

REGIME SWITCHING AND BOND PRICING

Christian GOURIEROUX ⁽¹⁾

Alain MONFORT ⁽²⁾

Fulvio PEGORARO ⁽³⁾

Jean-Paul RENNE⁽⁴⁾

June, 2013

Abstract Regime Switching and Bond Pricing

This paper proposes an overview of the usefulness of the regime switching approach for building various kinds of bond pricing models and of the roles played by the regimes in these models. Both default-free and defaultable bonds are considered. The regimes can be used to capture stochastic drifts and/or volatilities, to represent discrete target rates, to incorporate business cycles or crises, to introduce contagion, to reproduce zero lower bound spells, or to evaluate the impact of standard or non-standard monetary policies. From a technical point of view, we stress the key role of Markov chains, Compound Autoregressive (Car) processes, Regime Switching Car processes and multi-horizon Laplace transforms.

Keywords: Term Structure, Regime Switching, Affine Models, Car Process, Multi-horizon Laplace Transform, Contagion, Default Risk, Monetary Policy.

JEL classification: E43, G12.

JFEC Lecture, SoFiE, Oxford, June 20th, 2012.

¹CREST, Laboratoire de Finance-Assurance [E-mail: gourieroux@ensae.fr], and University of Toronto.

²CREST, Laboratoire de Finance-Assurance [E-mail: monfort@ensae.fr], Banque de France, Financial Economics Research Service [DGEI-DEMFI-RECFIN] and University of Maastricht.

³Banque de France, Financial Economics Research Service [DGEI-DEMFI-RECFIN; E-mail: Fulvio.PEGORARO@banque-france.fr], CREST, Laboratoire de Finance-Assurance [E-mail: pegoraro@ensae.fr] and HEC Lausanne, Department of Finance and Insurance.

⁴Banque de France, Financial Economics Research Service [DGEI-DEMFI-RECFIN; E-mail: Jean-Paul.RENNE@banque-france.fr].

1 INTRODUCTION

Regime switching models have been widely used in Financial Econometrics. The domains of applications include the analysis of stock returns [see e.g. Hamilton, Susmel (1994), Billio, Pelizzon (2000), Ang, Chen (2002)], exchange rates [see e.g. Engel, Hamilton (1990), Bekaert, Hodrick (1993)], asset allocations [see Ang, Bekaert (2002a, 2004), Guidolin, Timmerman (2008), Tu (2010)], electricity prices [Huisman, Mahieu (2003), Mount, Ning, Cai (2005), Monfort, Feron (2012)], or systemic risk [Billio, Getmansky, Lo, Pelizzon (2012)]. See also the survey paper by Ang, Timmerman (2011). However, it is in the modeling of default-free interest rates that the regime switching approach is the most frequent. A first stream of literature does not consider the pricing problem, but shows how the introduction of switching regimes can improve the properties of dynamic models of interest rates in terms of persistence, of fitting and forecasting of the yields or of their unconditional and conditional moments [see e.g. Hamilton (1988), Garcia, Perron (1996), Ang, Beckaert (2002b, 2002c)]. A second stream of literature focuses on the pricing problem and incorporates switching regimes in a simultaneous modeling of the historical dynamics, the risk-neutral dynamics and the stochastic discount factor, in order to evaluate market prices of risks, risk premia or term premia [see e.g. Bansal, Zhou (2002), Dai, Singleton, Yang (2007), Monfort, Pegoraro (2007), Ang, Bekaert, Wei (2008), Chib, Kang (2013)]. In both kinds of literature the switching regimes are latent, that is not observed by the econometrician.

More recently, regime-switching features have been introduced in the modeling of defaultable bond prices [see Monfort, Renne (2011, 2013)] and credit ratings. In the latter case the latent regimes are introduced to account for non-linear changes in the probabilities of credit-rating transition, extending the approach proposed for instance by Jarrow, Lando, Turnbull (1997).

The present paper focuses also on the applications to interest rate models. More precisely, we propose an overview of the usefulness of the regime switching approach for building various bond pricing models and of the roles of the regimes in these models. The objective of the pricing models can be to price default-free or defaultable bonds, or to analyse simultaneously credit ratings and

defaultable bonds prices. The regimes can be used to capture stochastic drifts and/or volatilities, to represent discrete values of a target rate, to incorporate business cycle or crises effects, to introduce contagion effects, to reproduce zero lower bound spells, or to evaluate the impact of standard or non-standard monetary policies. From a technical point of view, we stress the key role of Markov chains, Compound Autoregressive (Car) processes, Regime Switching Car processes and multi-horizon Laplace transforms.

In Section 2 we show that a key tool for pricing both default-free and defaultable bonds in discrete time is the multi-horizon Laplace transform of the underlying risk factors. These Laplace transforms can be computed in closed form for Markov chains and recursively for Regime Switching Compound Autoregressive (Car) processes. In order to justify the non-linear models chosen for the historical dynamics of interest rates, we conclude this section with an empirical exercise on the U.S. bond market. This shows the relevance of Regime-Switching Gaussian VAR(p) models in capturing linear and non-linear serial dependence in interest rates as well as their lack of Gaussianity.

Then we develop regime switching term structure models in various directions. We first consider in Section 3 the pricing of default-free bonds. We carefully distinguish the Regime Switching Term Structure Models (RSTSM), which provide affine formulas for the *yields* as functions of underlying risk factors, and the RSTSM for which the affine formulas are satisfied by the bond *prices*. In the latter case, we discuss the respective properties of models with exogenous and endogenous switching regimes and their ability to generate short rate paths staying at a lower bound. We also discuss the practical implementation of these models, where the bond prices can be easily computed recursively, and sometimes in closed form. We also propose numerical illustrations showing the potentialities of these models for reproducing zero lower bound spells, or for evaluating monetary policies.

In Section 4 we consider pricing models for defaultable bonds. In this framework there exist individual (specific) risk factors as well as common (systematic) risk factors including a global regime indicator. When the stochastic discount factor (s.d.f.) depends on the common factors only,

the causality features between individual and common factors are the same under the historical and risk-neutral distributions. Defaultable-bond pricing is illustrated by an application to sovereign bonds of the Euro-zone countries. A common regime variable is introduced to capture the crisis periods. The approach disentangles credit and liquidity risks incorporated in spreads. Historical and risk-neutral default probabilities are compared.

Section 5 concludes. Proofs and Tables are gathered in Appendices.

2 A TOOLBOX FOR REGIME SWITCHING TERM STRUCTURE MODELS

This section gathers the tools which are useful for the analysis of RSTSM. We first recall the pricing formulas for default-free and defaultable bond pricing and highlight the key role of the multi-horizon Laplace transform of the risk factors. Then, we compute Laplace transforms for Markov chains and for regime switching compound autoregressive processes.

2.1 Bond pricing

Let us adopt a discrete time setting in which the new information of the investors⁵ at date $t, t = 1, 2, \dots$, is a n -dimensional factor w_t . The whole *information* of the investors at date t is therefore $\underline{w}_t = (w'_t, w'_{t-1}, \dots, w'_1)'$. The *historical dynamics* of the factor process $\{w_t\}$ is characterized either by the sequence of conditional probability density functions (p.d.f.) $f^{\mathbb{P}}(w_t|\underline{w}_{t-1})$ (with respect to a dominating measure μ), or by the sequence of conditional Laplace transforms $\varphi_{t-1}^{(w)}(u) = E[\exp(u'w_t)|\underline{w}_{t-1}]$, defined on a convex set containing 0. Let us denote by $p_t[g(\underline{w}_{t+h})]$ the (spot) price at t of an asset providing at $t + h$ the payoff $g(\underline{w}_{t+h})$. Under standard assumptions, including the absence of arbitrage opportunity [see Harrison, Kreps (1979), Hansen, Richard (1987), Bertholon, Monfort, Pegoraro (2008)], there exists a sequence of positive random variables

⁵We focus more on bond pricing than on the estimation of the dynamic term structure models, which depends on the information available to the econometrician. This information can be different from the information of the investor.

$M_{t-1,t} = M_{t-1,t}(\underline{w}_t)$, called *stochastic discount factors* (s.d.f.), such that :

$$p_t[g(\underline{w}_{t+h})] = E[M_{t,t+1}(\underline{w}_{t+1}) \dots M_{t+h-1,t+h}(\underline{w}_{t+h})g(\underline{w}_{t+h})|\underline{w}_t]. \quad (2.1)$$

In particular the price at date t of a default-free zero-coupon bond with residual maturity h , delivering the unitary payoff at $t + h$, is :

$$B(t, h) = E_t(M_{t,t+1} \dots M_{t+h-1,t+h}).$$

The default-free *yield to maturity* h is :

$$R(t, h) = -\frac{1}{h} \log[B(t, h)].$$

For $h = 1$, we get the *short rate* $r_t = R(t, 1)$, defined by :

$$r_t = -\log[E_t(M_{t,t+1})] \iff E_t(M_{t,t+1}) = \exp(-r_t). \quad (2.2)$$

The *risk-neutral* (R.N.) dynamics of $\{w_t\}$ is defined by the sequence of conditional distributions of w_t given \underline{w}_{t-1} , whose p.d.f. with respect to the corresponding historical distribution is $M_{t-1,t} \exp(r_{t-1})$. In other words, the conditional R.N. p.d.f. of w_t given \underline{w}_{t-1} with respect to a dominating measure μ is :

$$f^{\mathbb{Q}}(w_t|\underline{w}_{t-1}) = f^{\mathbb{P}}(w_t|\underline{w}_{t-1}) M_{t-1,t} \exp(r_{t-1}), \quad (2.3)$$

This equality is equivalent to :

$$M_{t-1,t} = \frac{f^{\mathbb{Q}}(w_t|\underline{w}_{t-1})}{f^{\mathbb{P}}(w_t|\underline{w}_{t-1})} \exp(-r_{t-1}), \quad (2.4)$$

and implies :

$$E_{t-1}^{\mathbb{Q}}(M_{t-1,t}^{-1}) = \exp(r_{t-1}), \quad (2.5)$$

which is the R.N. analogue of equation (2.2) at date $t - 1$.

Thus, the pricing formula (2.1) can be also written as :

$$p_t[g(\underline{w}_{t+h})] = E_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{t+h-1}) g(\underline{w}_{t+h})], \quad (2.6)$$

and therefore the spot price is the conditional expectation under the R.N. distribution of the discounted cash-flow $g(\underline{w}_{t+h})$.

