX_1(t) &= \psi_1 X_1 dt + \sigma_1 dZ_1(t), \\ dX_2^i(t) &= \psi_2 X_2 dt + \sigma_2 \rho^i dZ_1(t) + \sigma_2 \sqrt{1-(\rho^i)^2} dZ_2(t), \end{aligned}$$

where $\psi_1, \psi_2 \in \mathbb{R},\ \sigma_1, \sigma_2 \in \mathbb{R}^+, \boldsymbol{\rho} \in [-1, 1]$ and $Z_1(t)$ and $Z_2(t)$ are orthogonal BMs. In this setting, the survival probability to time τ, for a life aged x at time $t = 0$, is given by

$$S^i(0, \tau) = \mathbb{E}\big[e^{-\int_0^\tau \mu^i(s) ds}\big],$$

which gives rise to the analytical solution (c.f. Duffie et al. (2000))

$$S^i(0, \tau) = e^{\hat{\alpha}^i(\tau) + \hat{\beta}_1(\tau) X_1^i(0) + \hat{\beta}_2(\tau) X_2^i(0)},$$

where we have for j in $\{1, 2\}$,

$$\hat{\beta}_j(\tau) = -\int_0^\tau e^{\psi_j(\tau - s)} ds = \frac{1}{\psi_j}\Big(1 - e^{\psi_j \tau}\Big)$$

and

$$\begin{aligned} \hat{\alpha}^i(\tau) = & \sum_{j=1}^{2} \frac{\sigma_j^2}{2\psi_j^3}\left(\psi_j \tau - 2e^{\psi_j \tau} + \frac{1}{2} e^{2\psi_j \tau} + \frac{3}{2}\right) \\ & + \frac{\rho^i \sigma_1 \sigma_2}{\psi_1 \psi_2}\left(\tau - \frac{e^{\psi_1 \tau}}{\psi_1} - \frac{e^{\psi_2 \tau}}{\psi_2} + \frac{e^{(\psi_1+\psi_2)\tau}}{\psi_1 + \psi_2} + \frac{\psi_1^2 + \psi_1\psi_2 + \psi_2^2}{\psi_1 \psi_2 (\psi_1 + \psi_2)}\right). \end{aligned}$$

A relevant feature of our model is that it enables the derivation of formulas for instantaneous correlations among mortality intensities of different generations.

2.2. CALIBRATION

We consider the male population of the United Kingdom. Our data set consists of cohort death rates for a life aged $x = 40$, which we examine until they have reached the age of 59 having $\tau = 1, \ldots, 19$. The generations i span from 1900-1950, with a 5-year increment.

In this model, four parameters are common to all generations, denoted by $[\psi_1, \psi_2, \sigma_1, \sigma_2]$, and three parameters are specific to each generation, denoted by $[\rho^i, X_1^i(0), X_2^i(0)]$. Since intensity is Gaussian distributed and thus and can become negative with a positive probability, we a priori set this probability to be maximum 1%, and during calibration procedure, along common, we calibrate only two parameters $[\rho^i, X_1^i(0)]$ specific to each generation.

To calibrate our model, we use the Differential Evolution algorithm, a stochastic search and optimization method which we adapt in order to suite the unique needs of our setting.

2.3. RESULTS

The residuals plot for the entire region demonstrates that the quality of calibration is exceptionally high as can be seen in Figure 1. Apart from one observation, all residuals stay in the range $[-2 \times 10^{-3}, 2 \times 10^{-3}]$. Moreover, no structural patterns can be observed.

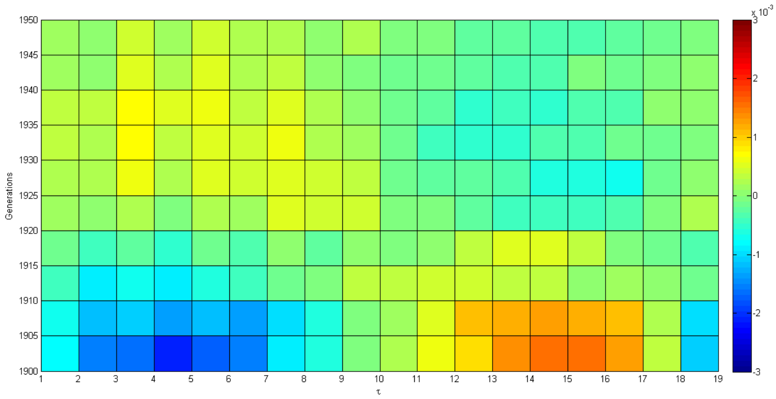


Figure 1: Calibration residuals plot

The constraint of 1% on the probabilities of negative intensities is respected for all relevant durations $\tau \in \{1, 2, ..., 69\}$ and for all relevant generations. In most cases, it is well below this level which can be observed in Figure 2.

In Figure 3, we can see the forecasting error is remarkably small and even below 1% for in-sample data. However, it increases up to 26% for out-of-sample data in case of $\tau = 40$ and has the tendency of an increase afterwards. This is in accordance with the existing literature.

In Figure 4, we plot the survival curve at time $S(1, \tau)$, for generation 1950, as a function of τ. The stochastic mortality framework is characterized by the fact that at $t = 0$ – when the calibration is performed – the survival curve which will apply one year later, at $t = 1$, is a random variable.

According to this formula, 100,000 simulations are made and we report of them the median (green line), the 5th percentile (red line) and the 95th percentile (blue line) as functions of τ. The figure also shows the survival probabilities as observed one year later (dots). The in-sample

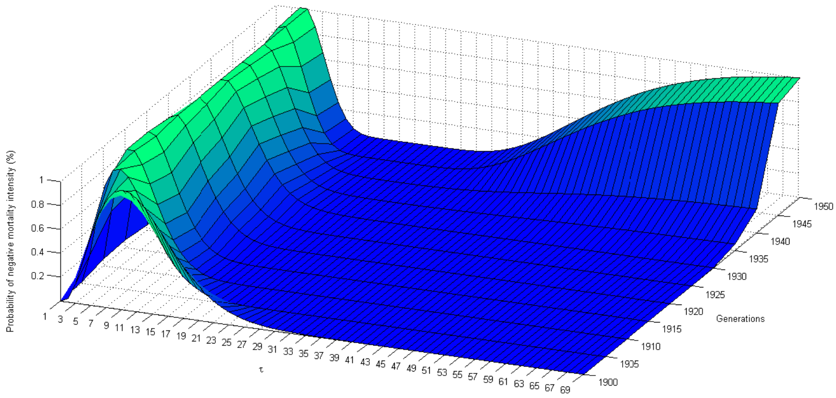


Figure 2: Probability of negative mortality intensities surface

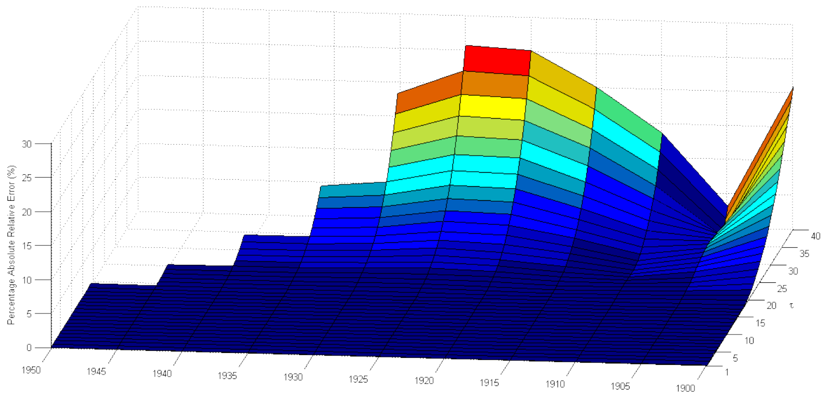


Figure 3: Percentage Absolute Relative Error of Survival Probability Surface

forecasting is very accurate. Almost all the survival probabilities lie on the median (or very close to it), and therefore they stay in the 90% confidence interval.

In Table 1, we observe that the instantaneous correlations among mortality intensities of different generations are positive and high, which is in accordance with actuarial intuition. They stay between 0.94 and 1.00 and, as expected, tend to decrease with the difference in years of birth. To the best of our knowledge, this is the first research that provides the actuary with a calibrated and sensible correlation of mortality intensity among different generations.

3. CONCLUSIONS

This paper is a first attempt to construct an effective cohort-based continuous-time factor model of the mortality surface. We cast the model first in the affine framework, and specialize it then to Ornstein-Uhlenbeck factors. The resulting longevity intensity model extends the G2++ interest-rate model, since the factors have different weights for each generation. The main novelty of the model with respect to existing literature is that it allows for imperfect correlation of mortality intensity across generations.

The model is implemented on UK data for the generations born between 1900 and 1950, using

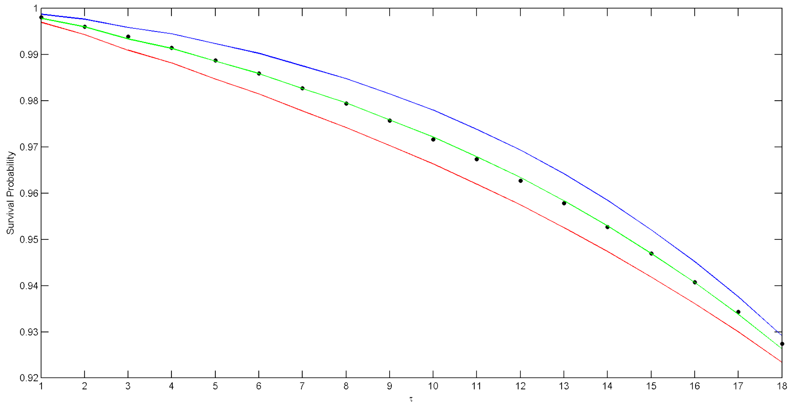


Figure 4: Survival probability curve at $t = 1$ $S(1;\tau)$ for generation 1950

	1900	1905	1910	1915	1920	1925	1930	1935	1940	1945	1950
1900	1.0000										
1905	0.9601	1.0000									
1910	0.9497	0.9993	1.0000								
1915	0.9489	0.9992	0.9999	1.0000							
1920	0.9515	0.9995	0.9999	0.9999	1.0000						
1925	0.9491	0.9993	0.9999	1.0000	0.9999	1.0000					
1930	0.9496	0.9993	1.0000	0.9999	0.9999	0.9999	1.0000				
1935	0.9584	0.9999	0.9995	0.9994	0.9997	0.9995	0.9995	1.0000			
1940	0.9693	0.9993	0.9975	0.9973	0.9979	0.9973	0.9975	0.9991	1.0000		
1945	0.9961	0.9810	0.9735	0.9729	0.9749	0.9731	0.9735	0.9798	0.9872	1.0000	
1950	1.0000	0.9601	0.9497	0.9489	0.9515	0.9491	0.9496	0.9584	0.9693	0.9961	1.0000

Table 1: Table of correlations

HMD data for the period 1900-2008. On these data, two factors are deemed as a reasonable first choice. Calibration by means of stochastic search and the Differential Evolution optimization algorithm proves to produce small errors and yields robust and stable parameters. Standard criteria desirable for a model of the mortality surface are satisfied.

The calibration confirms that correlation across generations is very high but smaller than one. Up to our knowledge, this is the first calibration of the correlation among mortality intensities of different generations in the academic literature. The calibrated correlations turn out to be sensible and intuitive. The possibility of capturing these correlations is owed to the combination of a generation-based model and DE driven calibrations, and is our major contribution.

References

E. Biffis. Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37(3):443–468, 2005.

E. Biffis and P. Millossovich. A bidimensional approach to mortality risk. *Decisions in Economics* *and Finance*, 29(2):71–94, 2006.

C. Blackburn and M. Sherris. Consistent dynamic affine mortality models for longevity risk applications. *UNSW Australian School of Business Research Paper No. 2011ACTL08*, 2012.

