A Positive Interest Rate Model with Sticky Barrier

Yuri Kabanov*, Masaaki Kijima† and Sofiane Rinaz‡§

June 12, 2006

Abstract

This paper proposes an efficient model for the term structure of interest rates when the interest rate takes very small values. We make the following choices: (i) we model the *short-term* interest rate, (ii) we assume that once the interest rate reaches zero, it stays there and we have to wait for a random time until the rate is reinitialized to a (possibly random) strictly positive value. This setting ensures that all term rates are strictly positive.

Our objective is to provide a simple method to price zero-coupon bonds. A basic statistical study of the data at hand indeed suggests a switch to a different mode of behavior when we get to a low level of interest rates. We introduce a variable for the time already spent at 0 (during the last stay) and derive the pricing equation for the bond. We then solve this partial integro-differential equation (PIDE) on its entire domain using a finite difference method (Cranck-Nicholson scheme), a method of characteristics and a fixed point algorithm. Resulting yield curves can exhibit many different shapes, including the S shape observed on the recent Japanese market.

Keywords: short-term interest rate models, partial integro-differential equation, zero-interest rate, finite difference methods

JEL Classification: E43, C63, G12

*Université de Franche-Comté, 16 Route de Gray, F-25030 Besancon Cedex, France, and Central Economics and Mathematics Institute, Moscow, Russia. This research was done during the stay of the author at the Daiwa Securities Chair of Graduate School of Economics, Kyoto University in Spring 2004.

†Graduate School of Social Sciences, Tokyo Metropolitan Univeristy, and Daiwa Securities Chair, Graduate School of Economics, Kyoto University.

‡Shinsei Bank Ltd., 1-8, Uchisaiwaicho 2-Chome, Chiyoda-ku, Tokyo 100-8501 (e-mail: Sofiane.Rinaz@Shinseibank.com, corresponding author)

§The third author acknowledges financial support from the Monbukagakusho. We express our gratitude for helpful comments to Rama Cont, Monique Jeanblanc, Yue Kuen Kwok, Yoshio Miyahara, Motoh Tsujimura, Nick Webber, two anonymous referees, and participants at the 2004 Daiwa International Workshop on Financial Engineering and QMF 2005 in Sydney. All remaining errors are ours, of course.

1 Introduction

Many attempts have been made to model the term structure of interest rates since the second half of the 1970s. However no model has managed to capture all the characteristics of the qualitative behavior of interest rates, and no model has proved superior to all the others on every point of comparison. Moreover, empirical studies (e.g. Sun, 2003) suggest that the qualitative behavior may be very different depending on the country and period of time. Our aim here is to propose an adequate model for the current Japanese short rate.

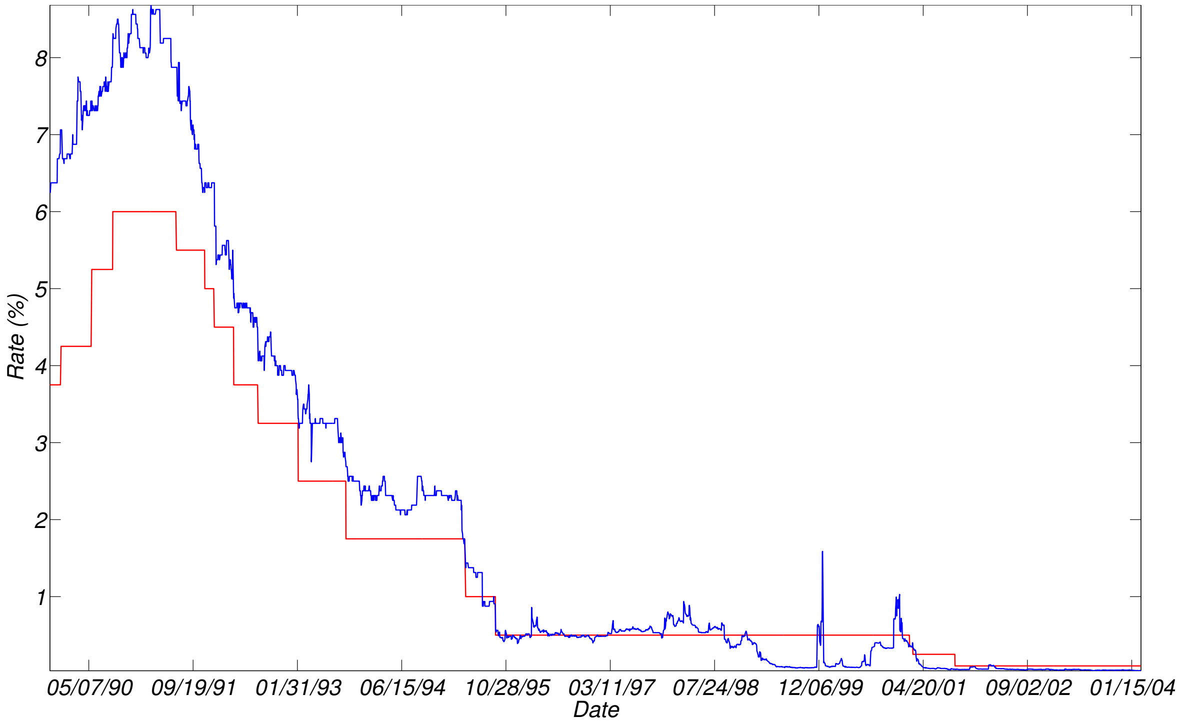

Figure 1: 1-month LIBOR in yen between November 1989 and April 2004 ($3,644$ daily observations). The stairs are the official lending rates from the Bank of Japan.

One of the striking specificities of the Japanese short rate is that it went from a period of oscillations in relatively high levels until 1995, to a period of near-zero rates since 1995. The switch between these two *regimes* is well-pronounced (see Figure 1) and, as we will explain in the next few paragraphs, this continuing near-zero regime is actually quite challenging to model: the existing ones do not capture such a feature.

First, we notice that all Gaussian interest rate models since the one of Vašíček[1] (1977)

[1]Vašíček (1977) assumed that the short rate process r_t evolves according to the dynamics

$$dr_t = \kappa[\theta - r_t]dt + \sigma W_t, \tag{1}$$

where θ is the long-term mean, κ the speed of convergence to the mean, σ the volatility and W_t a standard allow negative rates. This could appear not too problematic when the interest rate was high enough, especially with a strong mean-reverting drift, since then the probability of negative rates in the future was small. But even so, as shown in Rogers (1996), some derivatives' prices are extremely sensitive to the *possibility* of negative interest rates, and especially in the current situation this may greatly impair our computation of bonds and derivatives' prices in the Japanese market.

Several attempts have been made to solve the problem of possible negative rates. First, other (non-Gaussian) processes have been considered, implying different distributions for the short rate r; non-central χ^2 (Cox, Ingersoll and Ross, 1985) and lognormal (Dothan, 1978; Black and Karasinski, 1991). Second, the drift term has been set to be non-linear as in Aït-Sahalia (1996); it would depend on the interest rate and tend to infinity as the rate tends to 0. Third, Goldstein and Keirstead (1997) tried to set a special boundary condition at 0, reflecting or absorbing. Fourth, instead of modelling the short rate, Flesaker and Hughston (1996) proposed to model directly the bond price, and showed that they actually consider a class of positive HJM models (see also Cairns, 2004). At last, Black (1995) introduced the idea of considering interest rates as options.[2]

However, none of these solutions is perfect:

(a) Non-central χ^2 distributions are not as widely used in finance as normal distributions, and parameters have to satisfy certain conditions (the Feller condition) to ensure that 0 is unattainable and to grant that the short rate remains positive. Lognormal distributions also have a weakness; their fat tail induces the explosion of the bank account, and so prices can only be computed using trees in practice (thus truncating the tail). Moreover, in both cases (χ^2 and lognormal distributions), the absolute volatility vanishes as the rate approaches zero, and the mean-reverting drift, left without opposing force, pulls the rate back up, making the occurence of sustained periods of near-zero interet rates very improbable.[3] Even the popular market model of Brace et al. (1997), which is based

Brownian motion. The parameters θ, κ and σ are constant or deterministic.

[2] Black's idea was that the rate cannot become negative because if it does, investors can just invest in currency. The interest rate would thus be the *positive part* of, say, an Ornstein-Uhlenbeck process, like in the Vašíček model. The original Gaussian process is then called the *shadow interest rate*.

[3] This is a direct consequence of the dynamics of the short rate. Under the CIR model (Cox et al., 1985) the short rate has the following dynamics:

$$dr_t = \kappa[\theta - r_t]dt + \sigma\sqrt{r_t}W_t,$$

and under the BK model (Black and Karasinski, 1991):

$$d\ln r_t = \kappa[\theta - \ln r_t]dt + \sigma W_t,$$

where θ, κ, σ and W_t are defined as in (1).

on a lognormal assumption for the dynamics of the forward rates, cannot satisfactorily render the possibility of long periods of near zero short-rate.

(b)Non-linear drifts have been shown to be only a feature of daily data rather than weekly or monthly data (see Jones, 2003) and less essential in the modelling choices than stochastic volatility (see Sun, 2003).

(c) Reflecting or absorbing boundary conditions in 0 lead to closed-form solutions for bond prices, but are really not justifiable economically.

(d) The Flesaker and Hughston model can reproduce sustained periods of both high and low interest rates (see Cairns, 2004), but is driven by unobservable state variables and the fact that it is not a direct representation of the short-rate makes it more difficult to apprehend in practice. Moreover, once again, 0 is unattainable by the short rate.

(e) Black's idea of interest rates as options has been recently developed by Gorovoi and Linetsky (2004) who provided analytical expressions for both bond prices and bond options prices. Nevertheless, this implies a rate that stays exactly at 0 for a long time (with a zero volatility). Also, the shadow interest rate is by definition unobservable in the 0-interest rate state. It can then only be estimated through the current yield curve. In 2002, from actual bond prices, this shadow interest rate was estimated to be -5%, which is hard to reconciliate with any standard economic indicator.[4]

After an exploratory analysis of the data at hand, we conclude that the Japanese interest rate experiences two regimes. Up to 1995 or so, we observe a strong correlation with economic indicators like the consumer price index (CPI), the inflation rate and the dollar/yen exchange rate. After that, the short rate fell consistently below 1% and it did no longer exhibit correlation with such economic indicators, as if there had been a break in regime. However, it seems that there is no reason to believe that the second regime is a perpetual one.