Formula (2.6) can be used to derive an alternative expression of the (spot) price of the zero-coupon bond :

$$B(t, h) = E_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{t+h-1})]. \quad (2.7)$$

When the short rate is an affine function of risk factors w_t :

$$r_t = \beta_0 + \beta_1' w_t,$$

the zero-coupon price $B(t, h)$ becomes :

$$B(t, h) = \exp(-\beta_0 h - \beta_1' w_t) E_t^{\mathbb{Q}}[\exp(-\beta_1' w_{t+1} - \dots - \beta_1' w_{t+h-1})]. \quad (2.8)$$

For a defaultable zero-coupon bond, with residual maturity h , the payoff at $t + h$ is 1, if the issuing entity n has not defaulted, and 0, otherwise, when the recovery rate is zero. The price of the defaultable bond is (see Section 4) :

$$\begin{aligned} B_n(t, h) &= E_t^{\mathbb{Q}}[\exp(-r_t \dots - r_{t+h-1} - \lambda_{n,t+1}^{\mathbb{Q}} \dots - \lambda_{n,t+h}^{\mathbb{Q}})] \\ &= \exp[-h(\beta_0 + \alpha_{0n}) - \beta_1' w_t] \\ &\quad \times E_t^{\mathbb{Q}}\{\exp[-(\beta_1 + \alpha_{1,n})' w_{t+1} - \dots - (\beta_1 + \alpha'_{1,n}) w_{t+h-1} - \alpha'_{1,n} w_{t+h}]\}, \end{aligned} \quad (2.9)$$

where $\lambda_{n,t}^{\mathbb{Q}} = \alpha_{0,n} + \alpha'_{1,n} w_t$ denotes the R.N. default intensity.

These bond pricing formulas highlight the role of the conditional Laplace transforms of the risk-factors. More precisely, throughout the paper, we will have to compute for any date t , and given sequences $(\gamma_1^{(h)}, \dots, \gamma_h^{(h)})$, $h \in \{1, \dots, H\}$, the multi-horizon conditional Laplace transforms :

$$\varphi_{t,h}^{(w)} \left(\gamma_1^{(h)}, \dots, \gamma_h^{(h)} \right) := E_t[\exp(\gamma_1^{(h)'} w_{t+1} + \dots + \gamma_h^{(h)'} w_{t+h})]$$

in an efficient way. Note that, in formulas (2.8) and (2.9) the sequences $(\gamma_1^{(h)}, \dots, \gamma_h^{(h)})$, $h \in \{1, \dots, H\}$, have a *reverse order* structure in which, for any $h \in \{1, \dots, H\}$, we have :

$$\gamma_h^{(h)} = \delta_1, \gamma_{h-1}^{(h)} = \delta_2, \dots, \gamma_1^{(h)} = \delta_h,$$

for a given sequence $\delta_1, \dots, \delta_H$. In formula (2.8) we have $\delta_h = -\beta_1, \forall h$, whereas in formula (2.9) we have $\delta_1 = -\alpha_{1,n}$ and $\delta_h = -\beta_1 - \alpha_{1,n}, \forall h \geq 2$.

2.2 Markov chains

Switching regimes are usually represented by Markov chains. When there are J regimes, we can define a Markov chain process $\{z_t\}$ whose component $z_{j,t}$, for any $j \in \{1, \dots, J\}$, is the indicator function of regime j . In other words, z_t is valued in $\{e_1, \dots, e_J\}$, where e_j is the J -dimensional vector, whose components are all equal to zero except the j^{th} one which is equal to one. The dynamics of $\{z_t\}$ is characterized by its transition matrices Π_t , whose entries $\pi_{i,j,t}$ are defined by :

$$\pi_{i,j,t} = \mathbb{P}(z_t = e_j | z_{t-1} = e_i).$$

These probabilities may depend on time in a deterministic way, in order to incorporate exogenous variables or seasonal dummies. The conditional distribution of z_t given z_{t-1} can also be

characterized by its *conditional Laplace transform* :

$$\begin{aligned}\varphi_{t-1}^{(z)}(u) &= E[\exp(u'z_t)|z_{t-1}]. \\ &= \left[\sum_{j=1}^J \pi_{1,j,t} \exp(u_j), \dots, \sum_{j=1}^J \pi_{J,j,t} \exp(u_j) \right] z_{t-1} \\ &= e' P'_t(u) z_{t-1}\end{aligned}$$

where $P_t(u) = \Pi_t \text{diag}[\exp(u)]$, $\text{diag}[\exp(u)]$ being the diagonal matrix with components the exponential of the components of u and $e' = (1, \dots, 1)$. The conditional Laplace transform can alternatively be written as an exponential function of z_{t-1} :

$$\varphi_{t-1}^{(z)}(u) = \exp\left\{ \left[\log \sum_{j=1}^J \pi_{1,j,t} \exp(u_j), \dots, \log \sum_{j=1}^J \pi_{J,j,t} \exp(u_j) \right] z_{t-1} \right\}. \quad (2.10)$$

Moreover, this multi-horizon conditional Laplace transform has a closed form:

$$\varphi_{t,h}^{(z)}(\gamma_1^{(h)}, \dots, \gamma_h^{(h)}) = E_t \left[\exp \left(\gamma_1^{(h)'} z_{t+1} + \dots + \gamma_h^{(h)'} z_{t+h} \right) \right].$$

Proposition 1. *The multi-horizon conditional Laplace transform of a Markov chain is:*

$$\varphi_{t,h}^{(z)} = e' P'_{t+h}(\gamma_h^{(h)}) \dots P'_{t+1}(\gamma_1^{(h)}) z_t,$$

where $P_t(\gamma) = \Pi_t \text{diag}[\exp(\gamma)]$, $e' = (1, \dots, 1)$ and $\text{diag}[\exp(\gamma)]$ denotes the diagonal matrix with diagonal terms the exponential of the components of γ .

Proof : see Appendix 1.

According to Proposition 1, the multi-horizon conditional Laplace transform is of the form $\alpha' z_t$ (the component α_j of α being positive), i.e. linear in z_t ; it can also be considered as an exponential linear function of z_t , since $\alpha' z_t = \exp(\beta' z_t)$, where the components of β are $\beta_j = \log(\alpha_j)$. This remark will be useful for combining Markov chains with the Car processes considered in Section 2.3.

2.3 Regime Switching Car process

The usefulness of Car processes, or discrete-time affine processes, introduced by Darolles, Gouriéroux, Jasiak (2006) is now well documented [see for instance, Gouriéroux, Monfort (2007), Gouriéroux, Monfort, Polimenis (2006), Monfort, Pegoraro (2007), Le, Singleton, Dai (2010), Monfort, Renne (2011, 2013)]. A Car process of order one, Car (1), is defined as follows:

Definition 1. *A n -dimensional process $\{w_t\}$ is Car(1) if its conditional log-Laplace transform given the past $\underline{w}_{t-1} = (w'_{t-1}, \dots, w'_1)$, is affine in w_{t-1} , that is, of the form :*

$$\begin{aligned}\psi_{t-1}^{(w)}(u) &= \log E[\exp(u'w_t)|\underline{w}_{t-1}] \\ &= a_{t-1}(u)'w_{t-1} + b_{t-1}(u),\end{aligned}$$

where a_{t-1} and b_{t-1} may depend on time in a deterministic way.

It is known that we can also define Car processes of order p [Car(p)] and that, by extending the dimension of the process, a Car(p) process is also a Car(1) process. Therefore we only consider Car(1) processes in the next sections. It is also known that the family of Car(1) processes contains many important processes like autoregressive Gaussian processes, autoregressive Gamma processes, compound Poisson processes and autoregressive Wishart processes. Equation (2.11) shows that a Markov chain is Car(1). Let us now introduce new stochastic processes, namely the Regime Switching Car(1) processes [RSCar(1)] defined in the following way:

Definition 2. *Let us consider :*

i) a baseline family of Car(1) conditional log-Laplace transforms of the form :

$$a_{t-1}(u)' \tilde{w}_{t-1} + b_{t-1}^{(0)}(u)' \delta,$$

where δ is a K -dimensional vector and $b_{t-1}^{(0)}$ a K -dimensional vector of functions;

ii) a J -regime exogenous Markov chain $\{z_t\}$ with transition matrices Π_t ;

iii) a set of independent random K -dimensional vectors $\Delta_{i,t}^j, i \in \{1, \dots, J\}, j \in \{1, \dots, J\}$, identically distributed over time.

The stochastic process $\{y_t\}$ such that the conditional log-Laplace transform of y_t given $y_{t-1}, z_t = e_j, z_{t-1} = e_i, \Delta_{i,t}^j = \delta_{i,t}^j$ is given by :

$$a_{t-1}(u)'y_{t-1} + b_{t-1}^{(0)}(u)'\delta_{i,t}^j,$$

is called a *RSCar(1)*.

Regime Switching Car(1) processes are similar to diffusion models with jumps encountered in continuous time models. The baseline dynamics corresponds to the baseline diffusion equation and this diffusion equation involves several parameters which can switch. The underlying Markov chain defines the times of the jumps on the different parameters and the components of $\Delta_{i,t}^j$ define the stochastic sizes of the jumps.

Example: *Gaussian autoregressive process with switching drift and volatility.*

Let us consider a baseline Gaussian VAR(1) dynamics:

$$\tilde{w}_t = \mu + \Phi \tilde{w}_{t-1} + \varepsilon_t, \text{ where } \varepsilon_t \sim IIN(0, \Omega).$$

We have:

$$a_{t-1}(u) = \Phi'u, b_{t-1}(u) = u'\mu + \frac{1}{2}u'\Omega u = b_{t-1}^{(0)}(u)'\delta,$$

with:

$$b_{t-1}^{(0)}(u) = \left[u, \frac{1}{2} \text{vec}(uu') \right]', \quad \delta = [\mu, \text{vec}(\Omega)]'.$$

Therefore we can introduce regime switching drift and volatility parameters.

A RSCar(1) process $\{y_t\}$ is not Car(1), but the extended process $w_t = \{y_t', z_t'\}'$ is Car(1). Indeed, we have the following property :

Proposition 2. *The process $w_t = (y_t', z_t')$, where $\{z_t\}$ is a Markov chain and $\{y_t\}$ an associated $RSCar(1)$, is a $Car(1)$ process; its conditional log-Laplace transform is given by :*

$$\log E_{t-1}[\exp(u'y_t + v'z_t)] = a'_{t-1}(u)y_{t-1} + [A_1(u, v), \dots, A_J(u, v)]z_{t-1},$$

with :

$$A_i(u, v) = \log \sum_{j=1}^J \{\pi_{i,j,t} \exp[\psi_{i,j}(b_{t-1}^{(0)}(u)) + v_j]\},$$

$\psi_{i,j}(\cdot)$ being the log-Laplace transform of $\Delta_{i,t}^j$.