A.J.G. Cairns, D. Blake, and K. Dowd. Pricing death: Frameworks for the valuation and securitization of mortality risk. *Astin Bulletin*, 36(1):79–120, 2006.

M. Dahl. Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance: Mathematics and Economics*, 35(1):113–136, 2004.

D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

P. Jevtić, E. Luciano, and E. Vigna. Mortality surface by means of continuous time cohort models. *Insurance: Mathematics and Economics*, 53(1):122 – 133, 2013.

E. Luciano and E. Vigna. Mortality risk via affine stochastic intensities: calibration and empirical relevance. *Belgian Actuarial Bulletin*, 8(1):5–16, 2008.

M.A. Milevsky and D. Promislow. Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, 29(3):299–318, 2001.

D.F. Schrager. Affine stochastic mortality. *Insurance: Mathematics and Economics*, 38(1):81–97, 2006.

DEMOGRAPHIC RISK TRANSFER: IS IT WORTH FOR ANNUITY PROVIDERS?

Elisa Luciano and Luca Regis

Department of Economics and Statistics, Corso Unione Sovietica 218/bis, 10134, Torino.
Email: elisa.luciano@unito.it, luca.regis@unito.it

1. INTRODUCTION

Annuity providers are exposed to longevity risk – i.e. the risk of unexpected improvements in the survivorship of their insureds – as well as to financial risks on both assets and liabilities, as soon as the latter are fairly evaluated. In this paper, we aim at discussing the optimal management of such risks, when they are tackled together in an asset-liability model. In particular, we are interested in analysing whether an annuity provider should better transfer systematic longevity risk to a reinsurer or a special purpose vehicle - as most of the recent deals do - or remaining exposed to it, while saving on the costs of the transfer. We assess this trade-off in a model which allows us to obtain closed-form expressions for the expected financial return of the fund and its risk, measured through a value-at-risk measure. For the sake of simplicity we use first-order approximations and show that, if the transfer is fairly priced and the aim of the fund is to maximize returns, the funds' alternatives can be represented in the plane expected return-VaR. We build a risk-return frontier, along which the optimal transfer choices of the fund are located. We disentangle the demographic and financial component of the overall funds' risk.
Our paper departs from the literature on optimal longevity transfers (Biffis and Blake (2010), Barrieu and Loubergé (2013)) because we explore the choice of the fund in the context of an ALM model. Delong et al. (2008) studied the asset allocation problem of a pension fund in the accumulation phase in the presence of systematic mortality. Most of the papers analyzing the asset management of pension funds, instead, focused on idiosyncratic mortality risk, but neglected its systematic component (Hainaut and Devolder (2007), Battocchio et al. (2007)). Our paper focuses on the transfer of this systematic (or aggregate) risk, given that previous works (Hari et al. (2011)) found it to be far more important than the idiosyncratic one in large and well-diversified portfolios. More details and an application to UK data can be found in Luciano and Regis (2012).

2. SET UP

We consider a stylized ALM model of a pension fund which has issued a single annuity on a head aged x. The fund can either

- transfer demographic risk to a reinsurer - which in turn hedges and prices it fairly - or
- suffer it without hedging.

At the same time, the fund can invest its collected premiums either in bonds or keep them as cash. It is supposed to maximize expected financial returns. We assume that hedging of demographic risk on the part of the reinsurer and measurement of financial risk on the part of the fund is done up to first-order discrepancies between actual and forecasted interest and survival rates. We perform a Delta analysis, but extension to second order hedges, or Delta-Gamma hedging, is quite obvious. In order to make measurement and management of demographic and financial risk feasible in closed form, we place ourselves in a standard, continuous-time framework. Consider a time interval $\mathcal{T} = [0, T], T < \infty$, a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a multidimensional standard Wiener $W(\omega, t), t \in \mathcal{T}$. The space is endowed with the filtration generated by w, $\mathbf{F}^w = \{\mathcal{F}_t\}$. We adopt a stochastic extension of the classical Gompertz law for mortality description and we stick to the Hull-White model for interest-rate risk, as in Luciano et al. (2012a). For what concerns longevity risk modelling, we assume that the death of an individual belonging to a generation x is the first jump time of a Poisson process with stochastic intensity $\lambda_x(t)$, which has the following dynamics:

$$d\lambda_x(t) = a_x \lambda_x(t)dt + \sigma_x dW_x(t),$$

where $a_x > 0$, $\sigma_x \geq 0$, W_x is a standard one-dimensional Brownian motion in W. On top of being parsimonious, the model provides a closed-form expression for the survival probability of head x at any point in time t and up to any horizon T :

$$S_x(t,T) = \frac{S_x(0,T)}{S_x(0,t)} \exp\left[-X_x(t,T)I_x(t) - Y_x(t,T)\right],$$

where

$$X_x(t,T) := \frac{\exp(a_x(T-t)) - 1}{a_x},$$

$$Y_x(t,T) := -\frac{\sigma_x^2 \left[1 - \exp\left(2a_x t\right)\right] X_x(t,T)^2}{4a_x},$$

$$I_x(t) := \lambda_x(t) - f_x(0,t).$$

and $I_x(t)$ - the difference between the actual mortality intensity of generation i at time t and its forward value or forecast at time 0, $f_x(0,t)$ - is what we interpret as the *mortality or demographic risk factor*. It is the discrepancy between realization and forecast which makes the pension fund exposed to mortality risk.

For what concerns financial risk, we choose the standard Vasicek model for interest rates. The spot rate has the following dynamics under a measure $\mathbb{Q}$ equivalent to $\mathbb{P}$:

$$dr(t) = g(\theta - r(t))dt + \Sigma dW_F(t),$$

where $\theta, g > 0, \Sigma > 0$ and W_F is a univariate Brownian motion independent of $W_x(t)$ for all t. The corresponding zero-coupon bond price - if the bond is evaluated at time t and has maturity T - is

$$B(t,T) = \frac{B(0,T)}{B(0,t)} \exp\left[-\bar{X}(t,T)K(t) - \bar{Y}(t,T)\right],$$

where

$$\bar{X}(t,T) := \frac{1 - \exp(-g(T-t))}{g},$$
$$\bar{Y}(t,T) := \frac{\Sigma^2}{4g}\left[1 - \exp(-2gt)\right]\bar{X}^2(t,T),$$
$$K(t) := r(t) - R(0,t).$$

$K(t)$ is the *financial risk factor*, akin to the demographic factor $I_x(t)$. As in the longevity case, the financial risk factor is the difference between actual and forecasted rates for time t, where the forecast is done at time 0. It is the only source of randomness which affects bonds.

3. PORTFOLIO RISKS AND DEMOGRAPHIC RISK TRANSFER

Consider an annuity issued on an individual of generation x. Make the annuity payment per period equal to one. The fair price of the annuity - which lasts until the extreme age ω - is

$$V_i^A(t) = \sum_{T=t+1}^{\omega - x} S_i(t,T)B(t,T)$$

at time $t \geq 0$. It can be shown (see Luciano et al. (2012b)) that the change on the fair value due to changes in the longevity risk factor can be approximated up to the first order as follows:

$$\Delta V_x^{AM}(t) = \Delta_A^M(t)\Delta I_x(t),$$

where the Delta is

$$\Delta_A^M(t) = -\sum_{u=t+1}^{\omega - x} B(t,u)S_x(t,u)X_x(t,u) < 0.$$

From now on, we assume that the pension fund has issued such contract at a price $P \geq V_i^A(0)$ and can

- either run into demographic risk, evaluated at its first-order impact $\Delta_A^M(t)\Delta I_i(t)$, or
- transfer the risk to a reinsurer or to a special purpose vehicle at a fair cost C.

On top of being exposed to demographic risk, the fund is exposed to financial risk coming both from the asset side and the liabilities side. Any bond which enters the assets of the fund are subject to interest rate fluctuations. The first-order sensitivity to changes in K, denoted by ΔK, of a bond is given by:

$$\Delta_B^F(t,T) = -B(t,T)\bar{X}(t,T) < 0.$$

Also the annuity value, which enters the liabilities, is subject to financial risk, since it is fairly priced. The effect of a change in K on the annuity is:

$$\Delta V_i^{AF}(t) = \Delta_A^F(t)\Delta K(t),$$

where

$$\Delta_A^F(t) = -\sum_{u=t+1}^{\omega-x} B(t,u)S_i(t,u)\bar{X}(t,u) < 0.$$

4. RETURN MAXIMIZATION AND INVESTMENT STRATEGIES

At time t, the fund maximizes expected returns at $t+\Delta t$, by investing in bonds, if profitable to him, either $P-C$, if he transferred demographic risk, or P, if he did not. As a result, the fund has a portfolio made up by the annuity (short) and long n^* bonds, whose instantaneous expected return μ is

$$\mu = \mathbb{E}_t\left[-V_A^F(t+dt) + V_A^F(t) + n^*\left[B(t+dt,T) - B(t,T)\right]\right].$$

Since only the second part depends on n^*, the fund chooses this number as high as possible if $\mathbb{E}_t B(t+dt,T) > B(t,T)$, or equal to zero in the opposite case. Using first-order approximations for returns over the time interval Δt, this condition is verified if and only if

$$\mathbb{E}_t[K(t+\Delta t)] < 0. \tag{1}$$

Let us denote by C^* the amount paid for demographic risk transfer at time t, when the hedging strategy is set up. Depending on the fund's choice, we may have $C^* = C$ or $C^* = 0$. This choice of C^* and the asset allocation decision lead to the identification of four strategies, whose characteristics are described in Table 1. Financial returns are evaluated at a certain horizon $t+\Delta t$ and are net of the costs $C^*_{\Delta t}$ of demographic-risk transfer which can be imputed to the time interval Δt. Let us introduce the following notation:

$$\begin{aligned}
\alpha &:= \sum_{u=t+1}^{\omega-x} B(t,u)S_x(t,u)X_x(t,u) > 0,\\
\beta &:= \sum_{u=t+1}^{\omega-x} B(t,u)S_x(t,u)\bar{X}(t,u) > 0,\\
\gamma &:= \beta - P\bar{X} < \beta,\\
\delta &:= \gamma + C\bar{X} > \gamma.
\end{aligned}$$

Then for strategies 1 and 2, α is the Delta of the portfolio with respect to mortality risk, while β, γ and δ are the Deltas of the portfolios for the four strategies with respect to financial risk.