Our idea is to use a process with two different regimes for the short rate. While being strictly positive, the short rate is assumed to follow the Vašíček model with linear mean-reverting drift. When the rate reaches 0, a new regime is set; the rate stays at zero for a random time and then jumps to a strictly (possibly random) positive value and switches back to the usual Vašíček regime. In the Feller classification of boundaries for diffusions, 0 would be a *regular sticky* boundary (see for example Karlin and Taylor, 1981), but note that our process is actually not a diffusion in the strict sense, since its paths are not continuous in general.

This idea of regime switching has already been explored recently by e.g. Gray (1996), Hansen and Poulsen (2000) and Bansal and Zhou (2002). These models typically assume

[4]For example, at that time, the inflation rate (deflation actually) was never below -1%.

that the spot-rate process can shift randomly between two or more regimes, typically a high-mean high-volatility regime and a low-mean low-volatility regime. Although these models can technically reproduce sustained periods of near-zero interest rates, they still suffer from the possible flaws of the spot-rate model they are based on (e.g. non-positivity in the case of the Vasicek model), and they are often hard to calibrate to the current yield curve.

To draw a link between the model exposed here and all the previous ones, we may look at the asymptotic behavior of our model. When the time spent in 0 tends to infinity, the model is similar to the one with absorbing state in 0. Alternatively, if the time spent in 0 tends to zero, and the jump size is very small, we are close to the model with reflecting condition in zero. Also, Black's model with no drift (the shadow rate is a martingale) can be embedded in our model if we set the time of jumps to be normal inverse Gaussian (NIG) distributed, and the jump size to be very small.

Recently, Marumo et al. (2003), from Bank of Japan, proposed the following dynamics for the short rate:

$$dr_t^{\tau} = 1_{\{t \geq \tau\}}(\kappa[\theta - r_t]dt + \sigma W_t),$$

where W_t is a standard Brownian motion, $\tau > 0$ is a random stopping time, and θ, κ and σ are constant. This means that the short rate is initially 0 and, at a future random time, the short-rate will be reinitialized at a random value and from then on will follow the dynamics of an Ornstein–Uhlenbeck process forever. The authors argue that the Zero Interest Rate Policy (ZIRP) is transitory and that the Bank of Japan can and will put an end to it at its convenience. Of course this setting does not ensure positivity of the short rate at all. Our model can be seen as an extension of this work to the case when other ZIRPs happen in the future.

The rest of this paper is organized as follows. We examine shortly some data from the Japanese economy in Section 2 to further motivate our setting. In Section 3 we introduce our model and derive the pricing equation for zero-coupon bonds. An algorithm to solve this equation is explained in Section 4. Examples of yield curves are given in Section 5. Section 6 extends the model to introduce another state variable while the short rate is sleeping in 0. Section 7 concludes the article.

2 A Quick Look at Historical Data

We used data from several sources. For the short rate, we used the 1-month LIBOR rate in yens.[5] The LIBOR rate is determined once every day in London by averaging the results of a telephone poll of major banks. We had a total of 3, 644 daily observations. We also used the CPI index (monthly data) and the official central bank rate (CBRATE) for loans to affiliated banks (data available since 1970, and provided by the Bank of Japan). Finally, we obtained the yen-dollar exchange rate (FXRATE) from the web site of the US Federal Reserve in Saint-Louis (see Figure 2).

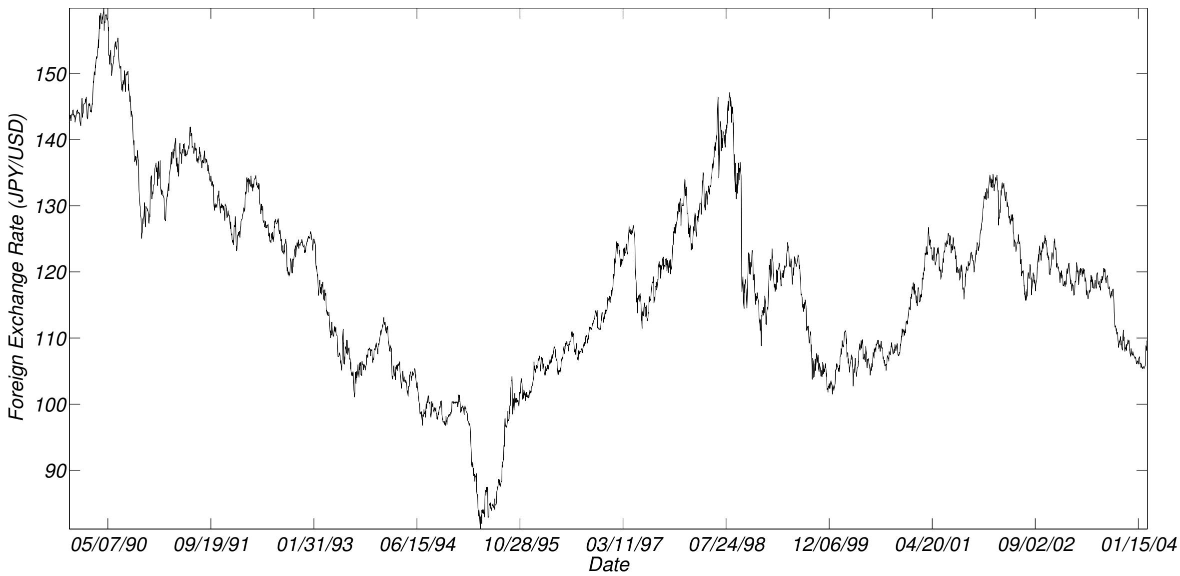

Figure 2: Exchange rate in yen per dollar in the period November 1989 to April 2004

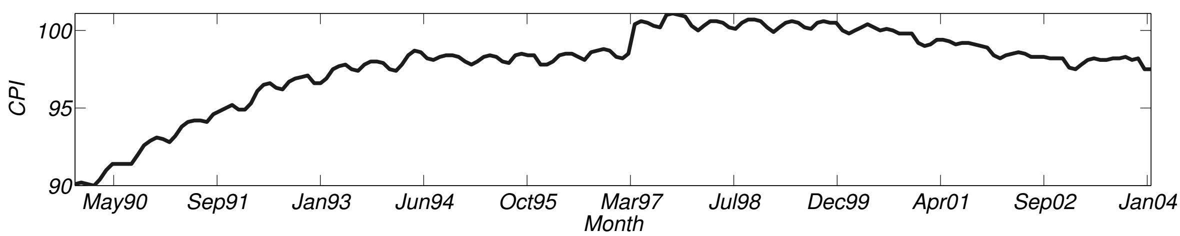

Figure 3: CPI since November 1989 in Japan (Base: 100 in January 2000)

[5]This is a questionable choice; it is not exactly the *short* rate, it has some counterparty risk in it (leading to probable presence of outliers, after large corporate bankruptcies for example), it is not directly related to the yields of Japanese Government Bonds (JGBs) which is the object we want to study, and -last but not least- it has a relatively short history (we only had the data since November 1989).

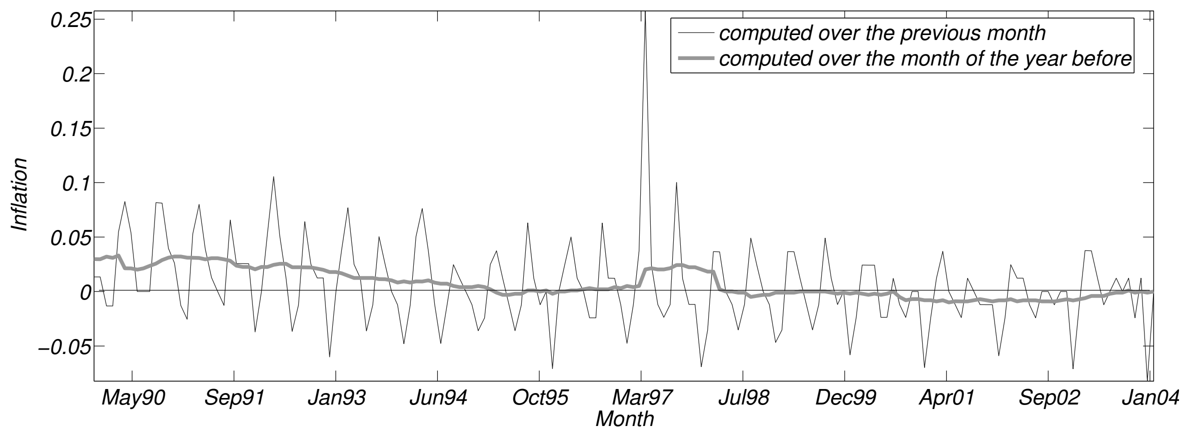

Figure 4: Monthly inflation rate compounded annually over the past month and over the past year since November 1989

First, to get series of the same length, we truncated the data to keep only the information in the period between November 1989 and April 2004. We also had to make the data match when on a national holiday. From the CPI index, we computed the monthly (compounded annually) and yearly inflation rates (respectively InfMonth and InfYear), and we transferred all these monthly data to a "daily" format, assuming the rates were constant for the full month (see Figures 3 and 4).

We first computed the summary of the obtained data (see Table 1), and the correlations between variables (see Table 2). Since the data set seems to be split in two, we decided to consider a first set of data until the first time we get below 1% interest rate, the rest of the data being considered as a second set of data. Sample correlation results for the two data sets can be found in Tables 3 and 4.