If the size of the jumps $\Delta_{i,t}^j$ is non random we have $\psi_{i,j}(b_{t-1}^{(0)}(u)) = b_{t-1}^{(0)}(u)' \Delta_{i,t}^j$.

Proof: see Appendix 2.

As stressed in subsection 2.1, an important issue is the computation of multi-horizon conditional Laplace transforms of factor process $\{w_t\}$. The importance of Laplace transforms has also been stressed in continuous time models (see, for instance, Duffie, Pan and Singleton (2000)). The following result shows that if $\{w_t\}$ is $Car(1)$ or, according to Proposition 2, $RSCar(1)$, these computations are easily done recursively.

Proposition 3. *If the conditional log-Laplace transform of $\{w_t\}$ is $\psi_{t-1}^{(w)}(u) = a_{t-1}(u)'w_{t-1} + b_{t-1}(u)$, the multi-horizon conditional Laplace transform (MLT) :*

$$\varphi_{t,h}^{(w)} = E_t[\exp(\gamma_1^{(h)'} w_{t+1} + \dots + \gamma_h^{(h)'} w_{t+h})],$$

is equal to :

$$\varphi_{t,h}^{(w)} = \exp(A'_{t,h} w_t + B_{t,h}),$$

where $A_{t,h} = A_{t,h}^{(h)}$, $B_{t,h} = B_{t,h}^{(h)}$, the $A_{t,i}^{(h)}, B_{t,i}^{(h)}, i = 1, \dots, h$ are defined recursively by :

$$\begin{cases} A_{t,i}^{(h)} &= a_{t+h-i}(\gamma_{h+1-i}^{(h)} + A_{t,i-1}^{(h)}), \\ B_{t,i}^{(h)} &= b_{t+h-i}(\gamma_{h+1-i}^{(h)} + A_{t,i-1}^{(h)}) + B_{t,i-1}^{(h)}, \\ A_{t,0}^{(h)} &= 0, B_{t,0}^{(h)} = 0. \end{cases}$$

Proof: see Appendix 3.

This proposition shows that the *MLT* is an exponential-affine function of w_t . To compute $\varphi_{t,h}^{(w)}$, for $t \in \{1, \dots, T\}$ and for given sequences of parameters $(\gamma_1^{(h)}, \dots, \gamma_h^{(h)})$, $h \in \{1, \dots, H\}$, we have, in general, to apply the above algorithm TH times. However, if functions a_t and b_t do not depend on t , we have to use it H times only. More importantly if the parameters $(\gamma_1^{(h)}, \dots, \gamma_h^{(h)})$, $h = 1, \dots, H$, have a *reverse order* structure $\gamma_{h+1-i}^{(h)} = \delta_i$ for $i = 1, \dots, h$ and $h = 1, \dots, H$, that is, if we want to compute :

$$E_t [\exp(\delta'_h w_{t+1} + \dots + \delta'_1 w_{t+h})], h \in \{1, \dots, H\}, t \in \{1, \dots, T\},$$

the algorithm has to be used only once for each date t . If, moreover, a_t and b_t do not depend on t , the algorithm has to be used *only once*.

2.4 Matching Interest Rates Statistical Properties

Before moving to the pricing of defaultable and non-defaultable bonds when the factor is regime switching Car, it is important to show that this class of non-linear models is appropriate to describe interest rates historical dynamics. More precisely, in this section we show that Regime-Switching Gaussian VAR(p) models can represent the observed strong interest rates linear and nonlinear serial dependence, as well as their non-Gaussianity. The proposed empirical analysis considers single-regime, 2-state (homogeneous and non-homogeneous) and 3-state (homogeneous) Regime-Switching Gaussian VAR(1) and VAR(2) models where the factor y_t (say) consists in a short rate (r_t , say), a long-term spread (S_t , say) and a butterfly spread ($S_t^{(b)}$, say), i.e. classical level, slope

and curvature factors, respectively. The family of Gaussian Regime-Switching VAR(p) processes (RS-VAR(p), say) is denoted by:

$$y_{t+1} = \nu + \Phi_1 y_t + \dots + \Phi_p y_{t+1-p} + \Omega(z_{t+1}) \varepsilon_{t+1}, \quad (2.11)$$

where ε_{t+1} is a 3-dimensional Gaussian white noise with $\mathcal{N}(0, I_3)$ distribution [I_3 denotes the (3×3) identity matrix], Φ_ℓ , for each $\ell \in \{1, \dots, p\}$, are (3×3) autoregressive matrices, while ν is a 3-dimensional vector; $\Omega(z_{t+1})$ is a (3×3) lower triangular matrix and (z_t) is the regime-indicator function following a J -state Markov chain (see Section 2.2). If the latter is homogeneous, the transition probabilities will be denoted by $\mathbb{P}(z_{t+1} = e_j | z_t = e_i) = \pi_{ij}$ while, in the non-homogeneous case, they will be denoted by $\mathbb{P}(z_{t+1} = e_j | z_t = e_i, r_t) = \pi(e_j, e_i; r_t)$. The single-regime case assumes a constant Ω . In our empirical analysis we consider $p \in \{1, 2\}$ and we assume $J = 2$ and 3 in the homogeneous case⁶, while in the non-homogeneous case only the two-state case is analyzed⁷ and the transition probabilities are specified by the following logistic function:

$$\mathbb{P}(z_{t+1} = e_j | z_t = e_j, r_t) = \pi(e_j, e_j; r_t) = \frac{e^{a_j + b_j r_t}}{1 + e^{a_j + b_j r_t}}, \quad j \in \{1, 2\}. \quad (2.12)$$

We use 408 monthly observations on U.S. Treasury 1-month, 5-year and 10-year interest rates, taken from the unsmoothed Fama and Bliss (1987) data set, covering the period from January 1970 to December 2003. The short rate is the 1-month yield, the long-term spread is the difference between the 10-year and 1-month yields, while the butterfly spread is given by $S_t^{(b)} = -r_t + 2 R_t(5y) - R_t(10y)$, where $R_t(5y)$ and $R_t(10y)$ denote the 5-year and 10-year yields, respectively. Relevant factors' summary statistics (see Table 1) highlight the lack of Gaussianity and the presence of linear and non-linear serial dependence in the data (see Appendix 4). In the single-regime case, parameters are estimated by OLS, while in the regime switching one they are

⁶In the 3-state case, we assume $\pi_{13} = \pi_{31} = 0$ given that a preliminary estimation of the entire transition matrix has clearly indicated their lack of significance.

⁷We have also estimated alternative specifications in which the constant term in the VAR was regime-dependent or there was a lack of contemporaneous causality from the Markov chain to the factor. The associated empirical performances are outperformed by the those of above mentioned specifications. Thus, for expository purpose, they are not given in the paper but are available upon request from the authors.

estimated by maximizing the log-likelihood function calculated with the Kitagawa-Hamilton filter. Tables 2 and 3 present parameter estimates, likelihood-based selection criteria and residual tests of the above mentioned models (see Appendix 4). In particular, the performances of the models are studied running the Ljung-Box test on any single-equation model residuals and squared residuals, in order to check if both linear and nonlinear serial dependence have been entirely captured, while the Jarque-Bera (JB, say) test is adopted to check the Gaussianity of the error terms ε_t .

The results from this exercise are the following. First, the single-regime Gaussian VAR(2) model, even if able to explain factors' linear dependence better than the VAR(1) case, is clearly not able to match the non-linear one and it is far from providing Gaussian residuals. Second, an important improvement is obtained by adopting a 2-state RS-VAR(2) model able to explain also the non-linear serial dependence, but still unable to provide Jarque-Bera test statistics smaller than critical values. Ljung-Box test on squared model residuals is satisfied if we introduce a second lag instead of moving from an homogeneous to a non-homogeneous Markov chain. Third, the regime-switching specification with 3 regimes and two lags completely satisfies the proposed tests.

3 REGIME SWITCHING AND DEFAULT-FREE BOND PRICING

In this section we describe two models for pricing default-free zero-coupon bonds. In the first model (see Section 3.1) the formulas for the *yields* are *affine* with respect to the factor w_t , whereas in the second model (see Section 3.2) the *affine* structure is obtained for the *prices*. In Section 3.3 we combine both kinds of formulas.

3.1 Regime Switching Affine *Yield* Term Structure Model

3.1.1 Regime switching risk-neutral dynamics and bond pricing

We assume that the new information of the investors at date t is :

$$w_t = (z'_t, y'_t)',$$

where $\{z_t\}$ is a time homogeneous Markov chain and $\{w_t\}$ is Car (1) in the risk-neutral (R.N.) world.

If the short rate r_t is an affine function of w_t :

$$r_t = \beta_0 + \beta'_1 w_t = \beta + \beta'_{11} z_t + \beta'_{12} y_t, \quad (3.1)$$

the price at t of a default-free zero-coupon bond of residual maturity h is :

$$\begin{aligned} B(t, h) &= E_t^{\mathbb{Q}} \exp(-r_t - \dots - r_{t+h-1}) \\ &= \exp(-\beta_0 h - \beta'_1 w_t) E_t^{\mathbb{Q}} \exp[-\beta'_1 (w_{t+1} + \dots + w_{t+h-1})] \end{aligned}$$

According to Proposition 3, the prices $B(t, h), t = 1, \dots, T, h = 1, \dots, H$, are of the form :

$$B(t, h) = \exp(c'_h w_t + d_h),$$

where the c_h, d_h are obtained from a simple recursive scheme (sometimes called Riccati recursive scheme). Therefore, we obtain the *switching affine yield term structure* :

$$R(t, h) = -\frac{c'_h}{h} w_t - \frac{d_h}{h} = -\frac{c'_{1,h}}{h} z_t - \frac{c'_{2,h}}{h} y_t - \frac{d_h}{h}. \quad (3.2)$$

Thus the stochastic term structure is obtained as a combination of baseline deterministic term structures, that are the components of $c_{1,h}, c_{2,h}, d_h$, with stochastic coefficients. An interesting

property of these affine term structure models is that some components of y_t can be chosen as yields of different residual maturities, while staying compatible with pricing formula (3.2). For instance, if the first component is $y_{1,t} = R(t, k)$, we have just to fix $c_{1,k} = 0, c_{2,k} = -ke_1, d_k = 0$, where e_1 is the vector selecting the first component of y_t , in the recursive scheme of Proposition 3. This clearly constrains the R.N. dynamics.