A risk evaluation of the VaR-type is constructed for the four strategies at a confidence level ϵ. Due to independence between financial and actuarial risk sources, if we sum up the appropriate scenario-based risks or VaRs (where appropriate stands for "based on the need of selecting VaR_ϵ

Strategy	n^*	C^*	**Dem risk**	**Fin risk**	**Net expected return**
1	0	0	$\alpha\Delta I$	$\beta\Delta K$	$\beta\mathbb{E}\left[\Delta K\right]$
2	P/B	0	$\alpha\Delta I$	$\gamma\Delta K$	$\gamma\mathbb{E}\left[\Delta K\right]$
3	0	C	0	$\beta\Delta K$	$\beta\mathbb{E}\left[\Delta K\right] - C_{\Delta t}$
4	(P-C)/B	C	0	$\delta\Delta K$	$\delta\mathbb{E}\left[\Delta K\right] - C_{\Delta t}$

Table 1: Risks and expected return

versus $VaR_{1-\epsilon}$") we obtain the strategy-VaR due to both sources of risk. Consider for instance the first strategy, which has risks $(\alpha\Delta I_i, \beta\Delta K)$. Since both coefficients α and β are positive, the VaR of the strategy is

$$\alpha VaR_{1-\epsilon}\left(\Delta I_i\right) + \beta VaR_{1-\epsilon}\left(\Delta K\right).$$

By applying a similar reasoning for the other strategies, we can compute for each one the *overall VaR*, which we report in Table 2 together with the strategy's net expected return.
It is natural now to represent the trade-offs of the strategies in a familiar way, by associating a point in the plane (Overall-VaR, net expected return) to each strategy. The risk-return preferences of the

Strategy	**(VaR,expected return) combination**
1	$(\alpha\text{VaR}_{1-\epsilon}(\Delta I_i) + \beta\text{VaR}_{1-\epsilon}(\Delta K), \beta\mathbb{E}\left[\Delta K\right])$
2	$(\alpha\text{VaR}_{1-\epsilon}(\Delta I_i) + \gamma\text{VaR}_{1-\epsilon}(\Delta K), \gamma\mathbb{E}[\Delta K])$ if $\gamma > 0$ $(\alpha\text{VaR}_{1-\epsilon}(\Delta I_i) + \gamma\text{VaR}_{\epsilon}(\Delta K), \gamma\mathbb{E}[\Delta K])$ if $\gamma < 0$
3	$(\beta\text{VaR}_{1-\epsilon}(\Delta K), \beta\mathbb{E}\left[\Delta K\right] - C_{\Delta t})$
4	$(\delta\text{VaR}_{1-\epsilon}(\Delta K), \delta\mathbb{E}\left[\Delta K\right] - C_{\Delta t})$ if $\delta > 0$ $(\delta\text{VaR}_{\epsilon}(\Delta K), \delta\mathbb{E}\left[\Delta K\right] - C_{\Delta t})$ if $\delta < 0$

Table 2: Overall VaR for the strategies

fund can be described through a utility function on the plane (Expected Financial Return, Overall VaR):

$$U = f(\mu, \text{VaR}(\Delta I, \Delta k), \eta),$$

where η is a risk aversion coefficient. The best strategy is identified as the one which maximizes the utility function U.
Actually, the fund could reinsure just a part of its liabilities against longevity risk, by choosing $C^* = \eta C, \eta \in [0,1]$. The fund can implement all the linear combinations of the two alternative strategies 1 and 3 or 2 and 4. It is then possible to represent the set of all the possible strategies with a line that goes from 1 to 3 or from 2 to 4. When $n^* = 0$, the set of possible strategies is characterized by a straight line that crosses 1 and 3. When instead condition (1) is met, the set of return maximizing strategies for different values of η is represented by a broken line between 2 and 4. In this case, indeed, there is no liquidity left, since the fund invests all its available resources in the bond. The kink of the line corresponds to the point at which the Delta of the portfolio of assets and liabilities – i.e. short the annuity and long the bond – is null with respect to the financial risk. Given U, the best strategy is identified by the point of the straight line that crosses the highest possible indifference curve. This point identifies the optimal level of reinsurance η^* demanded by

the fund.

5. CONCLUDING REMARKS

The paper explores the risk-return trade-off or a pension fund which can transfer longevity risk and optimally chooses its asset allocation. We measured this trade-off in terms of risk-return combinations and we assessed risk through value-at-risk from both financial and longevity shocks. We succeeded in quantifying the trade-off and we represented it in the plane expected return-VaR. The optimal transfer choices of the fund are located along the corresponding frontier and can be properly identified given its preferences. Our analysis could easily accommodate for the presence of regulatory capital requirements. The objective of the fund will then be to maximize its utility, subject to a solvency constraint, such as the ones descrived in detail in Olivieri and Pitacco (2003).

References

P. Barrieu and H. Loubergé. Reinsurance and securitisation of life insurance risk: the impact of regulatory constraints. *Insurance: Mathematics and Economics*, 52:135–144, 2013.

P. Battocchio, F. Menoncin, and O. Scaillet. Optimal asset allocation for pension funds under mortality risk during the accumulation and decumulation phases. *Annals of Operations Research*, 152:141–165, 2007.

E. Biffis and D. Blake. Securitizing and tranching longevity exposures. *Insurance: Mathematics and Economics*, 46:186–197, 2010.

L. Delong, R. Gerrard, and S. Habermann. Securitizing and tranching longevity exposures. *Insurance: Mathematics and Economics*, 42:107–118, 2008.

D. Hainaut and P. Devolder. Management of a pension fund under mortality and financial risks. *Insurance: Mathematics and Economics*, 41:134–155, 2007.

N. Hari, A. De Waegenere, B. Melenberg, and T. Nijman. Longevity risk in portfolios of pension annuities. *Insurance: Mathematics and Economics*, 42:505–519, 2011.

E. Luciano and L. Regis. Demographic risk transfer: is it worth for annuity providers? *ICER WP*, 11/12, 2012.

E. Luciano, L. Regis, and E. Vigna. Delta-Gamma hedging of mortality and interest rate risk. *Insurance: Mathematics and Economics*, 50(3):402–412, 2012a.

E. Luciano, L. Regis, and E. Vigna. Single and cross-generation natural hedging of longevity and financial risk. *Carlo Alberto Notebooks*, 257, 2012b.

A. Olivieri and E. Pitacco. Solvency requirements for pension annuities. *Journal of Pensions, Economics and Finance*, 2:127–157, 2003.

LINEAR PROGRAMMING MODEL FOR JAPANESE PUBLIC PENSION

Masanori Ozawa† **and Tadashi Uratani**§

†*Faculty of Science and Technology, Keio University, Kawasaki, Japan*
§*Faculty of Science and Engineering , Hosei University, Tokyo, Japan*
Email: `ozawa@ae.keio.ac.jp,uratani@hosei.ac.jp`

We consider a linear programming model to cut benefits to sustain a pension system with longevity and low fertility problem. Under PAYG we find optimal solutions of cutting benefit under given government subsidy. Solutions are obtained very fast for a 100 years planning period and passable to government planning alternative.

1. INTRODUCTION

In most countries, the public pension system is a pay-as-you-go scheme rather than a reserve funding scheme. Longevity with low fertility is one of the worldwide problems in the social security system. Individual pension is a reserve funding scheme but Bayraktar et al. (2007) demonstrated theoretically that reserving money in pension is not optimal for a certain class of individual utility functions. It is a puzzle to the traditional lifecycle hypothesis theory as discussed in Dutta et al. (2000).

We consider optimal strategies to sustain the Japanese public pension system by a linear programming model. We use government data and programs which are available in Japanese Ministry of Health, Labour and Welfare (2012). The Japanese government carries out an actuarial check of the financial status every five years. In the pension reform of 2004, the government has decided to increase premium till 2017 and to fix it afterward. For a reasonable period it is not possible to change the premium, however we need to cut benefits to sustain the system. Considering all the contributors and beneficiaries of a pension, it requires at least a fifty years planning period. In our model, demographic change is assumed to be deterministic in longevity and fertility data given by National Institute of Population and Social Security Research (2012). We simulate the pension system to assess the robustness in optimal solution by economic scenarios of growth rates in wage and return of reserve.

The paper is organized as follows. In section 2 the simple pension model is described by using per capita wage growth for premium and benefit. The growth model of per capita wage and rate of return are assumed to be mean-reversion processes. The budgetary balance of the pension system is formulated as a stochastic process of wage and rate of return of reserve. We formulate an optimalization problem of maximizing the average total benefit of all pensioners with assuring the benefit payment of the target year. In section 3 using Japanese government data we calculate the optimal cut and subsidy to pension in a linear programming model. In the simulation economic scenarios are evaluated by optimal cut and subsidy under constraint to sustain the pension system and budget constraint. Furthermore we simulate different fertility and longevity scenarios and finally we sum up simulation results.

2. MODEL OF PUBLIC PENSION SYSTEM

We define the following pension reserve process; Let $R(t)$ denote the reserve of a pension fund and $r(t)$ the stochastic process of rate of return of the reserve fund. Let $u(a,t)$ be the total premium payment of pension contributors, and $s(b,t)$ the total benefit amount to beneficiaries. The dynamics of the reserve is

$$dR(t) = r(t)R(t)dt + (u(a,t) - s(b,t))dt, \quad R(0) = R_0, \tag{1}$$

where R_0 is the initial reserve.
Let $a(t)$ be the premium rate and $Z_1(t)$ be the total wage of all contributors then $u(a,t) = a(t)Z_1(t)$. Let $\xi_1(t)$ the number of contributors which is estimated by National Institute of Population and Social Security Research (2012). Let $z_1(t)$ denote per capita average wage at t, then $Z_1(t) = z_1(t)\xi_1(t)$.

Suppose the per capita wage is determined by the scenario variable $x(t)$ which is the growth rate of average wage, then

$$z_1(t) = z_1(0)\exp\{\int_0^t x(s)ds\}. \tag{2}$$

The total benefit amount to beneficiaries is also determined by the number of beneficiaries $\xi_2(t)$ and per capita benefit $z_2(t)$. We use the observation of Ministry of Health, Labour and Welfare (2009) that the per capita benefit changes according to $x(t)$,

$$z_2(t) = z_2(0)\exp\{\int_0^t x(s)ds\}. \tag{3}$$

Let $b(t)$ be the cut rate of benefit, then

$$s(b,t) = (1 - b(t))z_2(t)\xi_2(t).$$

The balance of total premium and benefit is

$$u(a,t) - s(a,t) = \psi(t)\exp\{\int_0^t x(s)ds\}.$$

where $\psi(t) = a(t)z_1(0)\xi_1(t) - (1 - b(t))z_2(0)\xi_2(t)$, then (1) becomes

$$dR(t) = r(t)R(t)dt + \psi(t)\exp\{\int_0^t x(s)ds\}dt.$$

We can easily get the solution as follows,

$$R(T) = \exp\{\int_0^T r(s)ds\}\left(R(0) + \int_0^T \psi(t)\exp\{-\int_0^t [r(s) - x(s)]ds\}dt\right). \tag{4}$$

Let $r(t)$ be the rate of return of the pension fund satisfying

$$dr(t) = k_r(\theta_r - r(t))dt + \sigma_r dW_r(t). \tag{5}$$

Let $x(t)$ be rate of change in average salary as

$$dx(t) = k_x(\theta_x - x(t))dt + \sigma_x dW_x(t), \tag{6}$$

where W_r and W_x are Brownian motions satisfying $d < W_r, W_x >= \rho dt$. We assume that $k_v := k_x = k_r$ and $\sigma_v := \sigma_r = \sigma_x$. Then let $v(t) = r(t) - x(t)$ satisfying

$$dv(t) = k_v(\theta_v - v(t))dt + \sigma_v dW_v(t). \tag{7}$$

We consider control strategies of the pension fund (4) by government subsidy $\beta(t)$ and cut rate of benefit $b(t)$ but premium rate $a(t)$ is stipulated in the law and we set a constant value after 2017.

In order to sustain the pension for 100 years under longevity risk and low fertility, it is necessary to pour fund from the government budget unless the premium rate is increased. We set the first constraint for the sum of government subsidy as,

$$E[\int_0^T e^{-\int_0^t r(s)ds}\beta(t)dt] \leq \gamma \tag{8}$$

where we evaluate the minimal required reserve by the binomial model of Uratani and Ozawa (2012). We set the second constraint as a positive reserve at any $\tau \leq T$,

$$E[R_0 + \int_0^\tau e^{-\int_0^t v(s)ds} q_t dt + \int_0^\tau \beta_t e^{-\int_0^t r(s)ds}] \geq 0 \tag{9}$$

The cut rate of benefit is assumed to be less than the rate of population change $b(t) \leq C(t) := 1 - \frac{\xi_1(t)}{\xi_1(0)}$ The other requirement which is called a pension replacement ratio, $\pi_t := \frac{B_t}{I_t}$, is greater than 50%, ($\pi_{2004} = 59.3\%$), where the benefit is measured in household as $B_t = (1 - b(t))z_2(t) + NP$. Standard case in Japanese pension benefit includes house wife National pension benefit, which is denoted as NP. Let I_t denote the average disposable income.
The objective function is the expectation of the total benefit during the planning years,

$$\max_{b(t),\beta(t)} E[\int_0^T (1 - b(t))z_2(0)\xi_2(t)e^{-\int_0^t v(s)ds}]. \tag{10}$$

In order to have economic rationality, we consider the average value of the individual total balance of premium and benefit. The present value of the average pension balance at t is as follows,

$$Y(t) = \int_{t+T_p}^{t+T_d} (1 - b(u))z_2(u)e^{-\int_t^u r_s ds}du - \int_t^{t+T_p} az_1(u)e^{-\int_t^u r_s ds}du, \tag{11}$$

where t is a start time of premium, T_d is the life expectancy, and T_p is premium payment period.