	CBRate	FXRate	CPI	InfMonth	InfYear	LIBOR
Minimum	0.100	81.12	90.0	-0.0822	-0.0099	0.0375
1st Qu.	0.500	107.59	97.5	-0.0122	-0.0020	0.0912
Median	0.500	118.76	98.3	0.0000	0.0020	0.5585
Mean	1.611	118.58	97.8	0.0061	0.0070	2.0626
3rd Qu.	2.500	127.55	99.8	0.0248	0.0203	3.2500
Maximum	6.000	159.90	101.1	0.2576	0.0331	8.6850

Table 1: Summary of the data

During the first period, we observe that all the variables, except the monthly inflation, are strongly correlated; but during the second period (interest rates near zero) we can hardly notice any correlation (even the correlation between the LIBOR and the central bank rate is merely 74%). Note that the correlation between CPI (which is an index, and

	CBRate	FXRate	CPI	InfMonth	InfYear	LIBOR
CBRate	1	0.493	-0.843	0.251	0.839	0.990
FXRate	0.493	1	-0.541	0.174	0.479	0.532
CPI	-0.843	-0.541	1	-0.159	-0.664	-0.887
InfMonth	0.251	0.174	-0.159	1	0.307	0.260
InfYear	0.839	0.479	-0.664	0.307	1	0.846
LIBOR	0.990	0.532	-0.887	0.260	0.846	1

Table 2: Correlation between variables during the whole period

	CBRate	FXRate	CPI	InfMonth	InfYear	LIBOR
CBRate	1	0.835	-0.787	0.257	0.895	0.973
FXRate	0.835	1	-0.895	0.248	0.889	0.890
CPI	-0.787	-0.895	1	-0.183	-0.800	-0.886
InfMonth	0.257	0.248	-0.183	1	0.256	0.281
InfYear	0.895	0.889	-0.800	0.256	1	0.905
LIBOR	0.973	0.890	-0.886	0.281	0.905	1

Table 3: Correlation between variables from November 1989 to July 6, 1995

	CBRate	FXRate	CPI	InfMonth	InfYear	LIBOR
CBRate	1	-0.338	0.491	0.073	0.429	0.740
FXRate	-0.338	1	0.242	0.059	0.041	-0.071
CPI	0.491	0.242	1	0.188	0.441	0.256
InfMonth	0.073	0.059	0.188	1	0.233	0.093
InfYear	0.429	0.041	0.441	0.233	1	0.569
LIBOR	0.740	-0.071	0.256	0.093	0.569	1

Table 4: Correlation between variables after July 7, 1995

has the dimension of a price) and the LIBOR rate (which is a rate) is probably due to a pure coincidence. It could be the case for other correlations too. Also, note that we did not consider lags, here. Maybe the correlation between the monthly inflation and the LIBOR rate would be stronger if we considered the inflation rate a few months ago, since anyway the exact value of the inflation over a given month is only known a few months later by the market participants.

From these observations, it seems plausible to draw two conclusions:

- There is no evidence of the necessity to use a multifactor model. Even if we introduce an exogenous factor in the short rate model in the first period, it will not bring much since, due to the strong correlation, the short rate would essentially be equal to that exogenous factor, with just some kind of constant multiplicative factor. At least we can say that from these data, we cannot use inflation or foreign exchange rate as factors of a multifactor model.[6]

- There seems to be more than a "visible" break in the data when the short rate approaches zero. The LIBOR rate just looks like it is not obeying any rule any more (it is not even closely following the fluctuations of the central bank's interest rate).

In the next section, we try to associate some mathematical properties to these somewhat qualitative observations, and we introduce our model.

3 The Model

3.1 Basic Setting

We work in a filtered probability space $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\geq 0}, P)$. We assume that the spot rate $(r_t)_{t\geq 0}$ is a process evolving in $[0, \infty)$ which has two modes of behavior; when it is strictly positive, it behaves as an Ornstein–Uhlenbeck (OU) process[7], and when it reaches 0 it stays there for some time, the probability to get out of 0 at time t depending on Z_t, the time already spent at 0 (during the last stay). The "clock" process $(Z_t)_{t\geq 0}$ will be

[6]Note that we did not consider the correlation between *longer maturity* rates and macroeconomic factors. It still seems plausible that even during zero-interest rate periods, the inflation or foreign exchange rates have a large impact on long-term interest rates, and therefore multifactor models can certainly give a more realistic evolution of the whole yield curve. However, in this paper we chose to focus on the impact of the short rates on the whole yield curve.

[7]This process can be a general diffusion as long as we can employ a numerical technique similar to the one we use here.

called *latency*. We assume that P is a martingale measure[8] and the dynamics of (r_t, Z_t) is as follows:

$$\begin{cases} dr_t &= 1_{\{r_{t-}>0\}}[\kappa(\theta - r_t)dt + \sigma dW_t] + 1_{\{r_{t-}=0\}}dJ_t, \quad r_0 = \bar{r}, \\ dZ_t &= 1_{\{r_{t-}=0\}}[dt - Z_t \frac{dJ_t}{|dJ_t|}], \quad Z_0 = \bar{Z}, \end{cases} \tag{2}$$

where (J_t) is a pure-jump increasing process whose characteristics depend on the latency (Z_t), θ is the long term mean like in the standard Vašíček model, κ is the speed or rate of reversion to the mean, σ is the volatility and (W_t) is the standard Brownian motion. We assume that κ, θ, σ, $\bar{r}$, and $\bar{Z}$ are positive constants (at least for now). We add the additional requirement that at least one of $\bar{r}$ and $\bar{Z}$ is null. We also assume that the processes (W_t) and (J_t) are independent. Finally, it is understood that the notation

$$Z_t \frac{dJ_t}{|dJ_t|}$$

stands for $Z_{t-}1_{\{J_t \neq J_{t-}\}}$.

We can observe from the above formula that, indeed, as long as the short rate remains strictly positive, it evolves as an OU process (that is, as the standard Vašíček model). After hitting 0, the process r "sleeps" there and the latency Z_t starts increasing at the unit time-rate until a jump of the process J arrives. At the time of the jump, the process Z falls to zero and remains there while the short rate process is reinitialized at the (positive) value of the jump.

Of course, this setting implies that the short rate process evolves in the domain[9] $[0, \infty)$. We may call the endpoint 0 a *sticky barrier* following the denomination of Karlin and Taylor (1981). In general, (r_t) alone is not a Markov process since the probability to get out of zero depends on the time already spent in 0. But the couple (r_t, Z_t) has the Markov property.

[8]The assumption that P is already an equivalent martingale measure (EMM) implies certain simplifications. If we start from the real world measure, then we will need to make a change of measure to formulate our pricing equation in an EMM. Because of the presence of jumps with possibly infinite number of sizes, there might be an infinite number of EMMs. So we would have to specify a way to find one EMM, possibly discuss about the choice of the optimal EMM (e.g., minimum entropy martingale measure), and the definition of no-arbitrage in this market. We do not consider this aspect of the discussion here. For an example of selection of EMM in the case of geometric Lévy processes, see Miyahara and Fujiwara (2003).

[9]Note that we could choose to place the endpoint at any rate ϵ instead of 0 and let the short rate evolve in the domain $[\epsilon, \infty)$. This would not change the conclusions of this argument, but it would add one more ϵh term in equation (8) and would make the subsequent expressions of the bond price more complicated. To keep this paper as clear and concise as possible, as well as because economically 0 seems like the only *floor* that every agent would agree on for the short rate, we chose to fix the endpoint at 0.

Because we consider that the time already spent in 0 influences the intensity of the jumps, we are not limited to strictly memoryless (i.e. exponential) distributions for the interarrival times. This is needed because, as we will see later in the section about numerical applications, exponential distributions do not seem to be able to lead to S-shaped yield curves while the Japanese bond market exhibits such shapes. Let us define the jump intensity (or *hazard rate*) $\lambda_t(Z_t)$ such that

$$N_t - \int_0^t \lambda_s(Z_s)ds$$

is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$, where N_t is the number of jumps up to time t.

Note that in this general setting the jump size is random, but we may simplify this by fixing a constant jump size J_0. If the jump size is random, then we have

$$J_t := \sum_{i=1}^{N_t} U_i = \int_0^t \int_0^\infty x\mu(dx, ds),$$

where $(U_i)_{i\geq 1}$ are random variables representing the jump sizes, $\mu(dx, ds)$ is the measure of jumps and its compensator $\nu(dx, ds)$ is defined by

$$\nu(dx, ds) := \lambda_s(Z_s)G(Z_s, dx)ds.$$

Here we assume that the probability distribution function of jump sizes is $G(Z_s, dx)$. In the case when $G(Z_s, dx) = G(dx)$ the jump process considered is just compound renewal process.

To guarantee the existence of the measure $\mu(dx, ds)$, we must check that the jump process (J_t) does not have an explosive behavior. It is the case when λ does not depend on s and the arrival times of jumps form a renewal process which has always non-explosive realizations (see, for example, Last and Brandt, 1995, p. 9). Hereafter, we will assume that it is indeed the case.

3.2 Bond Prices

We are interested in computing the current price $P(\bar{r}, \bar{Z}, T)$ of a zero coupon bond maturing at time T. This price is given by

$$P(\bar{r}, \bar{Z}, T) := E[e^{-\int_0^T r_s ds} | r_0 = \bar{r}, Z_0 = \bar{Z}]. \tag{3}$$

Note that because (r, Z) is time-homogeneous, $P(r_t, Z_t, T - t)$ is the price at time t of a zero coupon bond maturing at T.

Now we want to establish the Feynman–Kac representation (or backward Kolmogorov equation) associated with (3). To do that, let us recall that the Itô formula with jumps applied to r (see, for example, Duffie, 2001, p.348) is given by

$$\begin{aligned} f(r_t) &= f(r_0) + \int_0^t f'(r_{s-})dr_s + \frac{1}{2}\int_0^t f''(r_s)\sigma^2 1_{\{r_{s-}>0\}}ds \\ &\quad + \sum_{0<s\le t}[f(r_s) - f(r_{s-}) - f'(r_{s-})(r_s - r_{s-})] \\ &= f(r_0) + \int_0^t f'(r_{s-})dr_s + \frac{1}{2}\int_0^t f''(r_s)\sigma^2 1_{\{r_{s-}>0\}}ds \\ &\quad + \int_0^t\int_0^\infty [f(r_{s-}+x) - f(r_{s-}) - f'(r_{s-})x]1_{\{r_{s-}=0\}}\mu(dx,ds) \end{aligned}$$

for any $f:\mathbb{R}\to\mathbb{R}$ twice continuously differentiable.