3.1.2 Back to the historical dynamics

Once the R.N. dynamics of $\{w_t\}$ is specified as well as the short rate function $r_t(w_t)$, the historical conditional p.d.f. of w_t given \underline{w}_{t-1} can be specified freely. Equivalently, we can specify any stochastic discount factor satisfying :

$$E_{t-1}^{\mathbb{Q}}[M_{t-1,t}^{-1}(w_t)] = \exp(r_{t-1}). \quad (3.3)$$

A convenient, flexible specification of the s.d.f. is the *exponential-affine* s.d.f. :

$$M_{t-1,t} = \exp\{-r_{t-1} + \gamma'(\underline{w}_{t-1})w_t + \psi_{t-1}^{\mathbb{Q}}[-\gamma(\underline{w}_{t-1})]\}, \quad (3.4)$$

where the vector of *risk sensitivity coefficients* $\gamma(\underline{w}_{t-1})$ is function of the past value of $w_t = (z'_t, y'_t)'$. This specification satisfies condition (3.3) or, equivalently, $E_{t-1}[M_{t-1,t}] = \exp(-r_{t-1})$. This large choice of risk sensitivity coefficients $\gamma(\underline{w}_{t-1})$ implies a large choice of historical dynamics, which in general are not Car. Nevertheless the conditional log-Laplace transform of $\{w_t\}$ in the historical world is easily obtained, since :

$$\begin{aligned} \psi_{t-1}^{\mathbb{P}}(u) &= \log E_{t-1}^{\mathbb{P}}[\exp(u'w_t)] = \log E_{t-1}^{\mathbb{Q}}[M_{t-1,t}^{-1} \exp(-r_{t-1} + u'w_t)] \\ &= -\psi_{t-1}^{\mathbb{Q}}[-\gamma(\underline{w}_{t-1})] + \log E_{t-1}^{\mathbb{Q}}\{\exp[u - \gamma(\underline{w}_{t-1})]'w_t\}, \end{aligned}$$

where $\psi_{t-1}^{\mathbb{Q}}(u)$ is the R.N. conditional log-Laplace transform of w_t . Therefore :

$$\psi_{t-1}^{\mathbb{P}}(u) = \psi_{t-1}^{\mathbb{Q}}[u - \gamma(\underline{w}_{t-1})] - \psi_{t-1}^{\mathbb{Q}}[-\gamma(\underline{w}_{t-1})], \quad (3.5)$$

where : $\psi_{t-1}^{\mathbb{Q}}(u) = a_{t-1}^{\mathbb{Q}}(u)'w_{t-1} + b_{t-1}^{\mathbb{Q}}(u)$,

since the factor process $\{w_t\}$ is Car(1).

3.1.3 A Gaussian Switching Affine Yield Term Structure Model

Let us assume that the R.N. dynamics of $w_t = (z_t', y_t)'$ is given by :

$$y_t = \mu(z_t, z_{t-1}) + \Phi y_{t-1} + \Omega(z_t, z_{t-1})\eta_t, \quad (3.6)$$

where $\{\eta_t\}$ is a standard Gaussian white noise and $\{z_t\}$ is a time homogeneous exogenous Markov chain valued in $\{e_1, \dots, e_J\}$, independent of $\{\eta_t\}$, and with transition matrix Π of general term π_{ij} . $\{y_t\}$ is a RSCar(1) under \mathbb{Q} , $\{w_t\}$ is Car(1) and its conditional log-Laplace transform is given by :

$$\psi_{t-1}^{\mathbb{Q}}(u_1, u_2) = \log E_{t-1}^{\mathbb{Q}}[\exp(u_1' z_t + u_2' y_t)] = [A_1(u_1, u_2), \dots, A_J(u_1, u_2)]z_{t-1} + u_2' \Phi y_{t-1}, \quad (3.7)$$

with $A_i(u_1, u_2) = \log\{\sum_j \pi_{ij} \exp[u_{1,j} + u_2' \mu(e_j, e_i) + \frac{1}{2} u_2' \Sigma(e_j, e_i) u_2]\}$, and $\Sigma(e_j, e_i) = \Omega(e_j, e_i) \Omega'(e_j, e_i)$.

Let us assume that the s.d.f. has the form :

$$M_{t-1,t} = \exp\left[-r_{t-1} + \frac{1}{2} \nu'(z_t, z_{t-1}, y_{t-1}) \nu(z_t, z_{t-1}, y_{t-1}) + \nu'(z_t, z_{t-1}, y_{t-1}) \eta_t + \delta'(z_{t-1}, y_{t-1}) z_t\right], \quad (3.8)$$

$$\begin{aligned} \text{with } \nu(e_j, e_i, y_{t-1}) &= \Omega^{-1}(e_j, e_i) [\tilde{\Phi} y_{t-1} + \tilde{\mu}(e_j, e_i)], \\ \delta_j(e_i, y_{t-1}) &= \log \left[\frac{\pi_{ij}}{\tilde{\pi}(e_j | e_i, y_{t-1})} \right], \end{aligned}$$

where the matrix $\tilde{\Phi}$, and the functions $\tilde{\mu}(z_t, z_{t-1})$, $\tilde{\pi}(z_t | z_{t-1}, y_{t-1})$ can be chosen arbitrarily. In this specification of the s.d.f., both the risks coming from the Gaussian white noise $\{\eta_t\}$ and from the stochastic regime $\{z_t\}$ are priced. The adjustment term $\frac{1}{2} \nu'(z_t, z_{t-1}, y_{t-1}) \nu(z_t, z_{t-1}, y_{t-1})$ and the

form of function δ ensures that the required constraint (3.3) on the s.d.f. $E_t^{\mathbb{Q}}(M_{t-1,t}^{-1}) = \exp(r_{t-1})$ is satisfied. Moreover, the historical dynamics is [see Monfort, Renne (2013)] :

$$y_t = \mu(z_t, z_{t-1}) - \tilde{\mu}(z_t, z_{t-1}) + (\Phi - \tilde{\Phi})y_{t-1} + \Omega(z_t, z_{t-1})\varepsilon_t, \quad (3.9)$$

where $\{\varepsilon_t\}$ is a standard Gaussian white noise under \mathbb{P} , z_t is valued in $\{e_1, \dots, e_J\}$ and such that $\mathbb{P}(z_t = e_j | z_{t-1} = e_i, \underline{y}_{t-1}) = \tilde{\pi}(e_j | e_i, y_{t-1})$. Since $\tilde{\mu}$ and $\tilde{\Phi}$ are free, the same is true for $\mu - \tilde{\mu}$ and $\Phi - \tilde{\Phi}$. The specific form of the s.d.f. provides R.N. and historical dynamics which can differ by their switching drift and autoregressive matrix, but share the same switching volatility matrix processes. In addition, in the historical world, the transition matrix of z_t may depend on the past values of y_t . Since z_t is valued in $\{e_1, \dots, e_J\}$, the s.d.f. given in (3.8) can be written as :

$$M_{t-1,t} = \exp\left[-r_{t-1} + \frac{1}{2}z_t' \tilde{\nu}'(z_{t-1}, y_{t-1}) \tilde{\nu}(z_{t-1}, y_{t-1}) z_t + z_t' \tilde{\nu}(z_{t-1}, y_{t-1}) \eta_t + \delta'(z_{t-1}, y_{t-1}) z_t\right], \quad (3.10)$$

where $\tilde{\nu}(z_{t-1}, y_{t-1})$ is the matrix whose j^{th} column is $\nu(e_j, z_{t-1}, y_{t-1})$. Therefore the s.d.f. $M_{t-1,t}$ is *exponential-quadratic* in (z_t, η_t) , and also exponential-quadratic in (z_t, y_t) ⁸ [see Monfort, Pegoraro (2012)]. If $\nu(z_t, z_{t-1}, y_{t-1})$ does not depend on z_t ⁹, $\tilde{\nu}(z_{t-1}, y_{t-1})$ is equal to $\nu_0(z_{t-1}, y_{t-1})e'$, where $\nu_0(z_{t-1}, y_{t-1})$ is a vector with the same dimension as y_t , and e the vector of size J whose components are all equal to one, and the s.d.f. becomes :

$$M_{t-1,t} = \exp\left[-r_{t-1} + \frac{1}{2}\nu_0'(z_{t-1}, y_{t-1})\nu_0(z_{t-1}, y_{t-1}) + \nu_0'(z_{t-1}, y_{t-1})\eta_t + \delta'(z_{t-1}, y_{t-1})z_t\right], \quad (3.11)$$

which is *exponential-affine* in (z_t, η_t) .

⁸The term $z_t' \tilde{\nu}'(z_{t-1}, y_{t-1}) \tilde{\nu}(z_{t-1}, y_{t-1}) z_t$ can also be written in the linear way $\frac{1}{2} \tilde{\nu}^{2'}(z_{t-1}, y_{t-1}) z_t$, where $\tilde{\nu}^2$ is understood componentwise.

⁹This condition is, in particular, satisfied if there is no instantaneous causality between $\{z_t\}$ and $\{y_t\}$ in both worlds.

3.2 Regime Switching Affine *Price* Term Structure Model

The models described in Section 3.1 provide term structures, where the *yields* are affine functions of the factor $w_t = (z'_t, y'_t)'$. In this section we consider a new kind of RSTSM in which the bond *prices* are affine functions of factors. Contrary to the Regime Switching Affine Yields Term Structure Models, these new models are able to reproduce a behavior of the short term rate staying equal to a lower bound during some spells. We distinguish two cases depending whether the Markov chain is *exogenous*, or *endogenous*.

3.2.1 *Exogenous* Markov chain

Let us consider a process $\tilde{w}_t = (z'_t, r'_t, y'_t)'$, where $\{z_t\}$ is an *exogenous* Markov chain, with transition matrices Π_t in the R.N. world. Thus we assume that the conditional distribution of z_t given $\underline{\tilde{w}}_{t-1}$ depends on z_{t-1} only and is characterized by Π_t , which implies that $\{r_t, y_t\}$ does not cause $\{z_t\}$. We assume that the R.N. conditional distribution of r_t given $\underline{z}_t, \underline{r}_{t-1}, \underline{y}_{t-1}$ depends on z_t only and has a conditional Laplace transform given by :

$$E[\exp(ur_t)|\underline{z}_t, \underline{r}_{t-1}, \underline{y}_{t-1}] = \exp[\gamma_t(u)'z_t],$$

where $\gamma_t(u)$ is the vector :

$$[\gamma_{1t}(u), \dots, \gamma_{Jt}(u)]'.$$

Finally, we assume that the R.N. conditional distribution of y_t given $\underline{z}_t, \underline{r}_t, \underline{y}_{t-1}$ depends on $z_t, \underline{r}_t, \underline{y}_{t-1}$ only. The information of the investors is either $\underline{\tilde{w}}_t$, if z_t is observed, or $\underline{w}_t = (\underline{r}_t, \underline{y}_t)$, if z_t is not observed. If we assume that z_t is not observed by the investors (or hidden), the zero-coupon price $B(t, h)$ is a *linear function* of the transformed factor $\hat{z}_t \exp(-r_t)$, where $\hat{z}_t = E^{\mathbb{Q}}(z_t|\underline{r}_t, \underline{y}_t)$. More precisely we have the following result :

Proposition 4. :

$$B(t, h) = e' P'_{t+h-1}(\tilde{\gamma}_{t+h-1}) \dots P'_{t+1}(\tilde{\gamma}_{t+1}) \hat{z}_t \exp(-r_t),$$

where $P_t(\gamma) = \Pi_t \text{diag}[\exp(\gamma)]$ and $\tilde{\gamma}_t = \gamma_t(-1)$; for $h = 1$, the product of the P matrices reduces to the identity matrix.