3. LINEAR PROGRAMMING MODEL

We discretize the time span from 2013 to 2112 in the above continuous modeling (10), (8), (9). The objective function is to minimize the present value of total cut benefit,

$$\min_{b(t_i),\beta(t_i)} \sum_{i=0}^{n} b(t_i)\xi_2(t_i)z_2(0)e^{f_v(t_i)}$$

$$\sum_{i=0}^{n} \beta(t_i)\exp(-\mu_r(t_i)+\sigma_r^2(t_i)/2) \leq \gamma$$

$$R(0)+\sum_{i=0}^{k} \beta(t_i)e^{f_r(t_i)} \geq -\sum_{i=0}^{k} q(t_i)e^{f_v(t_i)}$$

$$b(t_k) \leq b(t_{k+1}),\ 0 \leq b(t_k) \leq C(t_k),\ \text{ for } k=0,\ldots,n.$$

where we set for $h=\{v,r\}$, $f_h(t_i) := -\mu_h(t_i)+\sigma_h^2(t_i)/2$, $\mu_h(t) := \theta_h t + \frac{h(0)-\theta_h}{k}(1-e^{-k_h t})$, $\sigma_h^2(t) := \frac{\sigma_h^2}{k^2}\int_0^t (1-e^{-k_h(s-t)})^2 ds$.

3.1. Economic scenarios

For simulation scenarios of 100 years, we set annual average changes rates θ_r and θ_x as following Table 1. The volatility is assumed to be same value $\sigma = 0.01$ and mean-reversion $k = 0.1$. High case is high inflation and Middle case is the government inflation target and Low case is deflation. Present values of government subsidy for 100-years γ in constraint (8) are 4 cases from 400 to 550

%	inflation	nominal salary θ_r	rate of return θ_x
High	3	4.5	6.1
Middle	1	2.5	4.1
Low	-0.5	-0.5	1.1

Table 1: Economic scenarios in OU processes in (6) and (7)

trillion yen, which is calculated by Uratani and Ozawa (2012).
In Figure 1 each column represents respectively Low, Middle, High economic scenario. The first row depicts the cumulative cut ratio for 100 years. The more government subsidy is spent, the less cut ratio is required.
Second row is nominal subsidy which is the same for most years except beginning and ending years. The required subsidies are almost same in different economic cases.
The third and fourth rows are simulation results in respect to various values of rate of return. We assume that discount rate is equal to the rate of return. Therefore in the third row, high rate of return decreases the object function value. On the contrary low rate of return cannot sustain the pension system.

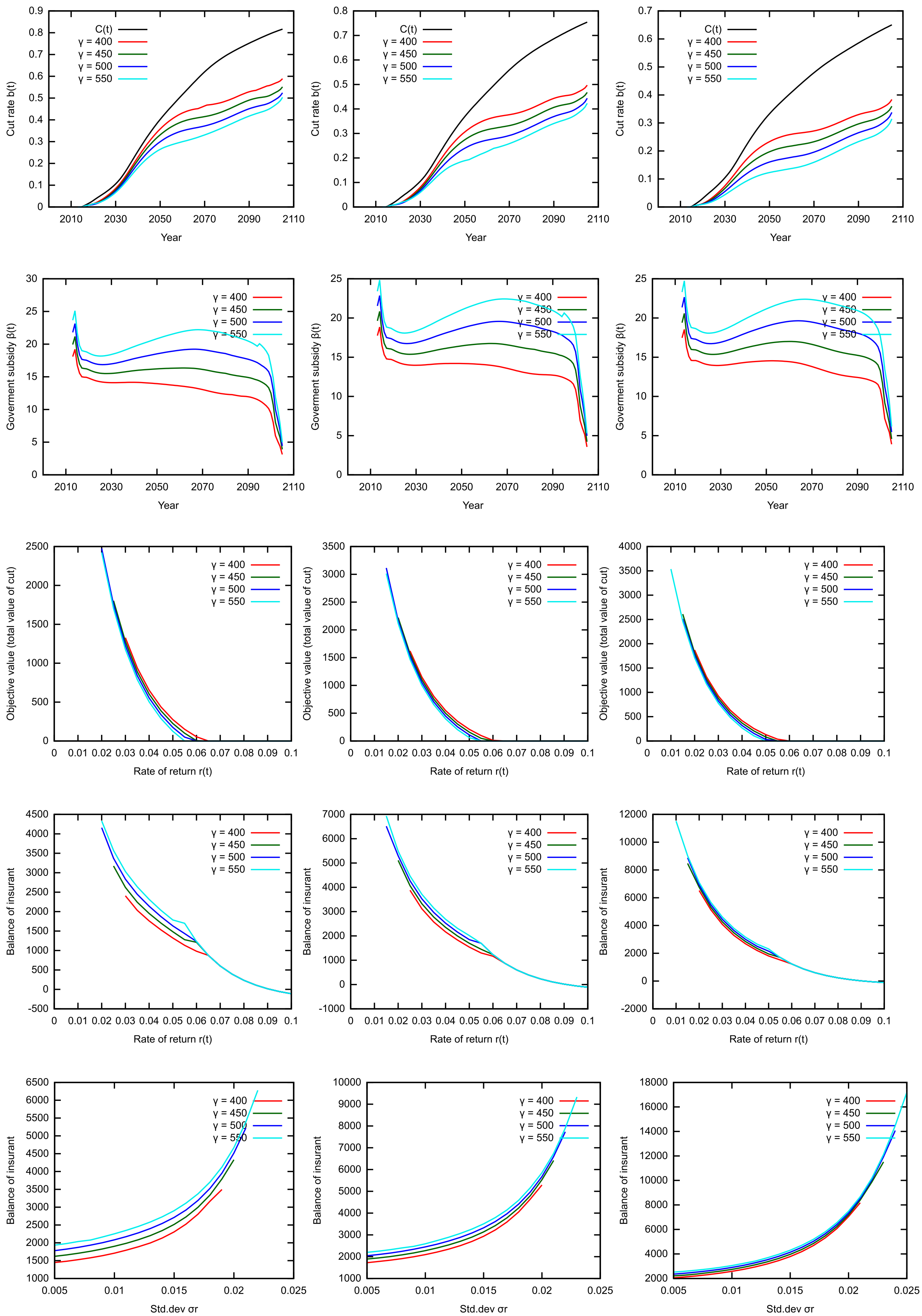


Figure 1: Economic simulation by linear programming

The fourth row depicts the present value of pension participant balance $Y(t)$ of (11). It shows that sustainability implies positive balance of insurants.
The last row of Figure 1 shows that the volatility does not affect the balance of insurants.

3.2. Concluding remarks

Concerning the cut rate $b(t)$, we conclude following points; (i) It is necessary to have a cut in benefit in order to sustain the pension system for 100 years. (ii) It is reasonable to set the cut ratio below the decreasing rate of population. (iii) Increasing government subsidy decreases optimal cut ratio.
Concerning the government subsidy $\beta(t)$ and total cut amount of benefit, we conclude the following points under the assumption that it is not allowed to increase premium after 2017; (i) High economy case: maximum subsidy of 40 trillion and cut amount of 10 trillion yen, (ii) Intermediate economy case: maximum subsidy of 10 trillion and cut amount of 14 trillion yen, (iii) Low economy case: Decreasing from 10 trillion subsidy and maximum cut amount of 14 trillion yen.

From simulation of demographic change the effect is significant as economic change but the effect is very similar. The public pension has economical rationality for future average generation.

References

E.M. Bayraktar, K. Moore, and V. Young. Minimizing the probability of lifetime ruin under borrowing constraints. *Insurance: Mathematics and Economics*, 41:196–221, 2007.

J. Dutta, S. Kapur, and J.M. Orszag. A portfolio approach to the optimal funding of pensions. *Economics Letters*, 69(2):201–206, 2000.

Japanese Ministry of Health, Labour and Welfare, 2012. URL `http://www.mhlw.go.jp`.

Ministry of Health, Labour and Welfare. The 2009 actuarial valuation of the employees' pension insurance and the national pension, 2009. (in Japanese).

National Institute of Population and Social Security Research. Annual report, 1 2012. URL `http://www.ipss.go.jp`.

T. Uratani and M. Ozawa. A simple model of japanese public pension and the risk management by option hedging strategy. *International Journal of Real Options and Strategy*, 2012. (in preparation).

SOME SIMPLE AND CLASSICAL APPROXIMATIONS TO RUIN PROBABILITIES APPLIED TO THE PERTURBED MODEL

Miguel J.M. Seixas† and Alfredo D. Egídio dos Reis§[1]

†*AXA MedLa, Camino Fuente de La Mora 1, 28050 Madrid, Spain*
§*ISEG & CEMAPRE, Technical University of Lisbon, Rua do Quelhas 6, 1200-781 Lisboa, Portugal*
Email: seixas.miguel@gmail.com, alfredo@iseg.utl.pt

We study approximations of the ultimate ruin probabilities under an extension to the classical Cramér-Lundberg risk model by adding a diffusion component. For the approximations, we adapt some simple, practical and well known methods that are used for the classical model. Under this approach, and for some cases, we are able to separate and to compute the ruin probability, either exclusively due to the oscillation, or due to a claim.

1. INTRODUCTION

We start by presenting **the model and the probability of ruin**. We study the **perturbed surplus process** as introduced by Dufresne and Gerber (1991) and defined for time t as:

$$V(t) = U(t) + \sigma W(t), \quad U(t) = u + ct - S(t), \quad t \geq 0\,,$$

where $U(t)$ defines the classical surplus process, c is the premium rate per unit time, $u = V(0) = U(0)$ is the initial surplus, $S(t) = \sum_{i=0}^{N(t)} X_i$, $X_0 \equiv 0$, are the aggregate claims up to time t, $N(t)$ is the number of claims received up to time t, X_i is the i-th individual claim, $W(t)$ is the diffusion component and σ^2 is the variance parameter. $\{W(t), t \geq 0\}$ is a standard Wiener process, $\{N(t), t \geq 0\}$ is a Poisson process with parameter λ and $\{X_i\}_{i=1}^{\infty}$ is a sequence of $i.i.d.$ random variables, independent from $\{N(t)\}$ with common distribution function $P(.)$ with $P(0) = 0$. The corresponding density function is denoted as $p(.)$. Denote by $p_k = E[X^k]$. The existence of p_1 is basic and essential, only in some of our methods the existence of higher moments is needed. We assume that $\{S(t)\}$ and $\{W(t)\}$ are independent. We also assume that $c = (1 + \theta)\lambda p_1$, where $\theta > 0$ is the premium loading coefficient.

[1]Support from FCT-Fundação para a Ciência e a Tecnologia (Programme PEst-OE/EGE/UI0491/2011) is gratefully acknowledged.