Now we can derive the term structure equation as in, for instance, Björk (1998). We assume that there are bonds for several maturities in the market. The price at time t of the discount bond maturing at time T (T-bond, for short) can be rewritten as

$$P(r_t, Z_t, T-t) =: f^T(t, r_t, Z_t),$$

where we assume that f is a smooth function of its 3 variables, and T is considered as a parameter.

From the Itô formula (slightly more general than that above), we get the following price dynamics for the T-bond:

$$df^T = \alpha^T(t, r_t, Z_t)dt + \int_0^\infty \gamma^T(t, r_t, Z_t, x)\lambda(Z_t)G(Z_t, dx)dt + dM_t,$$

where

$$M_t := \int_0^t \sigma^T(s, r_s, Z_s)dW_s + \int_0^t\int_0^\infty \gamma^T(t, r_{s-}, Z_{s-}, x)(\mu(dx,ds) - \nu(dx,ds))$$

and

$$\begin{aligned} \alpha^T(t,r,Z) &= f_t^T + 1_{\{r>0\}}[\kappa(\theta - r)f_r^T + \frac{1}{2}\sigma^2 f_{rr}^T] + 1_{\{r=0\}}f_Z^T, \\ \sigma^T(t,r,Z) &= 1_{\{r>0\}}\sigma f_r^T, \\ \gamma^T(t,r,Z,x) &= 1_{\{r=0\}}[f^T(t,x,0) - f^T(t,0,Z)]. \end{aligned}$$

Here the subscripts t and r denote partial derivatives with respect to the first and second variable respectively, and functions are evaluated at point (t,r,Z) unless otherwise specified.

We now apply the Itô formula to $e^{-\int_0^t r_s ds}P(r_t, Z_t, T-t)$ which should be a martingale under our choice of the reference measure. We use the fact that

$$d(e^{-\int_0^t r_s ds}) = -r_t e^{-\int_0^t r_s ds}dt,$$

and we get by direct computation that

$$\begin{aligned}
d(e^{-\int_0^t r_s ds}P(r_t, Z_t, T-t)) &= d(e^{-\int_0^t r_s ds})f^T + e^{-\int_0^t r_s ds}df^T \\
&= -r_t e^{-\int_0^t r_s ds}f^T dt + e^{-\int_0^t r_s ds}df^T \\
&= e^{-\int_0^t r_s ds}[-r_t f^T dt + df^T].
\end{aligned}$$

For $e^{-\int_0^t r_s ds}P(r_t, Z_t, T-t)$ to be a martingale, the dt-term obtained in the expression above by substitution of the stochastic differential df^T must be equal to zero. Thus, we arrive to the conclusion that the function $f^T(t, r, Z)$ satisfies the following partial integro-differential equation:

$$\begin{aligned}
-rf^T + f_t^T + 1_{\{r>0\}}[\kappa(\theta - r)f_r^T + \frac{1}{2}\sigma^2 f_{rr}^T] + 1_{\{r=0\}}f_Z^T \\
+ 1_{\{r=0\}}\lambda(Z)\int_0^\infty [f^T(t, x, 0) - f^T(t, 0, Z)]G(Z, dx) = 0,
\end{aligned} \tag{4}$$

with the associated boundary condition at T

$$f^T(T, r, Z) = 1. \tag{5}$$

Unfortunately, an analytical solution to the above equation (4)–(5) is difficult to obtain. The difficulty lies in the fact that the domain of definition of the bond pricing function is a non-standard one: the union of the two orthogonal semi-infinite rectangles

$$\Theta_1 = \{(t, r, Z) : \ 0 \le t \le T, \ r > 0, \ Z = 0\}$$

and

$$\Theta_2 = \{(t, r, Z) : \ 0 \le t \le T, \ r = 0, \ Z \ge 0\}.$$

The bond price should be continuous and we shall try to find the function which is continuous including the joint boundary

$$\Sigma = \{(t, r, Z) : \ 0 \le t \le T, \ r = 0, \ Z = 0\}.$$

We can solve the equation independently on each of the two rectangles without imposing a boundary condition on Σ but match them here. We shall employ a finite difference method on Θ_1, the (exact) method of characteristics on Θ_2, and then a fixed point algorithm to satisfy the continuity condition on Σ.

Note that at this stage we need to specify $\lambda_s(Z_s)$ and give an explicit expression of it. We shall assume in the next section that λ is of the form

$$\lambda(z) = \alpha + \beta\gamma z^{\gamma-1}, \tag{6}$$

where α and β are positive constants, and γ is a constant greater than 1. Jump interarrival times are exponentially distributed if and only if λ is constant (i.e. $\gamma = 1$ or $\beta = 0$). The case $\alpha = 0$ corresponds to the Weibull distribution which was the one considered by Marumo et al. (2003) and which is of a frequent use in many reliability applications.

4 Using Finite Difference Methods

In this section we study the case when λ is of the form given in (6) and there is only one possible jump size J_0. However, the same method can be used for any integrable function λ of Z and any discrete, finite spectrum of possible jump sizes.

From the form of equation (4) – (5), it seems natural to assume that f^T is of the form

$$f^T(t,r,Z) = 1_{\{r>0\}}g(t,r) + 1_{\{r=0\}}h(t,Z),$$

where g and h are solutions of the following equations (7) and (8) respectively:

$$\begin{cases} -rg + g_t + [\kappa(\theta - r)g_r + \frac{1}{2}\sigma^2 g_{rr}] = 0, \\ g(T,r) = 1, \end{cases} \tag{7}$$

and

$$\begin{cases} h_t + h_Z - \lambda(Z)h + \lambda(Z)g(t,J_0) = 0, \\ h(T,Z) = 1, \end{cases} \tag{8}$$

with the additional matching condition

$$\lim_{r\to 0} g(t,r) = h(t,0). \tag{9}$$

Equation (8) is simple to solve, provided we already know an expression for $g(.,J_0)$. We can use the method of characteristics. First we introduce the dummy variable s and we write

$$\frac{dh}{ds} = \frac{dt}{ds}h_t + \frac{dZ}{ds}h_Z.$$

Next, we identify the coefficients of the above equation with those of equation (8). We get

$$\begin{aligned} \frac{dt}{ds} &= 1, \\ \frac{dZ}{ds} &= 1, \\ \frac{dh}{ds} &= \lambda(Z)h - \lambda(Z)g(t,J_0). \end{aligned}$$

We also rewrite the boundary condition as

$$\begin{aligned} t(0,u) &= T, \\ Z(0,u) &= u, \quad u \in [0,\infty), \\ h(0,u) &= 1. \end{aligned}$$

From these equations, we can immediately get an expression for t and Z, function of (s,u), as follows:

$$\begin{aligned} t(s,u) &= T+s, && (10)\\ Z(s,u) &= u+s. && (11)\end{aligned}$$

The equation for $h(s,u)$ becomes

$$\frac{dh}{ds} = \lambda(u+s)h - \lambda(u+s)g(T+s,J_0),$$

which has the following solution:

$$h(s,u) = e^{\int_0^s \lambda(u+x)dx}\left(-\int_0^s \frac{\lambda(u+v)g(T+v,J_0)}{e^{\int_0^v \lambda(u+y)dy}}dv + K\right),$$

where $K = K(T)$ is a constant that may only depend on the parameter T and can be evaluated from the boundary condition. Since $h(0,u)=1$, we conclude that $K=1$.

Now we assume that $\lambda(Z)$ is of the form given in (6). Equation (4) becomes

$$h(s,u) = e^{\alpha s+\beta[(s+u)^\gamma-u^\gamma]}\left(-\int_0^s \frac{[\alpha+\beta\gamma(u+v)^{\gamma-1}]g(T+v,J_0)}{e^{\alpha v+\beta[(v+u)^\gamma-u^\gamma]}}dv+1\right).$$

Note that we can get an expression of h function of the original variables (t,Z). By using (10) and (11), the change of variable $w=v+T$ and after simplification, we obtain

$$h(t,Z) = \frac{e^{\beta Z^\gamma}}{e^{\alpha(T-t)}}\left(\frac{1}{e^{\beta(Z+T-t)^\gamma}} + \int_t^T \frac{e^{\alpha(T-w)}}{e^{\beta(Z+w-t)^\gamma}}[\alpha+\beta\gamma(Z+w-t)^{\gamma-1}]g(w,J_0)dw\right). \tag{12}$$

Note that equation (12) implies that in the case $\beta=0$, i.e. for the model with constant λ, the bond price tends to be constant (equal to 1) as λ tends to 0 (i.e. when the average time of stay at 0 tends to ∞).

We want to use a finite difference method to solve (7). We define a mesh $(g_i^j)_{0\le i,j\le N}$ such that for all i,j

$$g_i^j - g_{i-1}^j = \delta t$$

and

$$g_i^j - g_i^{j-1} = \delta r,$$

where $N\delta t = T$ and δr is chosen such that $J_0/(\delta r) = P$ is an integer, with M *sufficiently* larger than P. We can then rewrite equation (8) using some finite difference approximation instead of the derivatives, and try to solve it backward in time, starting from the boundary condition

$$g_N^j = 1.$$

The algorithm can be written as follows:

1. Take an arbitrary boundary condition at $r = 0$;
2. Use a finite difference method to find g on the nodes of our mesh using the lower boundary condition given in Step 1 at $r = 0$, as well as the usual boundary condition at $t = T$;
3. Use the values found at Step 2 for $g(., J_0)$ to find h using the formula (12);
4. Pick $h(., 0)$ as our new lower boundary condition at $r = 0$;
5. Go back to Step 2.

We tried to use an explicit finite difference method, but it almost always leads to explosion. Using an implicit or a Crank–Nicholson finite difference scheme, we observed convergence of the algorithm with reasonably small δt and δr.