Proof : see Appendix 5.

The price of the short term zero-coupon $B(t, 1)$ reduces to $e' \hat{z}_t \exp(-r_t) = \exp(-r_t)$ as expected. If the Markov chain is homogeneous, i.e. $\Pi_t = \Pi$, and the conditional distribution of r_t given z_t does not depend on t , i.e. $\gamma_t(u) = \gamma(u)$, we get the following result :

Corollary 1. : *If $\Pi_t = \Pi, \gamma_t(u) = \gamma(u)$, we have $B(t, h) = e' P'(\tilde{\gamma})^{H-1} \hat{z}_t \exp(-r_t)$, where $\tilde{\gamma} = \gamma(-1)$.*

The zero-coupon prices are explicit linear functions of the transformed factor $\exp(-r_t) \hat{z}_t$, which is nonlinear in $\underline{r}_t, \underline{y}_t$. Therefore it is important to have a simple way to compute the risk-neutral predictions \hat{z}_t . The following proposition shows that \hat{z}_t can be computed recursively using an algorithm similar to the Kitagawa-Hamilton' algorithm.

Proposition 5. :

$$\hat{z}_{t+1} = \frac{\text{diag}(f_t g_t) \Pi_t' \hat{z}_t}{e' \text{diag}(f_t g_t) \Pi_t' \hat{z}_t},$$

where $\text{diag}(f_t g_t)$ is the diagonal matrix, with the k^{th} diagonal term given by :

$$f_{k,t}(r_{t+1}) g_{k,t}(y_{t+1} | \underline{r}_{t+1}, \underline{y}_t),$$

where $g_{k,t}$ is the conditional p.d.f. of y_{t+1} given $z_{t+1} = e_k, \underline{r}_{t+1}, \underline{y}_t$, and $f_{k,t}(r_{t+1})$ is the p.d.f. of r_{t+1} given $z_{t+1} = e_k$.

Proof : see Appendix 6.

The proof in Appendix 6 includes the case where the conditional distribution of r_{t+1} given $z_{t+1} = e_1$ (say) is the point mass at a given value, for instance zero. This allows the short rate to stay at some *lower bound* during some spells.

3.2.2 *Endogenous* Markov chain

In the model of the previous section, the Markov chain $\{z_t\}$ is exogenous in the R.N. world, that is, it is not caused by the other processes $\{r_t, y_t\}$. In this section we consider a situation in which the process $\{z_t\}$ is *endogenous*, that is, caused by the process $\{r_t, y_t\}$.

More precisely, we assume that the risk-neutral conditional distribution of z_t given $(\underline{z_{t-1}}, \underline{r_{t-1}}, \underline{y_{t-1}})$ depends on (r_{t-1}, y_{t-1}) , i.e. is characterized by a J -dimensional vector of probabilities $\beta_{t-1}(r_{t-1}, y_{t-1})$. Moreover, we assume that the R.N. conditional distribution of (r_t, y_t) given $(\underline{z_t}, \underline{r_{t-1}}, \underline{y_{t-1}})$ depends on z_t only. We denote by $\alpha_t(r_t, y_t)$ the J -dimensional vector whose j^{th} component $\alpha_{j,t}$ is the p.d.f. of the conditional distribution of (r_t, y_t) given $z_t = e_j$, with respect to a given basic (dominating) probability. We assume that this probability has in turn a p.d.f. $\alpha_{0,t}(r_t, y_t)$ with respect to a given measure. In other words, for given values of (r_{t-1}, y_{t-1}) , z_t is drawn according to the vector of probabilities β_{t-1} and, then, if $z_t = e_j$, (r_t, y_t) is drawn in the distribution whose p.d.f. with respect to the dominating measure is $\alpha_{0,t}\alpha_{j,t}$. We assume that the information of the investors is $\underline{w}_t = (w'_t, \dots, w'_1)'$ with $w_t = (r'_t, y'_t)'$ and, therefore, z_t is not observed (or hidden).

The conditional p.d.f. or (r_t, y_t) given (r_{t-1}, y_{t-1}) w.r.t. μ is :

$$\alpha_{0,t}(r_t, y_t)\alpha'_t(r_t, y_t)\beta_{t-1}(r_{t-1}, y_{t-1}) \quad (3.12)$$

This kind of dynamics has been introduced by Gouriéroux, Jasiak (2000) and called Finite Dimensional Dependence (FDD) dynamics. It is easily seen that the conditional distribution of z_t given its own past $\underline{z_{t-1}}$ depends on z_{t-1} only; thus, z_t is marginally Markov.

Let us denote by $E_{0,t}$ the expectation with respect to the probability distribution with p.d.f. $\alpha_{0,t}$ and by Π_t the R.N. transition matrix of z_t , whose entries are $\pi_{i,j,t} = \mathbb{Q}_t(z_{t+1} = e_j | z_t = e_i)$. Note that $E_{0,t}$ is an unconditional expectation w.r.t. a distribution depending on time in a deterministic way. We have the following results :

Proposition 6. :

$$\begin{aligned}\Pi_t &= E_{0,t}(\alpha_t \beta'_t), \\ B(t, h) &= e' \tilde{P}'_{t+h-1} \dots \tilde{P}_{t+1} \beta_t \exp(-r_t),\end{aligned}$$

where $\tilde{P}_t = E_{0,t}[\exp(-r_t) \alpha_t \beta'_t]$, and the product of the \tilde{P} matrices reduces to the identity matrix for $h = 1$.

Proof : see Appendix 7.

The formulas obtained for $B(t, h)$ in the exogenous case (Proposition 4) and in the endogenous case (Proposition 6) are similar. The prices are *linear functions of factors*, the P_t matrices appearing in Proposition 4 are replaced by the \tilde{P}_t matrices in Proposition 6 and the factors $\exp(-r_t) \hat{z}_t$ are replaced by the factors $\exp(-r_t) \beta_t$. In both cases $B(t, h)$ is, for any h , a linear combination of factors, but a nonlinear function of the variables $(\underline{r}_t, \underline{y}_t)$ (in the endogenous case the factor β_t are functions of (r_t, y_t) only). In the stationary case where $\alpha_{0,t}, \alpha_t$ and β_t do not depend on t , we get a simplified formula.

Corollary 2. : *In the stationary case, that is, if $\alpha_{0,t}, \alpha_t$ and β_t do not depend on t , we have :*

$$\begin{aligned}B(t, h) &= e' \tilde{P}'^{h-1} \beta(r_t, y_t) \exp(-r_t), \\ \text{with : } \tilde{P} &= E_0[\exp(-r_t) \alpha(r_t, y_t) \beta'(r_t, y_t)].\end{aligned}$$

Two additional remarks are of interest. First, the basic probability may have a p.d.f. $\alpha_{0,t}$ with respect to a measure which is not the Lebesgue measure. For instance it could be such that the probability of the hyperplane $\{r_t = 0\}$ is strictly positive and, moreover, one of the p.d.f. $\alpha_{j,t}$, say $\alpha_{1,t}$, is non zero only in this hyperplane. Thus the short-term rate would be equal to zero in the first regime and would *remain equal to zero* for some time (see the illustration in the next subsection). Second, the FDD dynamics is rather general since it can approximate any Markov dynamics; indeed, any conditional p.d.f. $f(w_t|w_{t-1})$ of $w_t = (r_t, y_t)$ given w_{t-1} can be

approximated by the FDD dynamics :

$$\sum_{j=1}^J f(w_t | \tilde{w}_j) \frac{\mathcal{K}\left(\frac{w_{t-1} - \tilde{w}_j}{d}\right)}{\sum_{j=1}^J \mathcal{K}\left(\frac{w_{t-1} - \tilde{w}_j}{d}\right)}, \quad (3.13)$$

where \mathcal{K} is a kernel, $\tilde{w}_j, j = 1, \dots, J$ a fixed grid and d a bandwidth.

Finally, let us consider the historical dynamics. Since the R.N. and historical conditional distributions of w_t given the past are equivalent, the historical conditional distribution is absolutely continuous with respect to the probability defined by $\alpha_{0,t}$. We also have the following result:

Proposition 7. : *If the R.N. dynamics is FDD, the historical dynamics is also FDD if and only if the s.d.f. is factorized as $M_{1,t-1,t}(w_t) M_{2,t-1,t}(w_{t-1})$.*

Proof : See Appendix 8.

3.2.3 The zero lower bound problem

Both kinds of Regime Switching Affine Price Term Structure Models are able to generate paths of the short rate staying at a lower bound, zero for instance during some endogenous spells. As an illustration, let us consider a FDD model in which $\{y_t\}$ is univariate and the number of states is $J = 3$. The conditional risk-neutral probabilities of the regime are given by :

$$\beta_{j,t-1} = \frac{\varphi\left(\frac{r_{t-1} + y_{t-1} - k_j}{d}\right)}{\sum_{l=1}^3 \varphi\left(\frac{r_{t-1} + y_{t-1} - k_l}{d}\right)}, \quad (3.14)$$

where φ is the p.d.f. of the standard normal, $k_j, j = 1, 2, 3$ are given values and d is a bandwidth.

We assume that r_t and y_t are independent conditionally on (z_t, w_{t-1}) and:

i) the distribution of r_t is the point mass at zero, if $j = 1$, while it is the gamma distribution $\gamma(\nu_j, \mu_j)$, with mean m_j and variance σ_j^2 , if $j = 2$ or 3 , that is, with $\nu_j = m_j^2/\sigma_j^2$ and $\mu_j = m_j/\sigma_j^2$.

ii) the distribution of y_t is the gamma distribution $\gamma(\nu_2, \mu_2)$.