The diffusion component introduces an additional uncertainty into the classical model, so that if ruin occurs it may be caused either from a claim or by an (unfavorable) oscillation of the diffusion process. Let T be the time to ruin such that $T = \inf\{t : t \geq 0 \text{ and } V(t) \leq 0\}$, $T = \infty$ if $V(t) > 0$, $\forall t$. The ultimate ruin probability is given by

$$\psi(u) = \Pr(T < \infty | V(0) = u) = \psi_d(u) + \psi_s(u),$$

where $\psi_s(u)$ and $\psi_d(u)$ are the ruin probabilities due to a claim and to oscillation, respectively. The survival probability is $\delta(u) = 1 - \psi(u)$. We have that $\psi_d(0) = \psi(0) = 1$. Furthermore, $\delta(u)$, $\psi_s(u)$ and $\psi_d(u)$ follow defective renewal equations, respectively, for $u \geq 0$:

$$\begin{aligned}
\psi_s(u) &= (1-q)\left[H_1(u) - H_1 * H_2(u)\right] + (1-q)\int_0^u \psi_s(u-x) h_1 * h_2(x)dx\,, \\
\psi_d(u) &= 1 - H_1(u) + (1-q)\int_0^u \psi_d(u-x) h_1 * h_2(x)dx\,, \\
\delta(u) &= qH_1(u) + (1-q)\int_0^u \delta(u-x) h_1 * h_2(x)dx\,, \qquad (1)
\end{aligned}$$

with $q = 1 - \lambda p_1/c$, h_1 and $h_2(.)$ given by ($H_1(.)$ and $H_2(.)$ are the corresponding d.f.):

$$\begin{aligned}
h_1(x) &= \zeta e^{-\zeta x}, x > 0, \;\; \zeta = 2c/\sigma^2\,, \\
h_2(x) &= p_1^{-1}\left[1 - P(x)\right], x > 0.
\end{aligned}$$

We further introduce the **maximal aggregate loss** defined as $L = \max\{t \geq 0, L(t) = u - V(t)\}$. It can be decomposed as

$$\begin{aligned}
L &= L_0^{(1)} + \sum_{i=1}^{M}\left(L_i^{(1)} + L_i^{(2)}\right), \qquad (2) \\
L_i^{(1)} &= \max\{L(t), t < t_{i+1}\} - L(t_i), i = 0, 1, \ldots, M, \qquad (3) \\
L_i^{(2)} &= L(t_i) - L(t_{i-1}) - L_{i-1}^{(1)}, i = 1, \ldots, M, \qquad (4)
\end{aligned}$$

where M is the number of records of $L(t)$ that are caused by a claim, $L_i^{(1)}$ and $L_i^{(2)}$ are the *record highs* due to oscillation and a claim. $\{L_i^{(1)}\}_{i=0}^{\infty}$ and $\{L_i^{(2)}\}_{i=1}^{\infty}$ are independent sequences of $i.i.d$ random variables, with common d.f. $H_1(.)$, and $H_2(.)$, respectively. Also, $\delta(x) = \Pr\{L \leq x\}$ is a compound geometric d.f. and existing moments can be found easily.

We consider different **approximation methods that are adapted from the pure classical model.** We start with the method by **De Vylder (1978)**, that relies on the use of the exact ruin formula when the individual claim amount is exponential. We follow with a method by **Dufresne and Gerber (1989)** that produces upper and lower limits for the ruin probability and it is very useful to test the accuracy of the other methods presented, often simpler, for the cases where we do not have exact figures for the ruin probability. These two methods were already tried by Silva (2006), who presented no figures. After, we adapt an approximation known as **Beekman and Bowers'**, presented in Beekman (1969). It uses an appropriate gamma distribution in the defective renewal equation for $\delta(u)$. Jacinto (2008) also did some work on the previous methods. We further work two other models, **Tijms'** and the **Fourier transform** methods. The former was originally presented in the context of queueing theory by Tijms (1994), the latter is an adaptation of the work by Lima et al. (2002).

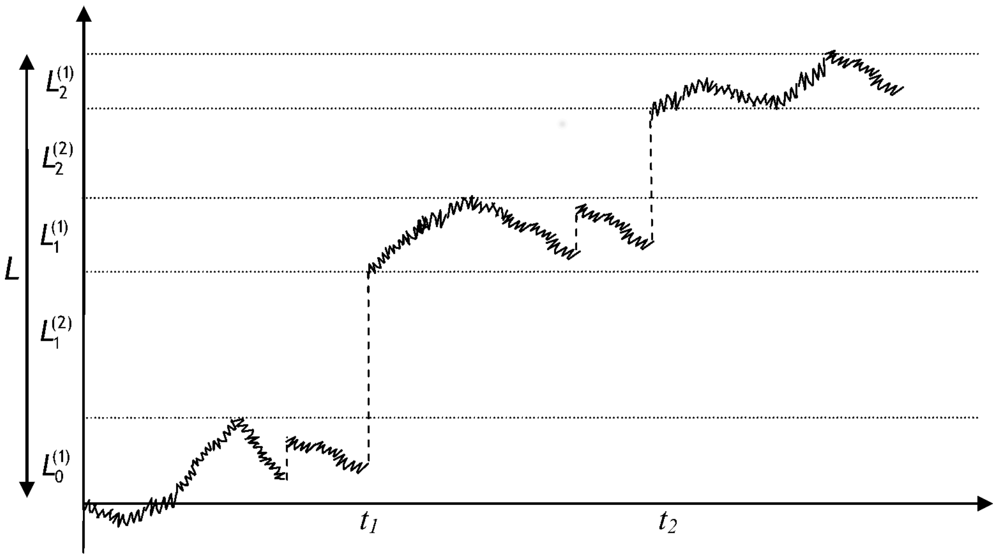


Figure 1: Decomposition of the maximal aggregate loss.

2. APPROXIMATIONS IN THE PERTURBED MODEL

We follow the order presented in the previous section and start with the **De Vylder's approximation**. Following De Vylder (1978), the original process, $V(t)$, is replaced by another process

$$V^*(t) = u + c^*t - S^*(t) + \sigma^* W(t),$$

where the individual claims follow an $exponential(\beta)$, and parameters β, c^*, λ^* and $\sigma^{*\,2}$ are calculated so that the existing lower four moments of $V(t)$ and $V^*(t)$ match:

$$\beta = 4\tfrac{p_3}{p_4}; \qquad \lambda^* = 32\lambda\tfrac{p_3^4}{3p_4^3}; \qquad c^* = 8\lambda\tfrac{p_3^3}{3p_4^2} + c - \lambda p_1; \qquad \sigma^{*\,2} = \sigma^2 + \lambda p_2 - 4\lambda\tfrac{p_3^2}{3p_4}.$$

Then, we use the exact ruin probability formula from Dufresne and Gerber (1991), so that approximation comes

$$\psi_{DV}(u) = C_1 e^{-r_1} + C_2 e^{-r_2}, \qquad C_1 = \tfrac{r_1-\beta}{\beta}\tfrac{r_2}{r_1-r_2}, \qquad C_2 = \tfrac{r_2-\beta}{\beta}\tfrac{r_1}{r_2-r_1}\ ,$$

where r_1 and r_2 are the solutions of equation, $r\sigma^{*2}/2 + \lambda^*/(\beta - r) = c^*$. Furthermore, we can obtain approximations for the decomposed probabilities $\psi_s(u)$ and $\psi_d(u)$, simply using the exact result for the case where the individual losses are exponential.

The second method is called the **Dufresne & Gerber's upper and lower bounds**. It is based on getting appropriate discrete distributions to replace on the convolution formula for the survival probability, Formula (7) in Dufresne and Gerber (1989). For the perturbed model, we use a similar method, now based on Formula (5.8) of Dufresne and Gerber (1991). Discrete random variables are defined followed by bounds computation for the ruin probabilities [see Sections 2.3 and 2.4 of Dufresne and Gerber (1989)]. We have

$$L^j = L_0^{j,(1)} + \sum_{i=1}^{M}\left(L_i^{j,(1)} + L_i^{j,(2)}\right),$$

with $L^j = L_0^{j,(1)}$ if $M = 0$ and $j = l, u$, $L_i^{l,(k)} = \vartheta\left[L_i^{(k)}/\vartheta\right]$, $L_i^{u,(k)} = \vartheta\left[(L_i^{(k)}+\vartheta)/\vartheta\right]$ for $\{k=1, i=0,...,M\}$, $\{k=2, i=1,...,M\}$, $\vartheta\epsilon(0,1)$ and $[x]$ is the integer part of x. Each summand of L, $L_i^{(k)}$, in (2), is correspondingly approximated by both the next lower and higher multiples of ϑ. We have then,

$$L^l \leq L \leq L^u \Rightarrow \Pr(L^l \geq v) \leq \psi(v) \leq \Pr(L^u \geq v).$$

We need the p.f. of the discrete r.v.'s $L_i^{l,(1)}, L_i^{l,(2)}, L_i^{u,(1)}$ and $L_i^{u,(2)}$, they are given by, respectively,

$$\begin{aligned} h_{n,k}^l &= \Pr\left(L_i^{l,(n)} = k\vartheta\right) = H_n(k\vartheta+\vartheta) - H_n(k\vartheta),\ n=1,2;\ k=0,1,...,\\ h_{n,k}^u &= \Pr\left(L_i^{u,(n)} = k\vartheta\right) = H_n(k\vartheta+\vartheta) - H_n(k\vartheta),\ n=1,2;\ k=0,1,... \end{aligned}$$

The following probability functions of L^l and L^u, f_k^l and f_k^u, can be computed using Panjer's recursion (for the compound geometric distribution)

$$f_k^j = \Pr\left(L^j = k\vartheta\right), k=0,1,... \text{ for } j=l,u\,.$$

We arrive to the following bounds for $\psi(.)$, where

$$1 - \sum_{k=0}^{m-1} f_k^l \leq \psi(m\vartheta) \leq 1 - \sum_{k=0}^{m} f_k^u,\quad m=0,1,...,v/\vartheta,\quad v=0,1,...$$

We consider now the **Beekman-Bowers' approximation**. We replace $\delta * h_2(.)$ in the renewal equation (1), $\delta(u) = qH_1(u) + (1-q)h_1 * \delta * h_2(u)$, by a d.f. of a gamma$(\alpha,\beta)$, denoted as $H_3(u)$. We arrive to the approximation

$$\delta_{BB}(u) = qH_1(u) + (1-q)h_1 * H_3(u),$$

Parameters α and β are got by equating the moments of $\delta_{BB}(u)$ with those of $\delta(u)$, respectively.

We address now **Tijms' approximation**. This method relies on the existence of the adjustment coefficient and an asymptotic formulae for $\psi(u)$, $\psi_d(u)$, and $\psi_s(u)$. Similarly to Tijms (1994) we consider the approximating expression

$$\psi_T(u) = Ce^{-Ru} + Ae^{-Su}, u \geq 0\,,$$

where A is chosen such that $\psi(0) = \psi_T(0)$. As $\psi(0) = 1$, then $A = (1-C)$. As $\psi(.)$ is the survival function of L, S is chosen in order that $\int_0^\infty \psi_T(u)du = E[L]$. Hence,

$$E[L] = \frac{C}{R} + \frac{(1-C)}{S} \Leftrightarrow S = \frac{R\,(1-C)}{RE[L]-C}\,.$$

The method we work and simply name as **Fourier** transform is not quite an approximation method but an exact formula that allows to compute numerically the ruin probability. This method uses the Fourier transform,

$$\phi_{f(x)}(s) = \int_0^{+\infty} e^{isx} f(x)dx = \underbrace{\int_0^{+\infty} \cos(sx)f(x)dx}_{\phi^r_{f(x)}(s)} + i\underbrace{\int_0^{+\infty} \sin(sx)f(x)dx,}_{\phi^c_{f(x)}(s)}$$

u	$\psi(u)\ (I)$	$\psi_{BB}(u)(II)$	$(I)/(II)$	$\psi_T(u)\ (III)$	$(I)/(III)$
1	0.40470	0.39819	1.01633	0.40470	1.00000
3	0.16674	0.17096	0.97529	0.16674	1.00000
5	0.06938	0.07089	0.97866	0.06938	1.00000
10	0.00775	0.00731	1.06010	0.00775	1.00000
15	0.00087	0.00072	1.19580	0.00087	1.00000

Table 1: Exact figures, Beekman-Bowers' and Tijms' approximations for $Exponential(1)$

so that for $F'(x) = f(x)$ we have

$$F(x) = F(0) + \frac{2}{\pi}\int_0^\infty \frac{\sin(xs)}{s}\phi^r_{f(x)}(s)ds\,. \tag{5}$$

From the integro-differential equation for $\psi(u)$ we get

$$\psi'(u) = -qh_1(u) + (1-q)\int_0^u \psi'(u-x)h_1 * h_2(x)dx\,,$$

and the transform can be written as

$$\phi_{\psi'(u)}(s) = \frac{A+iB}{C-iD} = \frac{AC-BD+i(BC+AD)}{C^2+D^2}\,,$$

with $A = -q\phi^r_{h_1(u)}(s)$, $B = -q\phi^c_{h_1(u)}$, $C = 1 - J(1-q)/sp_1$ and $D = I(1-q)/sp_1$. I and J depend only on the real and the complex part of $\phi_{h_1(u)}(s)$ and $\phi_{p(u)}(s)$:

$$\begin{aligned} I &= \phi^r_{h_1(u)}(s) - \phi^r_{h_1(u)}(s)\phi^r_{p(u)}(s) + \phi^c_{h_1(u)}(s)\phi^c_{p(u)}(s) \\ J &= \phi^r_{h_1(u)}(s)\phi^c_{p(u)}(s) - \phi^c_{h_1(u)}(s) + \phi^c_{h_1(u)}(s)\phi^r_{p(u)}(s)\,. \end{aligned}$$

Approximation $\psi_F(u)$ is then got computing numerically the inversion integral (5). Similar results can be derived for $\psi_{d,F}(u)$ and $\psi_{s,F}(u)$ (the index $_F$ refers to this method).