5 Obtained Yield Curves

In this section we tried to solve for the bond price on the whole domain using the algorithm of the previous section. We chose the model parameters close to the ones found by Gorovoi and Linetsky (2004) when fitting their model of interest rates as option. Namely, we set $\theta = 0.03$, $\sigma = 0.02$ and $\kappa = 0.2$. We computed the bond prices over a horizon of 30 years, using a time-step δt of 1 month (360 time-steps). We chose a unique jump size J_0 of 0.005 which seems realistic when we look at the recent central bank's interest rate policy. However the size of the jump proved to have a considerable impact on the long-term asymptotic value of the yield, with smaller jump sizes bringing lower long-term yields, as we might expect.[10] We used a rate-step δr of 0.001 (hence $J_0 = 5\delta r$). It seems optimal to use a jump size that is a multiple of the rate-step, so we do not have to use interpolation to compute the bond price $g(., J_0)$ at the jump size. Finally, concerning the fixed point algorithm, we used initial bond prices $h(., 0)$ so that the yield curve was a straight line from 0% to 1% initially, and we ran 30 iterations of the algorithm (Steps 2-5).

Figure 5 shows the convergence of the algorithm when we considered a constant $\lambda = 0.5$ (average stay at 0 of 2 years). We can see the successive yield curves that we obtain with an initial short rate of 0. The speed of convergence seems relatively fast, except for longer maturities. In that case, in bond price term, we are far away from the *exact* final boundary condition of $g(T, .) = 1$ and it takes time for the information to come back

[10]The smaller the jump size, the faster we expect the short rate to go back to zero after going out of it. As a result, the quantity J_0 may affect the price of interest rate derivatives. See Rinaz (2006) for a detailed discussion of the impact of each model parameter on bond prices.

to the line $t = 0$. This means that the longer-term yields take more time to converge than shorter-term yields. We always observed convergence starting from different initial conditions, but a judicious choice of initial condition, not too far from the exact yield curve, is essential to get closer to the solution in fewer steps.

The various shapes of the yield curves for different initial short rates are presented in Figure 6. Figure 7 is the solution for the yield curve on the half-plane of positive r. Note that, in the exponential case, no matter how we choose the value of λ, the obtained yield curve at zero is always concave, with no prolonged period of zero yield. This does not agree with the observation of the current Japanese yield curve, which is S-shaped with all yields for about the next 2 years equal to zero. This inadequacy of our model is due to the shape of the probability function of the exponential distribution, which is strictly positive at $t = 0+$ so that the probability that we get a jump immediately is always strictly positive. To reconcile our model with the observed yield curves, we need to consider other distributions for the interarrival times of jumps.

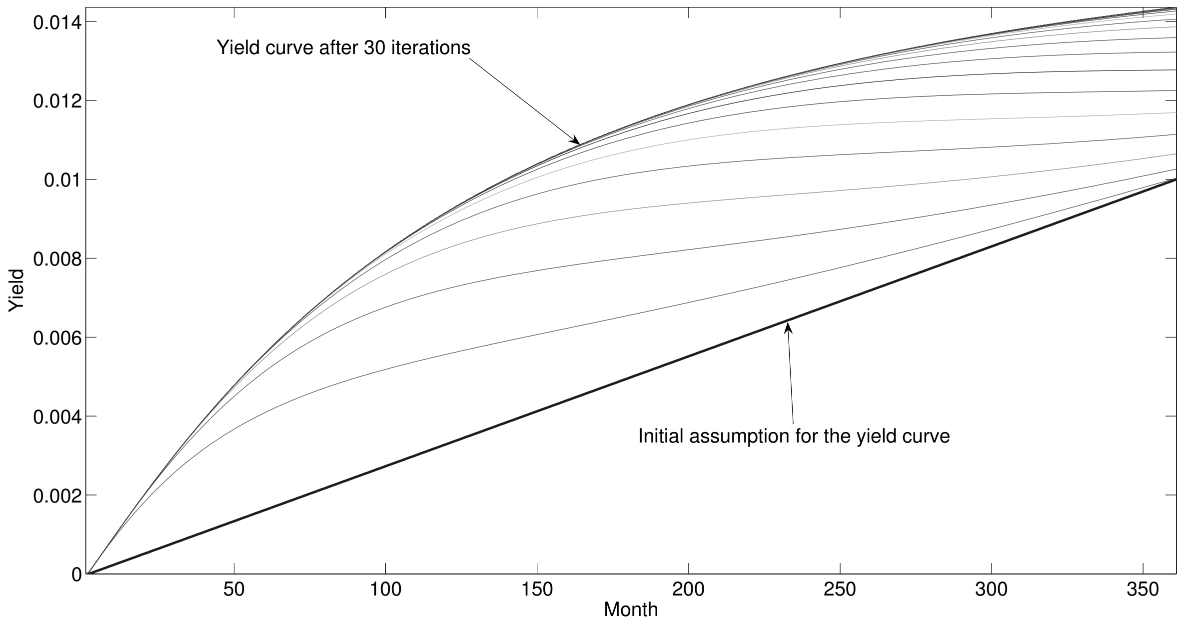

Figure 5: Successive yield curves starting from a current short rate of zero (and $Z = 0$) and obtained through the fixed point algorithm when the interarrival times of jumps are exponentially distributed ($\lambda = 0.5$) (the solution is the curve on the top)

The results in the case of a linear hazard rate ($\alpha = 0$ and $\gamma = 2$) are presented in Figures 8 and 9. We chose a small value, namely 0.025, for coefficient β, so that after one year spent in zero, λ is still 0.05 (ten times smaller than the constant λ of the first case). In this case, the probability density function for the interarrival times of jumps starts from zero and remains small for some time, so we do observe a prolonged zero-yield

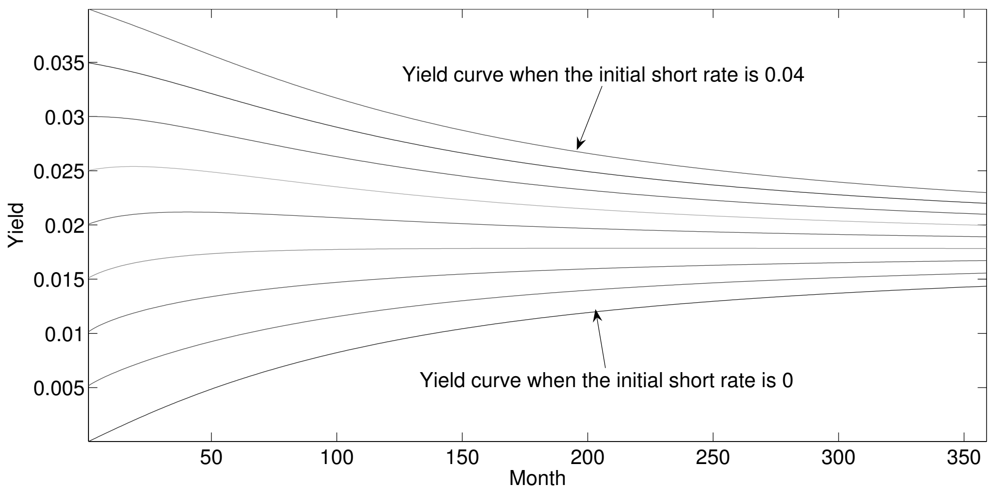

Figure 6: Yield curves obtained after 30 iterations of the algorithm for different initial values of r when the interarrival times of jumps are exponentially distributed ($\lambda = 0.5$)

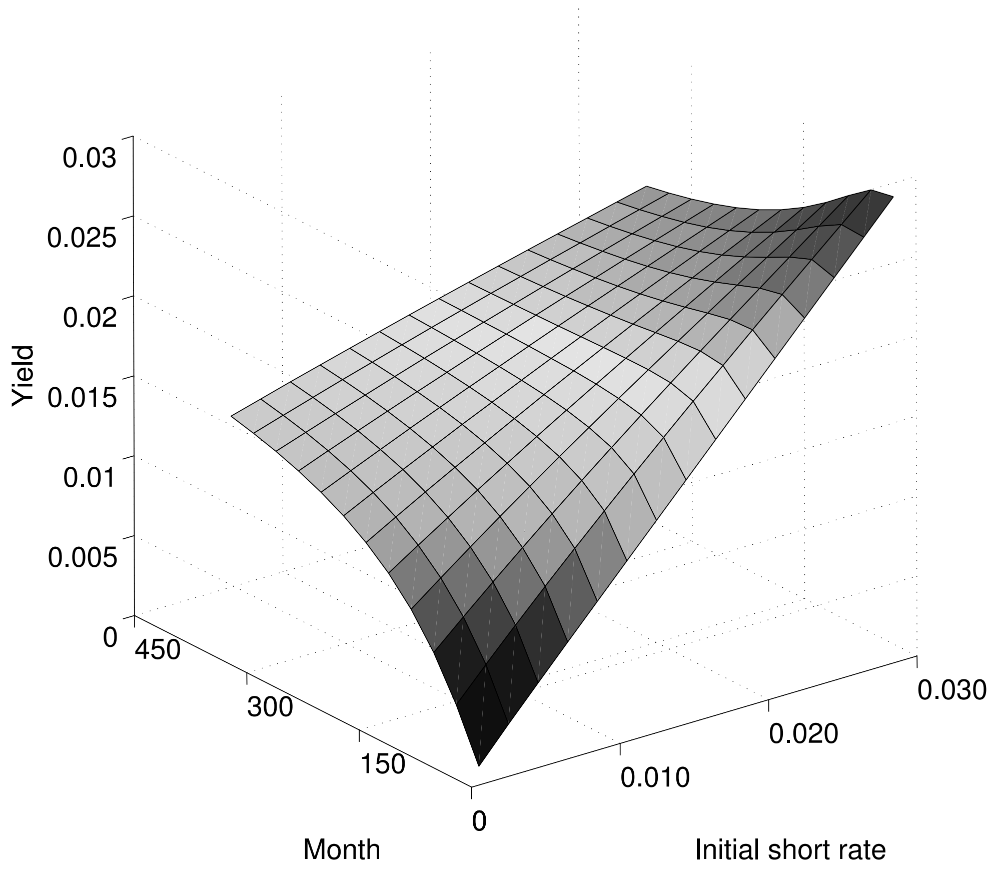

Figure 7: Surface formed by the yield curves for r between 0 and 3% when the jumps are exponentially distributed ($\lambda = 0.5$)

period of about a year (see Figure 8) and the curve is S-shaped (convex, then concave), with an inflexion point around 100 months. The long-term asymptotic value of the yield is still quite high (1% compared with an asymptotic value of 2% when $\lambda = 0.5$). The convergence seems faster than that in the exponential case, but maybe it is due to the different shape of the yield curve.