Since we are in a stationary case, the price at time t of a zero-coupon bond of residual maturity h is given by the formula of Corollary 2 :

$$B(t, h) = e' \tilde{P}'^{h-1} \beta(r_t, y_t) \exp(-r_t). \quad (3.15)$$

The matrix $\tilde{P}' = E_0[\exp(-r_t) \beta(r_t, y_t) \alpha'(r_t, y_t)]$ is easily computed by Monte-Carlo. More precisely its first column can be approximated by : $\frac{1}{S} \sum_{s=1}^S \beta(0, y^s)$, where the simulated y^s are drawn in $\gamma(\gamma_2, \mu_2)$. The columns $j = 2, 3$, can be approximated by : $\frac{1}{S} \sum_{s=1}^S \exp(-r^s) \beta(r^s, y^s)$, where the simulated y^s are drawn in $\gamma(\nu_2, \mu_2)$ and the simulated rates r^s in $\gamma(\nu_2, \mu_2)$, if $j = 2$, and $\gamma(\nu_3, \mu_3)$, if $j = 3$. For the Monte-Carlo analysis, we do not distinguish the R.N. and historical dynamics and the numerical values of the parameters are :

$$\begin{aligned} k_1 &= .03, k_2 = .05, k_3 = .07, d = .005, \\ m_2 &= .03, \sigma_2 = .01, m_3 = .04, \sigma_3 = .02. \end{aligned}$$

We simulate paths of length $T = 50$ for the factor (r_t, y_t) and for the yields $R(t, h) = -\frac{1}{h} \log B(t, h)$, for $h = 5, 10, 20, 100$ (and initial values $r_1 = y_1 = .001$). Figure 1 shows such paths. The short rate r_t is equal to zero in periods 2 to 8, 18 to 20 and 39 to 47. Within these periods, the rest of the yield curve is varying (see in particular the third period).

3.3 A simultaneous use of explicit and recursive pricing formulas

In Sections 3.1, 3.2, we have obtained either explicit, or recursive formulas for the prices of zero-coupon bonds. There are many ways to jointly use these results. In this section we consider such an approach and an application.

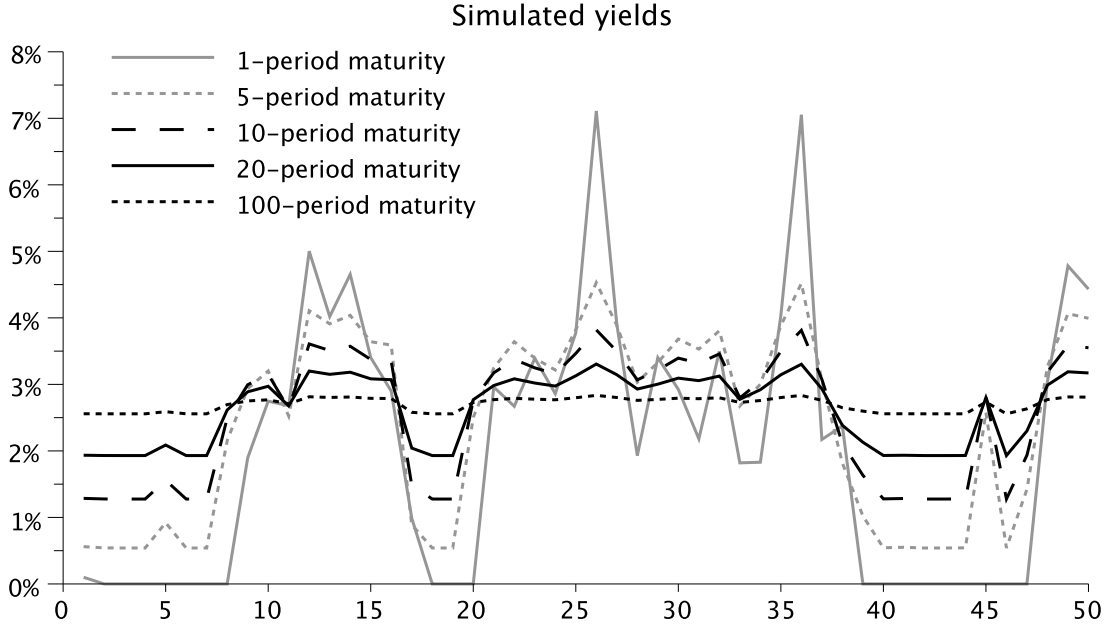


Figure 1: Interest rates paths and the lower bound: simulated paths of yields $R(t, h) = -\frac{1}{h} \log B(t, h)$, for $h = 5, 10, 20, 100$. Initial values: $r_1 = y_1 = .001$.

3.3.1 A flexible framework

Let us consider two independent Markov chains in the risk-neutral world, denoted by $\{z_t^{(1)}\}, \{z_t^{(2)}\}$, with J_1 and J_2 states and transitions matrices $\Pi_t^{(1)}$ and $\Pi_t^{(2)}$, respectively. Moreover, let us consider an independent Car(1) process $\{y_t\}$ and a sequence of $K \times J_2$ matrices Δ_t serially independent and independent of the other processes. Finally let us assume that the short rate between t and $t + 1$ is given by :

$$r_t = \mu'_1 z_t^{(1)} + \mu'_2 \Delta_t z_t^{(2)} + \mu'_3 y_t, \quad (3.16)$$

If we assume that $z_t^{(1)}, z_t^{(2)}$ and y_t are observed by the investor, the price of the zero coupon

bond $B(t, h)$ is :

$$\begin{aligned}
B(t, h) &= \exp(-r_t) E_t^{\mathbb{Q}} \exp(-r_{t+1} - \dots - r_{t+h-1}) \\
&= \exp(-r_t) B_{1,t}(h) B_{2,t}(h) B_{3,t}(h) \\
\text{where } B_{1,t}(h) &= E_t^{\mathbb{Q}} \exp\left(-\mu'_1 z_{t+1}^{(1)} - \dots - \mu'_1 z_{t+h-1}^{(1)}\right) \\
B_{2,t}(h) &= E_t^{\mathbb{Q}} \exp\left(-\mu'_2 \Delta_{t+1} z_{t+1}^{(2)} - \dots - \mu'_2 \Delta_{t+h-1} z_{t+h-1}^{(2)}\right) \\
B_{3,t}(h) &= E_t^{\mathbb{Q}} \exp\left(-\mu'_3 y_{t+1} - \dots - \mu'_3 y_{t+h-1}\right)
\end{aligned}$$

Using Proposition 1, $B_{1,t}(h)$ is an explicit linear function of $z_t^{(1)}$, or, equivalently, an explicit exponential linear function of $z_t^{(1)}$, since $z_t^{(1)}$ is valued in $\{e_1, \dots, e_J\}$,

$$B_{1,t}(h) = \exp\left[a'_{1,t}(h) z_t^{(1)}\right].$$

Similarly, conditioning first by $z_{t+1}^{(2)}, \dots, z_{t+h-1}^{(2)}$ and taking the expectation in $B_{2,t}(h)$ with respect to $\Delta_{t+1}, \dots, \Delta_{t+h-1}$, we get a closed form expression for $B_{2,t}(h)$:

$$B_{2,t}(h) = \exp\left[a'_{2,t}(h) z_t^{(2)}\right].$$

Using Proposition 3 we get :

$$B_{3,t}(h) = \exp[a'_{3,t}(h) y_t + a_{4,t}(h)],$$

where $a_{3,t}(h)$ and $a_{4,t}(h)$ can be computed recursively. Finally we get :

$$B(t, h) = \exp[a'_{1,t}(h) z_t^{(1)} + a'_{2,t}(h) z_t^{(2)} + a'_{3,t}(h) y_t + a_{4,t}(h)]$$

and :

$$R(t, h) = -\frac{1}{h} \left[a'_{1,t}(h) z_t^{(1)} + a'_{2,t}(h) z_t^{(2)} + a'_{3,t}(h) y_t + a_{4,t}(h) \right], \quad (3.17)$$

where $a_{1,t}(h), a_{2,t}(h)$ have closed forms and $a_{3,t}(h), a_{4,t}(h)$ can be computed recursively. Therefore

we get a very flexible framework which is able to take into account simultaneously many features :

- switching regimes with *deterministic values*
- switching regimes with *stochastic values*
- transition matrices *depending on time* in a deterministic way
- *quantitative factors*.

An application with these features is the multi-regime model developed in the next section.

3.3.2 A multi-regime model: the euro-area yield curve with discrete policy rates

While policy rates are known to be key drivers in the dynamics of the whole yield curve, only a few term-structure models explicitly consider monetary-policy rates (Rudebusch (1995), Balduzzi, Bertola and Foresi (1997), Piazzesi (2005) and Fontaine (2009), are notable exceptions). This rarity stems from the difficulties associated with the modeling of policy rates' dynamics. In particular, most central banks set their policy rates in multiples of 25 basis points, implying stepwise paths. This application illustrates how the flexibility of the short-term rate's specification given in (3.16) can be exploited in order to construct a term-structure model where the central-bank policy rate plays a central role. The main features and results of the model are reported here; a complete study can be found in Renne (2012).

A specificity of this model is the large number of states represented by the Markov chain $z_t^{(1)}$. Indeed, each state of $z_t^{(1)}$ is defined by (a) one of the possible values of the main policy rate of the European Central Bank (ECB) and (b) a monetary policy phase: tightening (T), status-quo (S) and easing (E). A tightening (resp. easing) monetary policy aims at restricting (resp. weakening) credit conditions; that is, during a tightening (easing) phase, the central bank is expected to raise (cut) its policy rate. The component $\mu'_1 z_t^{(1)}$ of the short rate (see 3.16) corresponds to the ECB policy rate, which implies that the entries of μ'_1 are of the form $\log[1 + k \times 0.25\%/360]$ with $k = 0, \dots, 40$, 25bp being the basic tick¹⁰. The probabilities of increases and cuts in the policy rate are defined by the matrix $\Pi_t^{(1)}$ (whose dimension is 123×123). These probabilities depend on

¹⁰Observe that 10% is assumed to be the maximum value of the (arithmetic) policy rate.

the level of the (geometric) policy rate as well as on the monetary-policy phase. During tightening phases (resp. easing phases), the probability of a cut (resp. a increase) in the policy rate is zero. No policy-rate move takes place during status-quo phases. Such features make it possible to model policy inertia, implying that policy-rate changes are often followed by additional changes in the same direction. This phenomenon is illustrated in the first panel of Figure 2, that shows successions of periods of increases and periods of decreases in the ECB policy rates. Monetary-policy phases turn out to affect significantly the yield curve: one can for instance observe in panel C of Figure 2 that the spread between the (short-term) policy rate and a longer-term rate (e.g. the 6-month rate) tends to be positive during periods of rising policy rates and negative during easing phases. These features are easily captured by this model.