3. NUMERICAL ILLUSTRATIONS

For the sake of illustration we show numerical results for three examples: when single amounts follow Exponential(1), Gamma$(2,2)$ or Pareto$(5,4)$ distributions (all with mean equal to one). The other parameters are: $c = 2$, $\lambda = 1$, $\sigma = 1$ and $\vartheta = 0.01$. Tables 1 and 2 show the results concerning the first example (De Vylder's method is exact in this case). Table 3 provides results for the Gamma$(2,2)$ case. Table 4 shows results for the Pareto$(5,4)$ case and all other methods except Tijms' one, as it doesn't apply. Table 5 shows the percentage of ruin due to oscillation for the worked cases.

u	$\psi(u)$ (I)	$\psi_F(u)$ (II)	$(I)/(II)$	$\psi_d(u)$ (III)	$\psi_{d,F}(u)$ (IV)	$(III)/(IV)$
1	0.40470	0.40470	1.00000	0.09688	0.09688	0.99999
3	0.16674	0.16674	1.00000	0.03655	0.03655	1.00000
5	0.06938	0.06937	1.00000	0.01521	0.01521	1.00000
10	0.00775	0.00775	1.00000	0.00170	0.00170	1.00000
15	0.00087	0.00087	1.00000	0.00019	0.00019	1.00002

Table 2: Exact figures and Fourier method for $Exponential(1)$

u	Lower Bound	$\psi_{DV}(u)$	$\psi_{BB}(u)$	$\psi_T(u)$	$\psi_F(u)$	Upper Bound
1	0.38643	0.39199	0.38231	0.39394	0.38867	0.39092
3	0.12024	0.12155	0.12660	0.12198	0.12196	0.12369
5	0.03696	0.03775	0.03825	0.03780	0.03780	0.03865
10	0.00194	0.00203	0.00167	0.00202	0.00202	0.00211
15	0.00010	0.00011	0.00007	0.00011	0.00011	0.00012

Table 3: Dufresne-Gerber's Bounds, De Vylder's, Beekman-Bowers', Tijms' & Fourier, $Gamma$.

u	Lower Bound	$\psi_{DV}(u)$	$\psi_{BB}(u)$	$\psi_F(u)$	Upper Bound
1	0.40867	0.45521	0.38282	0.41036	0.41206
3	0.19577	0.15464	0.20096	0.19707	0.19838
5	0.10339	0.08437	0.11286	0.10423	0.10509
10	0.02511	0.02879	0.02824	0.02537	0.02564
15	0.00727	0.01032	0.00730	0.00736	0.00744

Table 4: Dufresne-Gerber's Bounds, De Vylder's, Beekman-Bowers' & Fourier; $Pareto(5,4)$

	Exponential	Gamma			Pareto		
u	$\psi_d(u)/\psi(u)$	$\psi_{d,F}(u)$	$\psi_{s,F}(u)$	$\psi_{d,F}(u)/\psi_F(u)$	$\psi_{d,F}(u)$	$\psi_{s,F}(u)$	$\psi_{d,F}(u)/\psi_F(u)$
1	24%	0.11221	0.27647	29%	0.09042	0.31994	22%
3	22%	0.03570	0.08626	29%	0.03296	0.16411	17%
5	22%	0.01107	0.02673	29%	0.01590	0.08833	15%
10	22%	0.00059	0.00143	29%	0.00334	0.02203	13%
15	22%	0,00003	0,00008	29%	0.00085	0.00650	12%

Table 5: Weight of $\psi_d(u)$ for $Exponential(1)$, $Gamma(2,2)$ and $Pareto(5,4)$

4. CONCLUDING REMARKS

We underline the poor fit of the Beekman-Bowers' method no matter the examples we consider. The methods of De Vylder and Tijms appear capable of producing good results for light tail claims size distributions. In all cases Dufresne & Gerber's bounds method produces good approximations. The same observation holds true for the Fourier transform method which produces numerically exact figures. A final remark deals with the contribution of the oscillation component which plays a substantial role in the ruin probability, especially in the case that the claim size distribution is light tailed. We have chosen a volatility equal to one (equal to the mean claim size) in all examples. A deeper study can be performed choosing different values. For more details on the work please see Seixas (2012).

References

J. Beekman. A ruin function approximation. *Trans. Soc. Actuaries*, 21:41–48 and 275–279, 1969.

F. De Vylder. A practical solution to the problem of ultimate ruin probability. *Scandinavian Actuarial Journal*, 1978(2):114–119, 1978.

F. Dufresne and H. U. Gerber. Three methods to calculate the probability of ruin. *ASTIN Bulletin*, 19(1):71–90, 1989.

F. Dufresne and H. U. Gerber. Risk theory for the compound poisson process that is perturbed by diffusion. *Insurance: Mathematics and Economics*, 10:51–59, 1991.

A. C. F. S. Jacinto. Aproximações à probabilidade de ruína no modelo perturbado. Master's thesis, ISEG Technical University of Lisbon, 2008.

R. Kaas, M. Goovaerts, J. Dhaene, and M. Denuit. *Modern Actuarial Risk Theory: Using R*. Springer, Chichester, 2nd edition edition, 2008.

F. D. P. Lima, J. M. A. Garcia, and A. D. Egídio dos Reis. Fourier/laplace transforms and ruin probabilities. *ASTIN Bulletin*, 32(1):91–105, 2002.

M. J. M. Seixas. Some simple and classical approximations to ruin probabilities applied to the perturbed model. Master's thesis, ISEG Technical University of Lisbon, 2012.

M. D. V. Silva. Um processo de risco perturbado: Aproximações numéricas à probabilidade de ruína. Master's thesis, Faculty of Science University of Oporto, 2006.

H. Tijms. *Stochastic Models - An Algorithmic Approach.* John Wiley & Sons, Chichester, 1994.

BAYESIAN DIVIDEND MAXIMIZATION: A JUMP DIFFUSION MODEL

Michaela Szölgyenyi[1]

Institute of Financial Mathematics, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria
Email: `michaela.szoelgyenyi@jku.at`

Abstract

In this paper we study the valuation problem of an insurance company. We seek to maximize the discounted future dividend payments until the time of ruin. The surplus is modelled as a jump-diffusion process, where we assume to only have incomplete information. Therefore, we apply filtering theory to overcome uncertainty. Then we derive the associated Hamilton-Jacobi-Bellman equation. Finally, we study the problem numerically.

1. INTRODUCTION

De Finetti (1957) proposed the expected discounted future dividend payments as a valuation principle for an insurance portfolio. Standard references for diffusion models with complete observations are Shreve et al. (1984), Jeanblanc-Piqué and Shiryaev (1995), Radner and Shepp (1996), and Asmussen and Taksar (1997). For a jump-diffusion model with complete observations, see Belhaj (2010). For surveys about dividend optimization problems in various models we refer to Albrecher and Thonhauser (2009) and Avanzi (2009). However, all these papers treat the dividend maximization problem in full information setups. In Leobacher et al. (2013) we deal with the dividend maximization problem in a so-called Bayesian framework, i.e., the drift is modelled as an unobservable random variable expressing the insurer's uncertainty about the profitability of the portfolio.

In this note we extend the model proposed in Leobacher et al. (2013) by adding a jump component, and present a numerical study of the problem.
Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space and let the augmentation of the filtration generated by the later defined processes X, Z, and S, $\mathcal{F}^{X,Z,S}$ be our observation filtration.

[1]The author is supported by the Austrian Science Fund (FWF) Project P21943. The author would like to thank Gunther Leobacher, Stefan Thonhauser, and Jörn Sass for their helpful advice.

We model the surplus $X = (X_t)_{t\geq 0}$ of an insurance company by

$$X_t = x + \int_0^t (\theta - u_s)\, ds + \sigma B_t - S_t = Z_t - L_t - S_t\,, \qquad x > 0\,.$$

The ingredients of the model are the drift θ, the volatility σ, the Brownian motion $B = (B_t)_{t\geq 0}$, and the control $u = (u_t)_{t\geq 0}$.
The drift $\theta \in \{\theta_1, \theta_2\}$ with $0 < \theta_1 < \theta_2$ is constant, unobservable under $\mathcal{F}^{X,Z,S}$, but has given initial distribution $q := \mathbb{P}(\theta = \theta_1) = 1 - \mathbb{P}(\theta = \theta_2)$. σ is constant and observable, and $u_t \in [0, K]$ is the density at time t of the cumulated dividend process $L = (L_t)_{t\geq 0}$. $Z = (Z_t)_{t\geq 0}$ is the uncontrolled process. $S = (S_t)_{t\geq 0}$ is a compound Poisson process, i.e., $S_t = \sum_{i=1}^{N_t} D_i$, where $N = (N_t)_{t\geq 0}$ is a Poisson process with observable intensity ν and $D_i \sim \mathrm{Exp}(\lambda)$, so its (completely monotonic) density is given by $f_D(x) = \lambda e^{-\lambda x}$ with observable λ.

For applying the dynamic programming approach from optimal stochastic control, we have to apply filtering theory to overcome uncertainty. Our aim is to replace θ by an observable estimator $(\theta_t)_{t\geq 0}$ with

$$\theta_t = \mathbb{E}(\theta | Z_t \in [\bar{z}, \bar{z} + d\bar{z}])\,.$$

In Leobacher et al. (2013) we derived a filter for the problem without jumps, i.e., $S \equiv 0$. From the structure of our model the jumps are directly observable and the drift needs to be filtered from the continuous part only. Therefore, as in Leobacher et al. (2013), by using Bayes' rule we get

$$\mathbb{P}(\theta = \theta_1 | Z_t \in [\bar{z}, \bar{z} + d\bar{z}]) = \frac{1}{1 + \frac{1-q}{q} \exp\left(\frac{(\theta_2 - \theta_1)(\bar{z} - z - \frac{1}{2}(\theta_1 + \theta_2)t)}{\sigma^2}\right)}\,.$$

Thus, using Itô's formula we arrive at the following system:

$$X_t = x + \int_0^t (\theta_s - u_s)\, ds + \sigma W_t - S_t\,, \tag{1}$$

$$\theta_t = \vartheta + \int_0^t (\theta_s - \theta_1)(\theta_2 - \theta_s)\, dW_s\,, \tag{2}$$

where $(\theta_t)_{t\geq 0}$ is the estimator for the drift. One can show that $W = (W_t)_{t\geq 0}$ is a Brownian motion w.r.t. $\mathcal{F}^{X,Z,S}$, which replaces B (cf. (Liptser and Shiryaev 1977, Theorem 9.1)).
Considering (X_t, θ_t) we face full information. However, the price we have to pay is an extra dimension.

2. STOCHASTIC OPTIMIZATION

Now we can define the stochastic optimization problem and heuristically derive the Hamilton-Jacobi-Bellman equation.