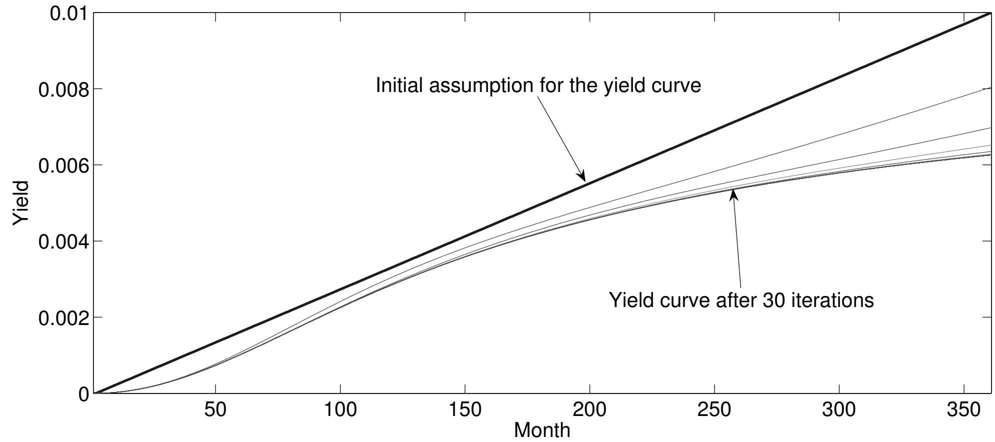

Figure 8: Successive yield curves starting from a current short rate of zero (and $Z = 0$) and obtained through the fixed point algorithm when λ is linear: $\lambda(z) = 0.05z$ (the solution is the curve at the bottom)

When we look at the solution on the half-plane of positive Z (i.e. when the ZIRP has been in place for some time), we see that the yield curve tends to evolve into the shape obtained with a constant λ as Z increases. It raises to make first the zero-yield period disappear, then the inflexion point gets closer to zero and eventually disappears too (see the dashed curves in Figure 9)

Using an affine λ (namely, $\lambda(z) = 0.5 + 0.05z$) we get an S-shaped yield curve, but no initial period of zero-yield. This is due to the constant part of λ which makes strictly positive the probability that r_t leaves zero immediately after reaching it. If we want to be able to observe yield curves with all initial term yields equal to zero, we must not consider hazard rate with constant terms. For $\gamma \geq 2$ and $\alpha = 0$, we obtained yield curve shapes similar to the linear case, only with a more pronounced S-shape. Note that we may use any (positive) hazard rate we want, plug it in formula (4) and obtain the yield curve through the same algorithm.

6 Extension to a Stochastic Latency

Since Section 3, we have assumed that the latency Z increased linearly with time, and that the time already spent in zero was the only determinant of the distribution of jump

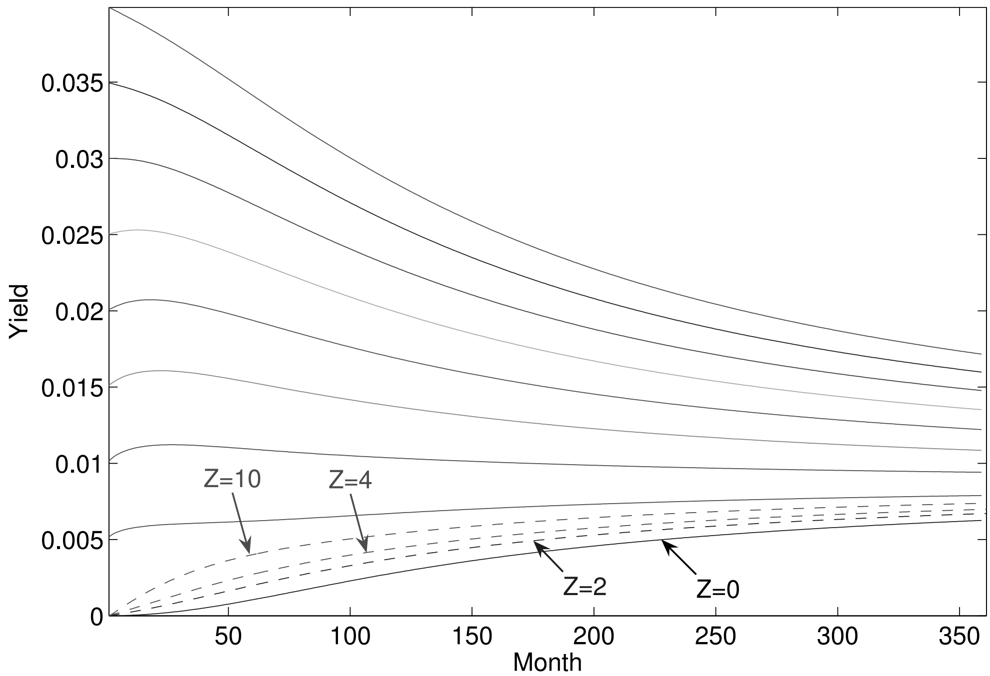

Figure 9: Yield curves obtained after 30 iterations of the algorithm for different initial values of r (solid curves) and Z (dashed curves) when $\lambda(z) = 0.05z$

times. Nevertheless this implies that the yield curve evolves somewhat linearly, in only one way, during the zero interest rate period, as seen in Figure 9; yields for all terms are always increasing for the whole duration of the ZIRP. Dynamically, this is not realistic; even when the short-rate is sleeping at zero, the state of the economy is still changing and information is still flowing in, thus the *prospect* of a return to strictly positive short rates may become alternately imminent or distant, term yields may increase or decrease.

To model this economic uncertainty, we will now consider that the latency has a diffusion component. Namely, we now consider that the dynamics of Z_t are of the form

$$dZ_t = 1_{\{r_{t-}=0\}}\left[\mu_Z dt + \sigma_Z dW_t - Z_{t-}\frac{dJ_t}{|dJ_t|}\right], \quad Z_0 = \bar{Z},$$

where μ_Z and σ_Z are constants and W_t is the same process as the one in the dynamics of r_t. To make the distinction easier, in the rest of the paper we will call the process Z_t *stochastic latency* under these new modelling assumptions and *linear latency* under our initial assumptions, although the latency was of course never fully deterministic.

Since Z can now become negative, we shall consider a hazard rate λ of the form given in formula (6), but with the additionnal restriction that $(\gamma - 1)$ is an even integer, so that $\lambda(Z)$ is always positive.

Now the dynamics of the bond price become

$$df^T = \alpha_T(t, r, Z)dt + \sigma_T(t, r, Z)dW_t + \int_0^\infty \gamma_T(t, r, x, Z)\mu(dx, dt),$$

where

$$\begin{aligned}
\alpha_T(t,r,Z) &= f_t^T + 1_{\{r>0\}}[\kappa(\theta-r)f_r^T + \frac{1}{2}\sigma^2 f_{rr}^T] + 1_{\{r=0\}}[\mu_Z f_Z^T + \frac{1}{2}\sigma_Z^2 f_{ZZ}^T], \\
\sigma_T(t,r,Z) &= 1_{\{r>0\}}\sigma f_r^T + 1_{\{r=0\}}\sigma_Z f_Z^T, \\
\gamma_T(t,r,Z,x) &= 1_{\{r=0\}}[f^T(t,x,0) - f^T(t,0,Z)].
\end{aligned}$$

The new f_{ZZ}^T-term correspond to the diffusion component of Z_t. Its presence prevents us from using the method of characteristics.

The PDE to solve on the r-halfplane is the same as in (7), but the PDE on the Z-plane is now

$$\begin{cases} h_t + \mu_Z h_Z + \frac{1}{2}\sigma_Z^2 f_{ZZ}^T - \lambda(Z)h + \lambda(Z)g(t,J_0) = 0, \\ h(T,Z) = 1 \quad \forall Z. \end{cases} \tag{13}$$

This time we need to use a numerical method to solve (13) as well as (7).

Using the same fixed point algorithm as in Section 4, only replacing Step 3 by a finite difference method to find h, we observed convergence for all parameters values we tried. In all the graphs reported here, we used a time horizon of 30 years, $\theta = 0.05$, $\sigma = 0.02$, $\kappa = 0.2$, $J_0 = 0.005$, $\delta r = 0.0005$, $\delta t = 30/500$ and $\delta Z = \delta t$.

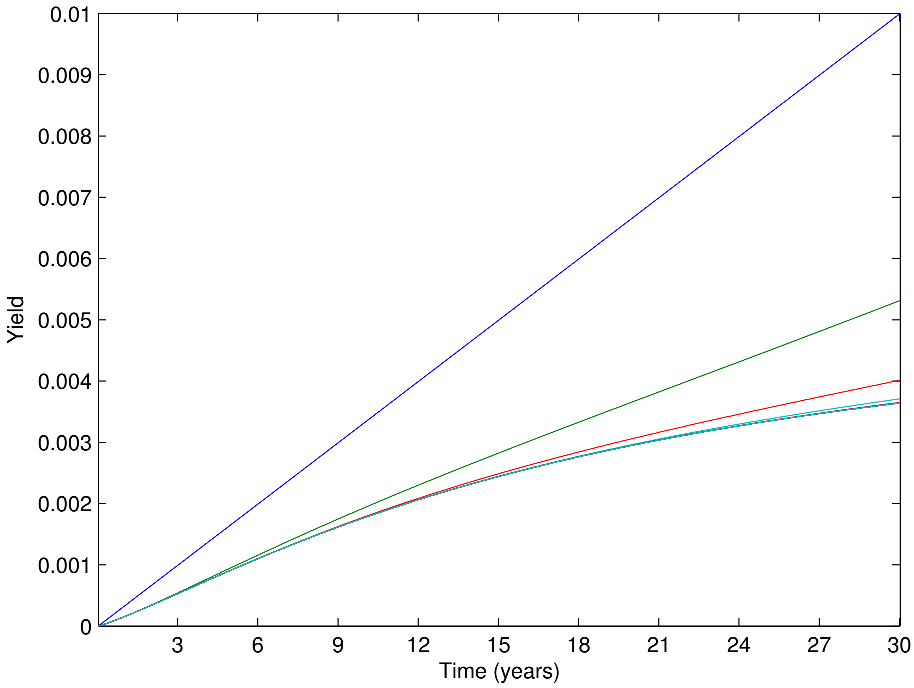

Figure 10: Convergence of the fixed point algorithm. Successive yields at $(r,Z) = (0,0)$ for 10 iterations of the algorithm when $\lambda(z) = 0.05 + z^2$ and $(\mu_Z, \sigma_Z) = (0, 0.002)$.