If one wants to model the overnight-indexed swap yield curve, the shortest-term (overnight) interest rate to consider is not the policy rate but the interbank rate, which is called EONIA (Euro Over-Night Index Average) in the euro area and denoted by r_t . Therefore, Renne (2012) introduces in the model a specification of the so-called EONIA spread, that is the spread between the euro-area overnight interbank rate and the policy rate. The evolution of this spread is displayed in Panel B of Figure 2. A dramatic change in the EONIA spread dynamics took place in Fall 2008, in the aftermath of Lehman's failure. While it was slightly positive on average before this failure, the EONIA spread suddenly dropped after the implementation of changes in the monetary-policy operational framework in Fall 2008. The latter led to an excess liquidity in the banking sector at the aggregate level. The impact of the "excess-liquidity" regime on the short-term rate is modeled through an additional two-state Markov chain $z_t^{(2)}$.¹¹ Using the notations of equation (3.16), Δ_t is a bivariate row vector of independent variables whose distributions are mixtures of beta distributions. Typically, the distribution of the entry of Δ_t that corresponds to the excess-liquidity regime has a negative mean and is positively skewed. This is consistent with the values of the EONIA spread that are observed during most of the periods between Fall 2008 and the end of the sample (see Panel B of Figure 2).¹²

¹¹One of these two states corresponds to the excess-liquidity regime.

¹²During the excess-liquidity regime, the overnight interbank rate tends to be low; since it can not be below the

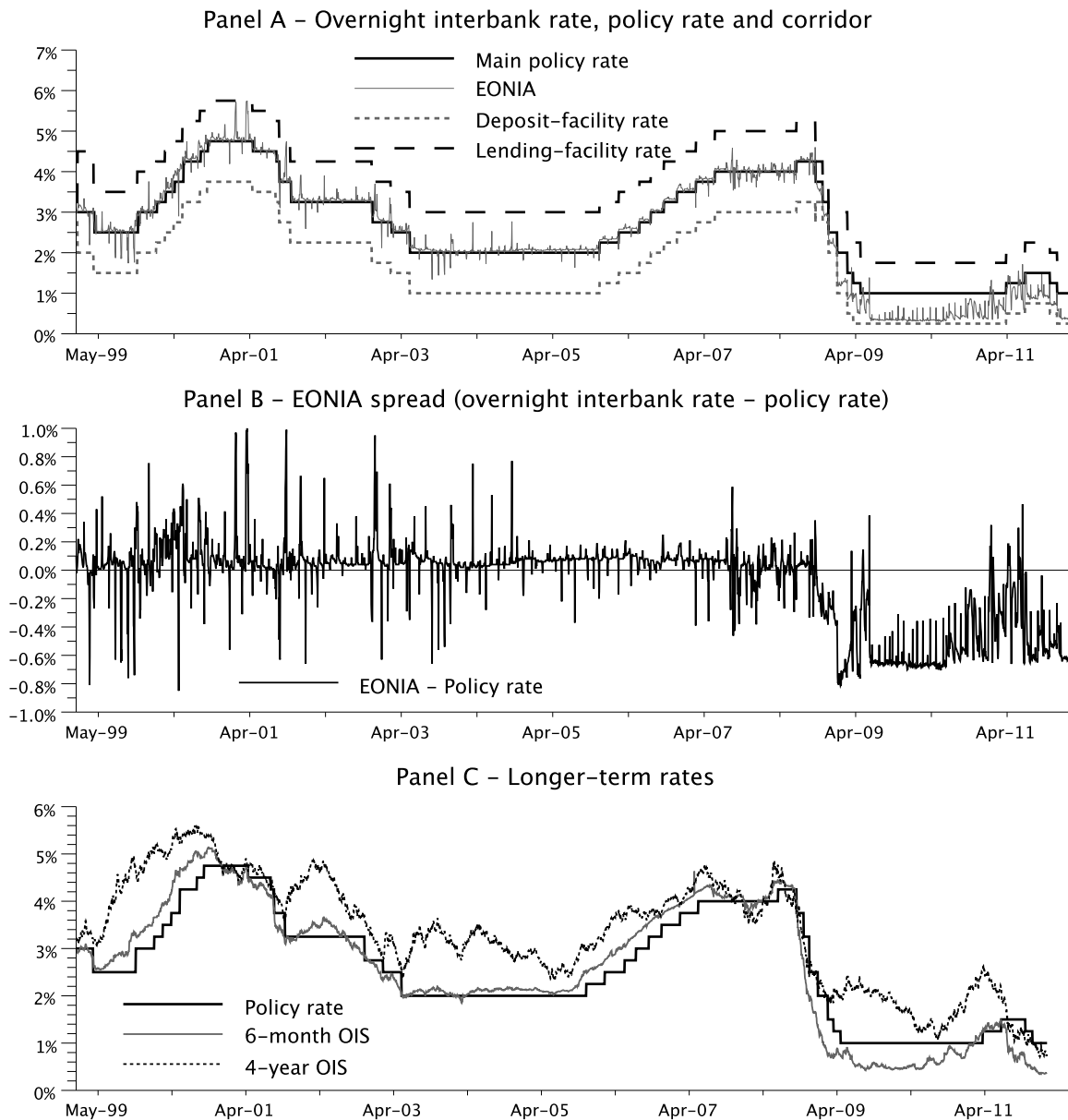


Figure 2: The first panel shows the target rate together with the overnight interbank interest rate (EONIA). The dashed lines define the monetary-policy "corridor" whose upper bound is the Eurosystem marginal-lending-facility rate and the lower bound is the Eurosystem deposit-facility rate. Since the Eurosystem's banks can lend at the former rate and borrow at the latter rate, the overnight interbank rate evolves between these two rates. The second panel displays the EONIA spread, which is the spread between the EONIA and the rate of the main refinancing operations (MRO) decided by the Governing council of the ECB. The third panel plots the main policy rate together with longer-term rates: the 6-month and the 4-year OIS rates.

marginal deposit facility rate (since banks can always deposit funds with the ECB using this facility), we get the positive skewness of the EONIA spread.

4 REGIME SWITCHING AND DEFAULTABLE BOND PRICING

The RSTSM can be extended to the modelling of defaultable bonds. In this framework, we distinguish the individual default indicators and associated individual risk factors from the common risk factors. This modeling is illustrated by an analysis of the Euro-zone sovereign bonds.

4.1 The setting

4.1.1 Risk-neutral dynamics and causality structure

The new information in the economy at date t is $w_t = (z'_t, y'_t, w'_{s,t}, d'_t)'$, where z_t is a regime variable valued in $\{e_1, \dots, e_J\}$, y_t is a vector of common factors, $w_{s,t}$ is a vector $(w_{s,t}^1, \dots, w_{s,t}^n, \dots, w_{s,t}^{N'})'$ of specific variables, $w_{s,t}^n$ corresponding to debtor n ($n = 1, \dots, N$) and $d_t = (d_t^1, \dots, d_t^n, \dots, d_t^N)'$ is a vector of default indicators, where $d_{n,t} = 1$, if entity n is in default at date t , $d_t^n = 0$, otherwise. Thus there are *two* kinds of regime variables: z_t is a *systematic* regime variable and d_t is a set of *individual binary* regime variables $d_t^n, n = 1, \dots, N$.

We use below the following notations : $w_{c,t} = (z'_t, y'_t)'$ for the common variables, $\tilde{w}_t = (w'_{c,t}, w'_{s,t})'$ for all common and specific variables, $\tilde{w}_t^n = (w'_{c,t}, w'_{s,t})'$ for common variables and specific variables of entity n only.

We make some assumptions about the R.N. dynamics of process $\{w_t\}$, in particular about its R.N. causality structure. Since these assumptions concern the risk-neutral distribution, their economic interpretation is in terms of pricing only, not in terms of historical prediction.

A.1. (R.N. Causality structure): $(w'_{s,t}, d'_t)'$ does not Granger cause $\{w_{c,t}\}$, and, $\{d_t\}$ does not cause $\{\tilde{w}_t\}$.

A.2. (R.N. Conditional independence of the entity behaviors): the variables $(w_{s,t}^n, d_t^n)'$, $n = 1, \dots, N$, are independent conditionally on $(w'_{c,t}, \underline{w'_{t-1}})'$, and the conditional distribution of $w_{s,t}^n$

only depends on $(\underline{w'_{c,t}}, \underline{w'^{n'}_{s,t-1}})$.

A.3. (Car distributed processes): process $\{w_{c,t}\}$ is Car(1) and the process of individual risk factors $\{w_{s,t}^n\}$ is conditionally Car(1), that is, the conditional Laplace transform of $w_{s,t}^n$ given $\underline{w_{c,t}}, \underline{w_{s,t-1}^n}$ is exponential affine in $w_{c,t}, w_{c,t-1}, w_{s,t-1}$ (which implies that \tilde{w}_t^n is Car(1)).

A.4. (R.N. default intensity): $\mathbb{Q}(d_t^n = 0 \mid d_{t-1}^n = 0, \tilde{w}_t) = \exp(-\lambda_{n,t}^{\mathbb{Q}})$, with $\lambda_{n,t}^{\mathbb{Q}} = \alpha_{0,n} + \alpha'_{1,n} \tilde{w}_t^n$, and $\mathbb{Q}(d_t^n = 1 \mid d_{t-1}^n = 1, \tilde{w}_t) = 1$, that is, the state $d_t^n = 1$ is absorbing. $\lambda_{n,t}^{\mathbb{Q}}$ is called the default intensity.

The exponential expression of the R.N. transition probability ensures its positivity, and the affine expression of the intensity is introduced to facilitate the computation of the term structure. Since the transition probability is also smaller than 1, the intensity has to be nonnegative, which induces restrictions on the R.N. dynamics of $\{\tilde{w}_t^n\}$.

A.5. (Risk-free rate) : The risk-free short rate between t and $t + 1$ is : $r_t = \beta_0 + \beta'_1 w_{c,t}$.

Since the individual risk factors do not appear in the expression of the risk-free rate, no individual entity has an impact on the risk-free prices. Under Assumptions A.1 and A.5, the spot price of any derivative written on $w_{c,t}$ depends on the past of the common factor only.