Since $(X_t, \theta_t)_{t\geq 0}$ is a Markov process, it is natural to consider Markov controls of the form $u_t = u(X_{t-}, \theta_t)$. Our aim is to find the optimal value function

$$V(x, \vartheta) = \sup_{u\in A} J^{(u)} = \sup_{u\in A} \mathbb{E}_{x,\vartheta} \left(\int_0^\tau e^{-\delta t} u_t \, dt \right)$$

and the optimal control law $u \in A$, where $\tau := \inf\{t \geq 0 | X_t \leq 0\}$, i.e., the stopping time of ruin. A is the set of admissible controls, which imposes technical conditions such that the control process exists and the value function is well-defined. $\mathbb{E}_{x,\vartheta}$ denotes the expectation given the initial values $X_0 = x$, $\theta_0 = \vartheta$.

Let $\eta > 0$ be an arbitrary stopping time. Heuristically applying (Protter 2004, Chapter II, Theorem 32) we can write

$$\begin{aligned} V(x, \vartheta) &= \sup_{u\in A} \mathbb{E}_{x,\vartheta} \left[\int_0^{\eta\wedge\tau} e^{-\delta t} u_t \, dt + e^{-\delta(\eta\wedge\tau)} V(X_{\eta\wedge\tau}, \theta_{\eta\wedge\tau}) \right] \\ &= \sup_{u\in A} \mathbb{E}_{x,\vartheta} \left[\int_0^{\eta\wedge\tau} e^{-\delta t} u_t \, dt + e^{-\delta(\eta\wedge\tau)} \left(V(x, \vartheta) + \int_0^{\eta\wedge\tau} \mathcal{L}V(X_t, \theta_t) \, dt \right.\right. \\ &\qquad \left.\left. - \int_0^{\eta\wedge\tau} u_t V_x(X_t, \theta_t) \, dt - \int_0^{\eta\wedge\tau} \Delta V(X_t, \theta_t) \, dt \right) \right] , \end{aligned} \tag{3}$$

with

$$\mathcal{L}V = \vartheta V_x + \frac{\sigma^2}{2} V_{xx} + \frac{1}{\sigma^2} (\theta_2 - \vartheta)^2 (\vartheta - \theta_1)^2 V_{\vartheta\vartheta} + (\theta_2 - \vartheta)(\vartheta - \theta_1) V_{x\vartheta} \, .$$

Using that S is a compound Poisson process, we obtain

$$\begin{aligned} -\mathbb{E}_{x,\vartheta} \left[\int_0^{\eta\wedge\tau} \Delta V(X_t, \theta_t) \, dt \right] &= \mathbb{E}_{x,\vartheta} \left[\sum_{0<t\leq\eta\wedge\tau} (V(X_t, \theta_t) - V(X_{t-}, \theta_t)) \right] \\ &= \nu(\eta \wedge \tau) \int_0^\infty \left((V(x-y, \vartheta) - V(x, \vartheta)) \lambda e^{-\lambda y} \right) \, dy \, , \end{aligned}$$

where $V(x-y, \vartheta) = 0$ for $y \geq x$. Dividing (3) by $\eta \wedge \tau$ and letting $\eta \to 0$ we arrive at

$$-(\delta + \nu) V(x, \vartheta) + \mathcal{L}V(x, \vartheta) + \sup_{u\in[0,K]} (1 - V_x(x, \vartheta)) u + \nu\lambda \int_0^x V(x-y, \vartheta) e^{-\lambda y} \, dy = 0 \, . \tag{4}$$

(4) is the HJB equation corresponding to the optimization problem where the underlying surplus process has jumps. The natural boundary conditions are given by

$$V(0, \vartheta) = 0 \, , \qquad V(B, \vartheta) = \frac{K}{\delta} \text{ for } B \to \infty \, .$$

$V(x, \theta_1)$ and $V(x, \theta_2)$ are obtained by solving the corresponding one-dimensional problems.

Calculating the supremum in the HJB equation yields a maximum of the form

$$\sup_{u\in[0,K]} (u(1 - v_x)) = \begin{cases} K, & v_x \leq 1 \\ 0, & v_x > 1 \, . \end{cases}$$

3. NUMERICAL SOLUTION

In this section we illustrate a method for solving the stochastic optimization problem numerically. The convergence results from (Fleming and Soner 2006, Chapter IX) imply that for the jump-free case we can compute a value function corresponding to an ε-optimal dividend policy using a finite difference method. Basically, we can use a similar numerical method as proposed in Leobacher et al. (2013) for the jump-free case. However, since here we have an IPDE instead of a PDE, we use a simple quadrature rule for numerical integration. Since the quadrature rule is simply added to the discretized problem from the jump-free case and since it converges to the integral part of the HJB equation, convergence will be preserved. Of course, for making these statements rigorous one needs to prove that V is the unique viscosity solution of the HJB equation in advance to the numerical treatment, but this theoretical treatment is beyond the scope of this paper.

We follow the following procedure:

- We start with a simple (threshold) strategy:

$$u^{(0)}(x,\vartheta) = K\,1_{\{x \geq b(\vartheta)\}}\,,$$

  where b denotes an initial threshold level, i.e., following the initial strategy means paying dividends at the maximum rate if $x \geq b(\vartheta)$, and otherwise paying no dividends.

- We use policy iteration to improve the strategy.
  - For a given Markov strategy $u^{(k)}$ we calculate its associated value $V^{(k)}$ by solving

$$(\mathcal{L}^G - \delta - \nu)V + u^{(k)}(1 - \mathcal{D}_x^G V) + \nu\lambda\mathcal{I}^G V = 0$$

    on a finite grid, where $\mathcal{L}^G$ is the operator $\mathcal{L}$ with differentiation operators replaced by suitable finite differences, $\mathcal{D}_x^G$ is the finite difference operator w.r.t. x, and $\mathcal{I}^G$ is the integral operator replaced by a quadrature rule.
  - Now we determine $u^{(k+1)}$ as the function that maximizes $u(1 - \mathcal{D}_x^G V)$, which is given by the rule $u^{(k+1)}(x,\vartheta) = K\,1_{\{\mathcal{D}_x^G V(x,\vartheta) \leq 1\}}$.

- The iteration stops as soon as $u^{(k+1)} = u^{(k)}$, i.e., one can not achieve any further improvement.

Figure 1 shows the resulting value function and dividend policy for the parameter set $\theta_1 = 1.5$, $\theta_2 = 2$, $\sigma = 1$, $\delta = 0.5$, $K = 1.25$, $\nu = 0.3$, $\lambda = 0.5$.

One can see that in our example for the jump-diffusion case, an ε-optimal dividend policy is of threshold type. This means there is a sufficiently smooth threshold level b such that no dividends are paid, if the surplus is less than b, and dividends are paid at the maximum rate, if the surplus is greater than b. Furthermore, the threshold level naturally depends on the estimate for θ. In our example it decreases monotonically in θ. Thus, our results fit very well to other results on the dividend maximization problem.

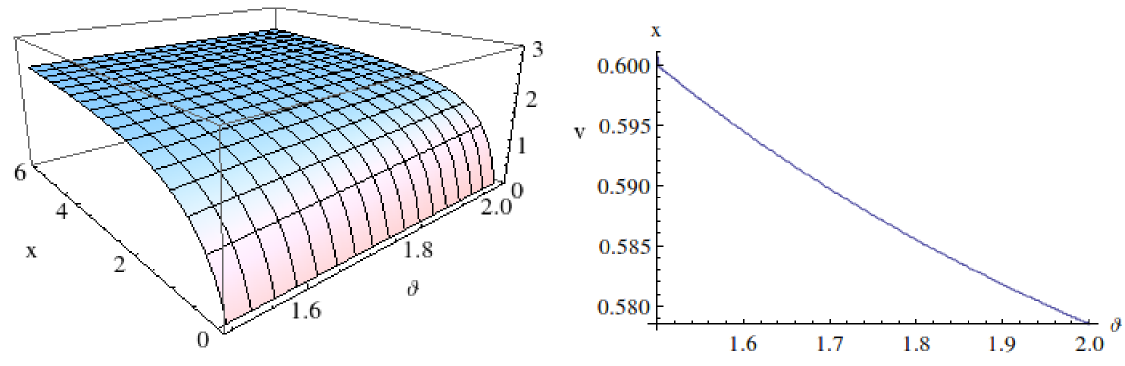


Figure 1: The resulting value function and the dividend policy.

4. CONCLUSION

We have presented a jump-diffusion model in a Bayesian setup for the surplus of an insurance company. In this setup, we have formulated the valuation problem of the company in terms of maximization of the discounted future dividend payments until the time of ruin. To overcome uncertainty we have found a Bayesian filter. We have derived the associated HJB equation, which is an IPDE. Finally, we have presented a way to study the problem numerically.
The numerical example has suggested that a threshold strategy, the threshold level of which is a function of the estimator of the drift, is at least ε-optimal.

References

H. Albrecher and S. Thonhauser. Optimality Results for Dividend Problems in Insurance. *RACSAM Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas*, 103(2):295–320, 2009.

S. Asmussen and M. Taksar. Controlled Diffusion Models for Optimal Dividend Pay-Out. *Insurance: Mathematics and Economics*, 20(1):1–15, 1997.

B. Avanzi. Strategies for Dividend Distribution: A Review. *North American Actuarial Journal*, 13 (2):217–251, 2009.

M. Belhaj. Optimal Dividend Payments when Cash Reserves Follow a Jump-Diffusion Process. *Mathematical Finance*, 20(2):313–325, 2010.

B. De Finetti. Su un'impostazione alternativa della teoria collettiva del rischio. *Transactions of the XVth International Congress of Actuaries*, 2:433–443, 1957.

W. Fleming and H. Soner. *Controlled Markov Processes and Viscosity Solutions*. Stochastic Modelling and Applied Probability. Springer, second edition, 2006.

M. Jeanblanc-Piqué and A.N. Shiryaev. Optimization of the Flow of Dividends. *Russian Math. Surveys*, 50(2):257–277, 1995.

G. Leobacher, M. Szölgyenyi, and S. Thonhauser. Bayesian Dividend Optimization and Finite Time Ruin Probabilities. 2013. Submitted.

R.S. Liptser and A.N. Shiryaev. *Statistics of Random Processes I - General Theory*. Applications of Mathematics. Springer, 1977.

P. Protter. *Stochastic Integration and Differential Equations*. Stochastic Modelling and Applied Probability. Springer, 2004.

R. Radner and L. Shepp. Risk vs. Profit Potential: A Model for Corporate Strategy. *Journal of Economic Dynamics and Control*, 20(8):1373–1393, 1996.

S.E. Shreve, J.P. Lehoczky, and D.P. Gaver. Optimal Consumption for General Diffusions with Absorbing and Reflecting Barriers. *SIAM Journal on Control and Optimization*, 22(1), 1984.

NON-RANDOM OVERSHOOTS OF LÉVY PROCESSES

Matija Vidmar

Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom
Email: m.vidmar@warwick.ac.uk

The class of Lévy processes for which overshoots are almost surely constant quantities is precisely characterized.

1. INTRODUCTION

Fluctuation theory represents one of the most important areas within the study of Lévy processes, with applications in finance, insurance, dam theory etc. (Kyprianou 2006). It is particularly explicit in the spectrally negative case, when there are no positive jumps, a.s. (Sato 1999, Section 9.46) (Bertoin 1996, Chapter VII).

What makes the analysis so much easier in the latter instance is the fact that the overshoots (Sato 1999, p. 369) over a given level are known *a priori* to be constant and equal to zero. As we shall see, this is also the only class of Lévy process for which this is true (see Lemma 3.1). But it is not so much the exact values of the overshoots that matter, as does the fact that these values are non-random (and known). It is therefore natural to ask if there are any other Lévy processes having constant overshoots (a.s.) and, moreover, what *precisely* is the class having this property.