Figure 10 shows that the convergence of the yield curve at the intersection of the two planes is quite as fast as before, when the latency was simply increasing linearly

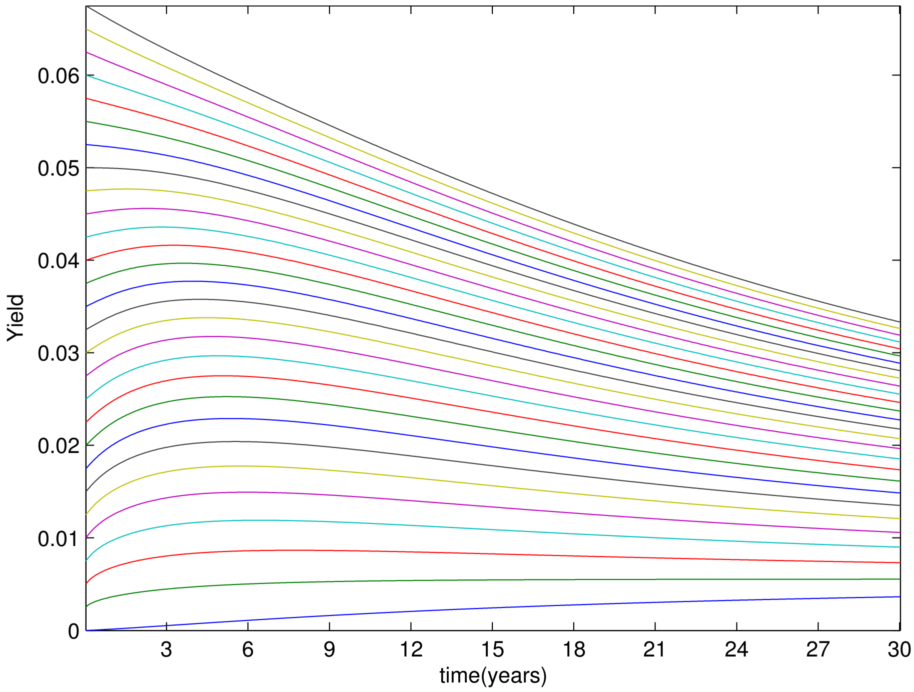

Figure 11: Different yield curves obtained for different initial values of r in the case when $\lambda(z) = 0.05 + z^2$ and $(\mu_Z, \sigma_Z) = (0, 0.002)$.

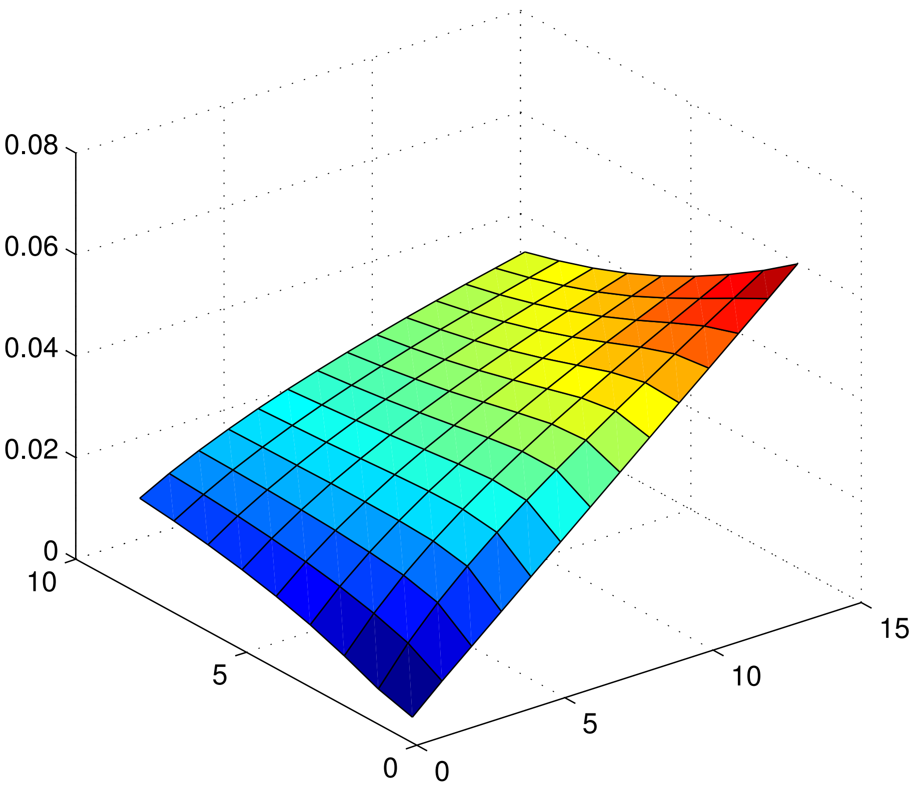

Figure 12: Yield curve surface in the r half-plane in the case when $\lambda(z) = 0.05z^2$ and $(\mu_Z, \sigma_Z) = (0, 2)$.

with time and we had an exact solution on the Z-plane. Nevertheless, the computation time to complete one iteration of the algorithm is almost doubled, since we must use the Cranck–Nicholson method twice.

Figures 11 and 12 show that the shapes of the yield curves we obtain for different initial values of r are about the same as in the previous model. Again we see that a constant term in $\lambda(.)$ prevents us from observing a prolonged period of zero yield.

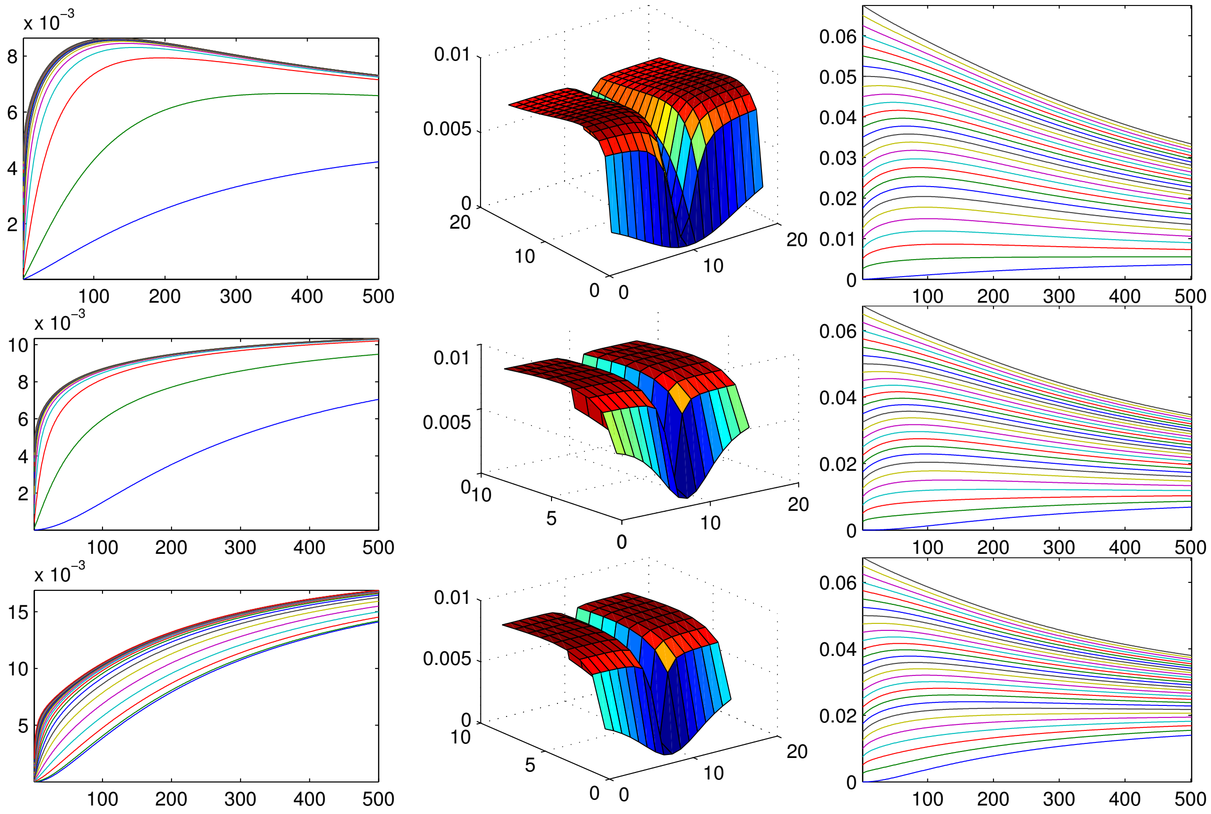

Figure 13: From left to right: yield curves at zero obtained for different initial values of Z, yield curve surface in the Z-plane, yield curves obtained for different initial values of r. The first line of graphs corresponds to the case $\lambda(z) = 0.05 + z^2$ and $(\mu_Z, \sigma_Z) = (0, 0.002)$, the second line to the case $\lambda(z) = z^2$ and $(\mu_Z, \sigma_Z) = (0, 0.2)$ and the last line to the case $\lambda(z) = 0.05z^2$ and $(\mu_Z, \sigma_Z) = (0, 2)$ (all other parameters being the same).