To be sure, in the existing literature one finds expressions regarding the distribution of the overshoots. Unfortunately, these do not seem immediately useful in answering the question posed above, and the following result is proved directly: for the overshoots of a Lévy process to be (conditionally on the process going above the level in question) almost surely constant quantities, it is both necessary and sufficient that *either* the process has no positive jumps (a.s.) *or* for some $h > 0$, it is compound Poisson, living on the lattice $\mathbb{Z}_h := h\mathbb{Z}$, and which can only jump up by h.

The precise and more exhaustive statement of this result is contained in Theorem 2.1 of Section 2, which also introduces the required notation. Section 3 supplies the main line and idea of the proof and Section 4 concludes. A full exposition may be found in Vidmar (2013).

2. NOTATION AND STATEMENT OF RESULT

Throughout we work on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$, which satisfies the standard assumptions (i.e. the σ-field $\mathcal{F}$ is P-complete and the filtration $\mathbb{F}$ is right continuous and $\mathcal{F}_0$ contains all P-null sets). We let X be a Lévy process on this space with characteristic triplet (σ^2, λ, μ) relative to some cut-off function (Sato 1999). $\overline{X}_t := \sup\{X_s : s \in [0,t]\}$ $(t \geq 0)$ is the supremum process of X.

Next, for $x \in \mathbb{R}$ introduce $T_x := \inf\{t \geq 0 : X_t \geq x\}$ (resp. $\hat{T}_x := \inf\{t \geq 0 : X_t > x\}$), the *first entrance time* of X to $[x, \infty)$ (resp. (x, ∞)). We will informally refer to T_x and $\hat{T}_x$ as the *times of first passage* above the level x.

$\mathcal{B}(S)$ will always denote the Borel σ-field of a topological space S, $\mathrm{supp}(m)$ the support of a measure m thereon. For a random element $R : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$, $R_\star\mathsf{P}$ is the image measure.

The next definition introduces the concept of an upwards skip-free Lévy chain, which is the continuous-time analogue of a right-continuous random walk (cf. e.g. Brown et al. (2010)).

Definition 2.1 (Upwards-skip-free Lévy chain) *A Lévy process X is an* upwards skip-free Lévy chain *if it is a compound Poisson process, and for some (then unique) $h > 0$, $\mathrm{supp}(\lambda) \subset \mathbb{Z}_h$ and* $\mathrm{supp}(\lambda|_{\mathcal{B}((0,\infty))}) = \{h\}$.

Finally, the following notion, which is a rephrasing of "being almost surely constant conditionally on a given event", will prove useful:

Definition 2.2 (P-triviality) *Let $S \neq \emptyset$ be any measurable space, whose σ-algebra contains the singletons. An S-valued random element R is said to be* P-trivial *on an event $A \in \mathcal{F}$ if there exists $r \in S$ such that $R = r$ a.s.-P on A (i.e. the push-forward measure $R|_{A*}\mathsf{P}(\cdot \cap A)$ is a weighted (possibly by 0, if $\mathsf{P}(A) = 0$) δ-measure). R may only be defined on some $B \supset A$.*

We can now state succinctly the main result of this paper:

Theorem 2.1 (Non-random position at first passage time) *The following are equivalent:*

1. *For some $x > 0$, $X(T_x)$ is P-trivial on $\{T_x < \infty\}$.*
2. *For all $x \in \mathbb{R}$, $X(T_x)$ is P-trivial on $\{T_x < \infty\}$.*
3. *For some $x \geq 0$, $X(\hat{T}_x)$ is P-trivial on $\{\hat{T}_x < \infty\}$ and a.s.-P positive thereon (in particular the latter obtains if $x > 0$).*
4. *For all $x \in \mathbb{R}$, $X(\hat{T}_x)$ is P-trivial on $\{\hat{T}_x < \infty\}$.*
5. *Either X has no positive jumps, a.s.-P or X is an upwards skip-free Lévy chain.*

If so, then outside a P-negligible set, for each $x \in \mathbb{R}$, $X(T_x)$ (resp. $X(\hat{T}_x)$) is constant on $\{T_x < \infty\}$ (resp. $\{\hat{T}_x < \infty\}$), i.e. the exceptional set in (2) (resp. (4)) can be chosen not to depend on x.

Finally, notation-wise, we make the following explicit: $\mathbb{R}^+ := (0, \infty)$, $\mathbb{R}_+ := [0, \infty)$, $\mathbb{R}^- := (-\infty, 0)$ and $\mathbb{R}_- := (-\infty, 0]$.

3. MAIN LINE OF THE PROOF

Remark 3.1 *T_x and $\hat{T}_x$ are $\mathbb{F}$-stopping times for each $x \in \mathbb{R}$ (apply the début theorem (Kallenberg 1997, p. 101, Theorem 6.7)) and $\mathsf{P}(\forall x \in \mathbb{R}_-(T_x = 0)) = 1$. Moreover, $\mathsf{P}(\forall x \in \mathbb{R}(T_x < \infty)) = 1$, whenever X either drifts to $+\infty$ or oscillates. If not, then it drifts to $-\infty$ (Sato 1999, p. 255, Proposition 37.10) and on the event $\{T_x = \infty\}$ one has $\lim_{t \to T_x} X(t) = -\infty$ for all $x \in \mathbb{R}$, a.s.-P.*

For the most part we find it more convenient to deal with the $(T_x)_{x \in \mathbb{R}}$, rather than $(\hat{T}_x)_{x \in \mathbb{R}}$, even though this makes certain measurability issues more involved.

Remark 3.2 *Note that whenever 0 is regular for $(0, \infty)$, then for each $x \in \mathbb{R}$, $T_x = \hat{T}_x$ a.s.-P (apply the the strong Markov property (Sato 1999, p. 278, Theorem 40.10) at the time T_x). For conditions equivalent to this, see (Kyprianou 2006, p. 142, Theorem 6.5). Conversely, if 0 is irregular for $(0, \infty)$, then with a positive P-probability $\hat{T}_0 > 0 = T_0$.*

The following lemma is shown, for example by appealing to the Lévy-Itô decomposition (Applebaum 2009, p. 108, Theorem 2.4.16).

Lemma 3.1 (Continuity of the running supremum) *The supremum process $\overline{X}$ is continuous (P-a.s.) iff X has no positive jumps (P-a.s). In particular, if $X(T_x) = x$ a.s.-P on $\{T_x < \infty\}$ for each $x > 0$, then X has no positive jumps, a.s.-P.*

The first main step towards the proof of Theorem 2.1 is the following:

Proposition 3.2 (P-triviality of $X(T_x)$) *$X(T_x)$ on $\{T_x < \infty\}$ is a P-trivial random variable for each $x > 0$ iff either one of the following mutually exclusive conditions obtains:*

1. *X has no positive jumps (P-a.s.) (equivalently: $\lambda((0, \infty)) = 0$).*
2. *X is compound Poisson and for some $h > 0$, $\mathrm{supp}(\lambda) \subset \mathbb{Z}_h$ and $\mathrm{supp}(\lambda|_{\mathcal{B}((0,\infty))}) = \{h\}$.*

If so, then $X(T_x) = x$ on $\{T_x < \infty\}$ for each $x \geq 0$ (P-a.s.) under (1) and $X(T_x) = h\lceil x/h \rceil$ on $\{T_x < \infty\}$ for each $x \geq 0$ (P-a.s.) under (2).

The main idea of the proof here is to appeal first to Lemma 3.1 in order to get (1) and then to treat separately the compound Poisson case; in all other instances the Lévy-Itô decomposition and the well-established path properties of Lévy processes yield the claim. Intuitively, for a Lévy process to cross over every level in a non-random fashion, either it does so necessarily continuously when there are no positive jumps (cf. also (Kolokoltsov 2011, p. 274, Proposition 6.1.2)), or, if there are, then it must be forced to live on the lattice $\mathbb{Z}_h$ for some $h > 0$ and only jump up by h.

The second (and last) main step towards the proof of Theorem 2.1 consists in taking advantage of the temporal and spatial homogeneity of Lévy processes. Thus the condition in Proposition 3.2 is strengthened to one in which the P-triviality of the position at first passage is required of one only $x > 0$, rather than all. To shorten notation let us introduce:

Definition 3.1 *For $x \in \mathbb{R}$, let $\mathsf{Q}^x := X(T_x)_\star \mathsf{P}(\cdot \cap \{T_x < \infty\})$ be the (possibly subprobability) law of $X(T_x)$ on $\{T_x < \infty\}$ under P on the space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We also introduce the set*

$$\mathcal{A} := \{x \in \mathbb{R} : \mathsf{Q}^x \textit{ is a weighted (possibly by } 0\textit{) } \delta\textit{-distribution}\}.$$

Remark 3.3 *Clearly* $(-\infty, 0] \subset \mathcal{A}$.

With this at our disposal, we can formulate our claim as:

Proposition 3.3 *Suppose* $\mathcal{A} \cap \mathbb{R}^+ \neq \emptyset$. *Then* $\mathcal{A} = \mathbb{R}$.

The proof of this proposition proceeds in several steps, but the gist of it consists in establishing the intuitively appealing identity $\mathsf{Q}^b(A) = \int d\mathsf{Q}^c(x_c)\mathsf{Q}^{b-x_c}(A - x_c)$ for Borel sets A and $c \in (0, b)$ (where Q^c must be completed). This is used to show that $\mathcal{A}$ is dense in the reals, and then we can appeal to quasi-left-continuity to conclude the argument. The main argument is thus fairly short, and a substantial amount of time is spent on measurability issues.

Finally, Proposition 3.2 and Proposition 3.3 are easily combined into a proof of Theorem 2.1.

4. CONCLUSION

Theorem 2.1 characterizes the class of Lévy processes for which overshoots are known *a priori* and are non-random. Moreover, the original motivation for this investigation is validated by the fact that upwards skip-free Lévy chains admit for a fluctuation theory just as explicit and almost (but not entirely) analogous to the spectrally negative case — but this already falls outside the strict scope of this work, rather it presents its natural continuation.

Acknowledgment: The support of the Slovene Human Resources Development and Scholarship Fund under contract number 11010-543/2011 is recognized.

References

D. Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2009.

J. Bertoin. *Lévy Processes*. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1996.

Mark Brown, Erol A. Peköz, and Sheldon M. Ross. Some results for skip-free random walk. *Probability in the Engineering and Informational Sciences*, 24:491–507, 2010.

O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, New York Berlin Heidelberg, 1997.

V. N. Kolokoltsov. *Markov Processes, Semigroups, and Generators*. De Gruyter Studies in Mathematics. De Gruyter, 2011.

A. E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer-Verlag, Berlin Heidelberg, 2006.

K. I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge studies in advanced mathematics. Cambridge University Press, Cambridge, 1999.

M. Vidmar. Non-random overshoots of Lévy processes. 2013. arXiv:1301.4463 [math.PR].

De Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten coördineert jaarlijks tot 25 wetenschappelijke bijeenkomsten, ook contactfora genoemd, in de domeinen van de natuurwetenschappen (inclusief de biomedische wetenschappen), menswetenschappen en kunsten. De contactfora hebben tot doel Vlaamse wetenschappers of kunstenaars te verenigen rond specifieke thema's.

De handelingen van deze contactfora vormen een aparte publicatiereeks van de Academie.

Contactforum "Actuarial and Financial Mathematics Conference" (7-8 februari 2013, Prof. M. Vanmaele)

Deze handelingen van de "Actuarial and Financial Mathematics Conference 2013" geven een inkijk in een aantal onderwerpen die in de editie van 2013 van dit contactforum aan bod kwamen. Zoals de vorige jaren handelden de voordrachten over zowel actuariële als financiële onderwerpen en technieken met speciale aandacht voor de wisselwerking tussen beide. Deze internationale conferentie biedt een forum aan zowel experten als jonge onderzoekers om hun onderzoeksresultaten ofwel in een voordracht ofwel via een poster aan een ruim publiek voor te stellen bestaande uit academici uit binnen- en buitenland alsook collega's uit de bank- en verzekeringswereld.