Figure 13 compares the yield curves obtained in the r- and Z-planes for different values of the parameters. We can observe (since $\lambda(.)$ is symmetric) that the yield surface on the Z-plane is symmetric (see the central column of Figure 13). Therefore, without loss of generality, we can only consider yield curves obtained for positive values of Z (see the first column of Figure 13). As we might expect, when Z increases, the corresponding yield curves converge to the yield curve starting at $(r, Z) = (J_0, 0)$ (here $J_0 = 0.005$) since for large values of Z an immediate jump occurs for sure.

7 Example of calibration

In this section we are interested in the calibration of the "linear" latency model to bond prices on the market. To simplify the notations, we regroup all the parameters into a vector-valued parameter $Y \in \mathbb{R}^6$ and we assume that it is of the form

$$Y = (\theta, \sigma, \kappa, J_0, \beta, \gamma),$$

where β and γ are the parameters in the hazard rate λ, namely $\lambda(z) = \beta\gamma z^{(\gamma-1)}$.

We are interested in solving the optimization problem

$$\min_{Y \in \mathbb{R}^6} \sum_{i \in I} (P(r, Z, T_i, Y) - \bar{P}(r, Z, T_i))^2, \text{ subject to } l_y \leq Y \leq u_y, \tag{14}$$

where $P(r, Z, T_i, Y)$ is the price of the zero-coupon bond maturing at T_i computed using our model, and where the parameters are set to level Y, the terms r and Z denote the current short rate and latency respectively, and $\bar{P}(r, Z, T_i)$ is the price of the zero-coupon bond maturing at T_i observed on the market.

We used zero-coupon bond prices in the objective function, but traded JGBs are actually coupon-bearing bonds. A bootstrap method is used to derive zero-coupon bond prices from coupon-bond prices, or more frequently from a set of deposit rates and swap rates. We estimated the model parameters using such bootstrapped prices of 20 zero-coupon bonds maturing between 3 months and 10 years.

In this numerical example we used a trust region reflexive Newton algorithm to compute the solution to Problem (14) . This algorithm is based on Coleman and Li (1996) and is already implemented in Matlab as the function `lsqnonlin`. We chose a constant jump size J_0 of 50 basis points, and fitted only the 5 remaining parameters in Y.

Table 5 reports the final parameters estimates that we obtained using different starting points. The trust region algorithm requires the computation of the Jacobian of the objective function, so that estimating 5 parameters requires 6 function evaluations at every iteration of the algorithm (in the case of forward differences). As we can see, the final estimates may vary depending on the starting points, but β is always estimated as very small and γ as close to 2. This is due to the S-shape of the yield curve at that date.

Naturally, the computational time may vary greatly depending on the number of iterations required to reach the termination criteria, the number of points used in the finite difference algorithm, the plateform used for computation and the programming language of implementation.

Figure (14) displays the yields of the 40 bonds that were used as well as the yield curve implied by the model after calibration.

Parameter	starting point	final estimates	starting point	final estimates	starting point	final estimates
θ	0.5	0.0581	0.03	0.1198	0.03	0.0479
σ	0.02	0.0156	0.2	0.0137	0.02	0.0008
κ	0.2	0.2400	0.2	0.0563	0.2	0.1126
β	0.05	0.0868	0.05	0.0832	1	0.0821
γ	2	2.0770	2	2.7884	1	2.3706
# of Iterations		6		19		33
Residuals		7.0158e-6		3.0696e-5		9.6245e-6

Table 5: Parameter estimation on May 2, 2001, using 40 zero-coupon bonds maturing between 3 months and 10 years.

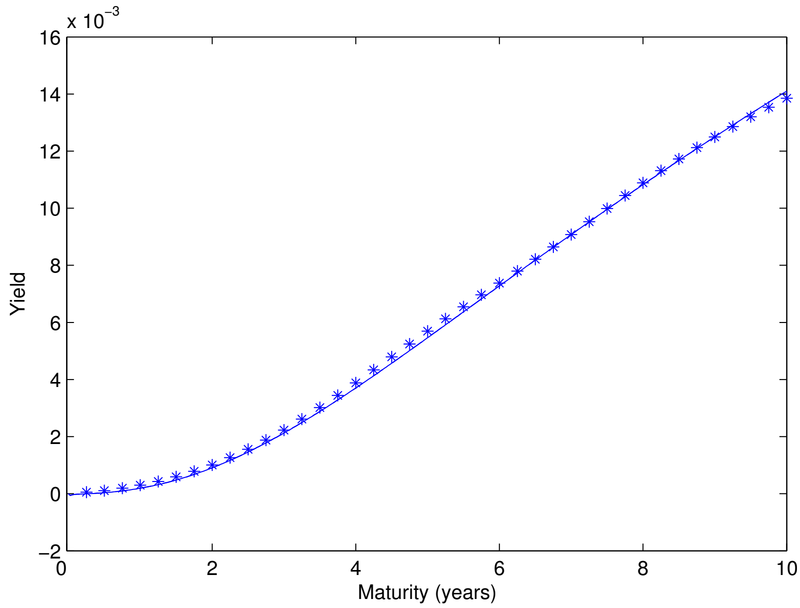

Figure 14: Fitted yield in the first case of Table 5. Solid lines: fitted bond prices and associated yields in the first case of Table 5. The "star" marks represent the yields derived from discount factors. The time-step used in the finite difference method is .025, so we have ten time steps between each bond maturity date.

8 Conclusion

We have introduced a new class of positive interest rate models and proposed a stable numerical method to compute the bond price when the intensity of the compensator of the jump interarrival times λ_t is a function of the latency Z_t. First, Z_t was considered to be the time during which the short rate has been sitting at 0 (*linear latency*); then it was considered to be an unobservable random process, characteristic of the latency or imminence of the jump (*stochastic latency*).

Statistically, the yield curves obtained in the case of linear λ are satisfactory in the sense that they exhibit sustained periods of zero yield (zero forward spot rate) as the Japanese market does at the moment but more complex forms of the intensity λ can be considered as well.

Dynamically, we obtain better result with a stochastic latency, but have to pay a price for it in terms of increased computational time.

The model calibration to current market prices seems a bit involved at the moment, because we do not have a closed-form solution for the bond price, and is left for further research. Qualitatively, we can say that the parameters having a strong influence on the long-term yield are θ, J_0 and λ, while the only parameter having a noticeable influence on the shape of the curve is λ, which gives us a hope that we can fit well the model to any current yield curve of the shapes shown in Figures 6 and 9.

References

Aït-Sahalia, Y.: 1996, Testing continuous-time models of the spot interest rate, *Review of Financial Studies* **9**, 385–426.

Bansal, R. and Zhou, H.: 2002, Term structure of interest rates with regime shifts, *Journal of Finance* **57**, 1997–2043.

Björk, T.: 1998, *Arbitrage Theory in Continuous Time*, Oxford University Press, Oxford, New York.

Black, F.: 1995, Interest rates as options, *Journal of Finance* **50**(7), 1371–1376.

Black, F. and Karasinski, P.: 1991, Bond and option pricing when short rates are lognormal, *Financial Analysts Journal* **47**, 52–59.

Brace, A., Gatarek, D. and Musiela, M.: 1997, The market model of interest rate dynamics, *Mathematical Finance* **7**, 127–155.

Brigo, D. and Mercurio, F.: 2001, *Interest Rate Models - Theory and Practice*, Springer-Verlag, Berlin.

Cairns, A. J.: 2004, A family of term-structure models for long-term risk management and derivative pricing, *Mathematical Finance* **14**(3), 415–444.

Coleman, T. F. and Li, Y.: 1996, An interior trust region approach for nonlinear minimization subject to bounds, *SIAM Journal on Optimization* **6**, 418–445.

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

Cox, J., Ingersoll, J. and Ross, S.: 1985, A theory of the term structure of interest rates, *Econometrica* **53**, 385–407.

Dothan, L.: 1978, On the term structure of interest rates, *Journal of Financial Economics* **6**, 59–69.

Duffie, D.: 2001, *Dynamic Asset Pricing Theory - Third Edition*, Princeton University Press, Princeton and Oxford.

Flesaker, B. and Hughston, L.: 1996, Positive interest, *Risk* **9**(1), 46–49.

Goldstein, R. and Keirstead, W.: 1997, On the term structure of interest rates in the presence of reflecting and absorbing boundaries. Working paper, Ohio State University.

Gorovoi, V. and Linetsky, V.: 2004, Black's model of interest rates as options, eigenfunction expansions and Japanese interest rate, *Mathematical Finance* **14**(1), 49–78.

Gray, S. F.: 1996, Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics* **42**, 27–62.

Hansen, A. T. and Poulsen, R.: 2000, A simple regime switching term structure model, *Finance and Stochastics* **4**(4), 409–429.

Jones, C. S.: 2003, Nonlinear mean reversion in the short-term interest rate, *Review of Financial Studies* **16**(3), 793–843.

Karlin, S. and Taylor, H. M.: 1981, *A Second Course in Stochastic Processes*, Academic Press, San Diego.

Last, G. and Brandt, A.: 1995, *Marked Point Processes on the Real Line - The Dynamic Approach*, Springer-Verlag, Berlin.

Marumo, K., Nakayama, T., Nishioka, S. and Yoshida, T.: 2003, Extracting market expectations on the duration of the zero interest rate policy from Japan's bond prices. Working Paper. Bank of Japan, Financial Markets Department.

Miyahara, Y. and Fujiwara, T.: 2003, The minimal entropy martingale measures for geometric Lévy processes, *Finance and Stochastics* **7**, 509–531.

Rinaz, S.: 2006, *Positive Interest Rate Models with Sticky Barrier for the Japanese Market*, PhD thesis, Kyoto University.

Rogers, L.: 1996, Gaussian errors, *Risk* **9**, 42–45.

Sun, L.: 2003, Non-linear drift and stochastic volatility: an empirical investigation of short-term interest rate models, *Journal of Financial Research* **26**(3), 389–404.

Vašíček, O.: 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**, 177–188.