4.1.2 Pricing defaultable bonds

Let us consider the case where the recovery rate is zero. The price at time t of a zero-coupon bond issued by entity n , with residual maturity h , is :

$$B_n(t, h) = E_t^{\mathbb{Q}}[\exp(-r_t \dots - r_{t+h-1})(1 - d_{t+h}^n)]. \quad (4.1)$$

Although $(\tilde{w}'_t, d_t^n)'$ is not R.N. Car (1), the causality structure assumed above implies that $B_n(t, h)$ can still be expressed as a multi-horizon Laplace transform of the process $\{\tilde{w}_t^n\}$, with reverse ordered coefficients (see Subsection 2.1). More precisely we have the following Proposition,

which justifies formula (2.10) :

Proposition 8. : *Under Assumptions A.1-A.2, A.4-A.5:*

$$\begin{aligned} B_n(t, h) &= \exp(-r_t) E_t^{\mathbb{Q}}[\exp(-r_{t+1} - \dots - r_{t+h-1} - \lambda_{n,t+1}^{\mathbb{Q}} - \dots - \lambda_{n,t+h}^{\mathbb{Q}})] \\ &= \exp[-h(\beta_0 + \alpha_{0,n}) - \tilde{\beta}'_1 \tilde{w}_t^n] \\ &\quad \times E_t^{\mathbb{Q}}\{\exp[-(\tilde{\beta}_1 + \alpha_{1,n})' \tilde{w}_{t+1}^n - \dots - (\tilde{\beta}_1 + \alpha_{1,n})' \tilde{w}_{t+h-1}^n - \alpha'_{1,n} \tilde{w}_{t+h}^n]\}, \end{aligned}$$

where $\tilde{\beta}_1 = (\beta'_1, 0)'$.

Proof : see Appendix 9.

If moreover Assumption A.3 is satisfied, $\{\tilde{w}_t^n\}$ is Car(1) and, since the Laplace transform is with a reverse order structure :

$$\delta_1 = -\alpha_{1,n}, \delta_j = -(\tilde{\beta}_1 + \alpha_{1,n}), \forall j \geq 2,$$

the prices $B_n(t, h), t = 1, \dots, T, h = 1, \dots, H$ can be computed recursively by using only *once* the algorithm of Proposition 3. So the yield $R_n(t, h)$ of residual maturity h associated with entity n is an *affine function* of \tilde{w}_t^n :

$$R_n(t, h) = c'_n(h) \tilde{w}_t^n + b_n(h), \text{ say.} \quad (4.2)$$

The risk-free rate of residual maturity h is obtained by the same algorithm, with $\alpha_{0,n} = 0, \alpha_{1,n} = 0$, and is an affine function of $w_{c,t}$:

$$R^*(t, h) = c'(h) w_{c,t} + b(h), \text{ say,} \quad (4.3)$$

as are the spreads :

$$R_n(t, h) - R^*(t, h) = [c_n(h) - c^*(h)]' \tilde{w}_t^n + b_n(h) - b(h),$$

where $c^*(h) = [c'(h), 0]'$.

The risk-free and defaultable term structures are all affine. They differ by the baseline term structures and the set of factors involved in their affine expressions. Also note that a direct impact of the regime variable appears since $w_{c,t} = (z'_t, y'_t)'$. This result can be extended to a "market value" recovery rate [see Duffie, Singleton (1999), Monfort, Renne (2013), and Appendix 10].

4.1.3 The historical dynamics

Once the R.N. distribution $f^{\mathbb{Q}}$ and the short rate r_{t-1} are specified, the historical p.d.f. $f^{\mathbb{P}}$ can be chosen arbitrarily and the s.d.f. $M_{t-1,t}$ is deduced from (2.4). In this section, we assume that the s.d.f. $M_{t-1,t}$ depends on the common variables $\underline{w_{c,t}}$ only :

Assumption A.6: $M_{t-1,t}$ is a function of $\underline{w_{c,t}}$.

Assumption A.6 means that the individual variables $w_{s,t}$ and d_t have no impact on the adjustment for risk. This assumption has important consequences. Let us first show a lemma.

Lemma : *If w_t is partitioned into $w_t = (w'_{1,t}, w'_{2,t})'$ and if the s.d.f. $M_{t-1,t}$ is a function of $(w_{1,t}, \underline{w_{t-1}})$:*

i) the R.N. and the historical conditional distributions of $w_{1,t}$ given $\underline{w_{t-1}}$ satisfy the relation :

$$f^{\mathbb{P}}(w_{1,t}|\underline{w_{t-1}}) = f^{\mathbb{Q}}(w_{1,t}|\underline{w_{t-1}})M_{t-1,t}^{-1}(w_{1,t}, \underline{w_{t-1}}) \exp(-r_{t-1}).$$

ii) the R.N. and the historical conditional distributions of $w_{2,t}$ given $(w_{1,t}, \underline{w_{t-1}})$ are the same.

Proof : Equation (2.3) can be written as

$$\begin{aligned} f^{\mathbb{Q}}(w_{1,t}|\underline{w_{t-1}})f^{\mathbb{Q}}(w_{2,t}|w_{1,t}, \underline{w_{t-1}}) &= f^{\mathbb{P}}(w_{1,t}|\underline{w_{t-1}})f^{\mathbb{P}}(w_{2,t}|w_{1,t}, \underline{w_{t-1}}) \\ &\times M_{t-1,t}(w_{1,t}, \underline{w_{t-1}}) \exp(r_{t-1}). \end{aligned}$$

Integrating both sides of this equation with respect to $w_{2,t}$ gives the equality *i)* of the Lemma,

and *ii*) follows. \square

The Lemma above shows the consequences of the absence of some risk factors in the s.d.f. Let us now apply it to see the consequences of the additional Assumption A.6 on the joint R.N. and historical analysis of default.

Proposition 9. :

Under Assumption A.6 on the s.d.f. and Assumption A.1 of non-causality from $(w'_{s,t}, d'_t)'$ to $w_{c,t}$:

i) $\{w'_{s,t}, d'_t\}'$ does not cause $\{w_{c,t}\}$ in the historical world.

ii) the R.N. and the historical conditional distributions of $(w'_{s,t}, d'_t)'$ given $(w_{c,t}, \underline{w_{t-1}})$ are the same.

Proof :

Proposition 9-*i*) is obtained from Lemma-*i*) by taking $w_{1,t} = w_{c,t}$ and $w_{2,t} = (w'_{s,t}, d'_t)'$, and noting that $f^{\mathbb{Q}}(w_{1t} | \underline{w_{t-1}}) M_{t-1,t}$ and r_t depend on $\underline{w_{t-1}}$ through $\underline{w_{c,t-1}}$ only, Proposition 9-*ii*) is obtained from Lemma-*ii*) \square .

Proposition 9-*ii*) implies that Assumptions A.2 and A.4 are also valid in the historical world. In particular the historical and R.N. default intensities are the same :

$$\lambda_{n,t}^{\mathbb{P}} = \lambda_{n,t}^{\mathbb{Q}} = \alpha_{0,n} + \alpha'_{1,n} \tilde{w}_t^n. \quad (4.4)$$

However, equality (4.4) does not imply that the historical intensity $\lambda_{n,t}^{\mathbb{P}}$ (or the R.N. intensity $\lambda_{n,t}^{\mathbb{Q}}$) has the same dynamic behavior in both worlds since the R.N. and historical dynamics of common risk factor $w_{c,t}$ are *different* and such that :

$$f_c^{\mathbb{P}}(w_{c,t} | \underline{w_{c,t-1}}) = f_c^{\mathbb{Q}}(w_{c,t} | \underline{w_{c,t-1}}) M_{t-1,t}^{-1}(\underline{w_{c,t}}) \exp[-r_{t-1}(\underline{w_{c,t-1}})]. \quad (4.5)$$

Once $f_c^{\mathbb{Q}}$ and r_{t-1} have been specified, $f_c^{\mathbb{P}}$ can be chosen arbitrarily and the s.d.f. is deduced from

(4.5). However, we can specify $M_{t-1,t}$ in a way which makes it easily interpretable, while giving a tractable historical dynamics for $w_t = (z'_t, y'_t)'$. The general method based on an *exponential-affine* s.d.f. presented in Section 3.1.2. remains valid, as well as the case of a switching VAR model described in Section 3.1.3.

4.2 Credit vs liquidity risks in euro-area sovereign yield curves

The following application is detailed in Monfort, Renne (2011). Its objective is to model the sovereign yield curves of ten euro-area countries in order to disentangle the impacts of the credit and liquidity risks, and to evaluate the historical and risk-neutral evolutions of the probabilities of default. We introduce a hidden Markov chain $\{z_t\}$ with two regimes in order to capture crisis periods.

Estimation data include monthly yields with residual maturities 1,2,5 and 10 years, for the period between July 1999 and March 2011. The sovereign issuers are Austria, Belgium, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal and Spain. The German bonds, known as *Bunds*, are considered as risk-free. The identification of liquidity-pricing effects relies on the spreads between the German sovereign bonds and those issued by KfW (Kreditanstalt für Wiederaufbau), a German agency whose bonds are fully and explicitly guaranteed by the Federal Republic of Germany. The credit qualities of German sovereign and KfW bonds are the same, implying that the KfW-Bund spread is essentially liquidity-driven [see Monfort and Renne (2011) for a detailed treatment of this point].

We use a Regime-Switching VAR (1) model with a five dimensional factor y_t . The observable entries of y_t are: the 10-year German yield, the slope of the German yield curve (10 year – 1 month), the convexity of the German yield curve (2×3 year – 10 year – 1 month), the first and second principal components of the spreads of four countries (France, Italy, Spain and the Netherlands) versus Germany of 10-year maturity.

Both the risk-neutral and historical models for the factor y_t are RSCar(1):

$$y_t = \mu' z_t + \Phi y_{t-1} + \Omega(z_t) \varepsilon_t,$$

where $\{z_t\}$ is a two-regime exogenous time homogenous Markov chain. The default-free yields are given by:

$$R(t, h) = -\frac{1}{h} \log E_t^{\mathbb{Q}}[\exp(-r_t \dots - r_{t+h-1})], \quad (4.6)$$

where r_t is the one-month risk-free yield. The risky yields are given by:

$$R_n(t, h) = -\frac{1}{h} \log E_t^{\mathbb{Q}}[\exp(-r_t \dots - r_{t+h-1} - \lambda_{n,t+1} - \dots - \lambda_{n,t+h})], \quad (4.7)$$

where the intensity $\lambda_{n,t}$ is decomposed into $\lambda_{n,t} = \lambda_{n,t}^c + \lambda_{n,t}^l$, $\lambda_{n,t}^c$ and $\lambda_{n,t}^l$ being the credit (or default) intensity and the illiquidity intensity, respectively [see Liu, Longstaff, Mandell (2006), Feldhütter, Lando (2008), Fontaine, Garcia (2012)].

The disentangling of the credit and illiquidity effects is based on the above-mentioned interpretation of the KfW-Bund spread and on the assumption according to which the $\lambda_{n,t}^l$ are affine functions of the illiquidity intensity obtained for KfW bonds. The intensities are assumed to be affine functions in z_t and y_t , and, since y_t is RSCar (1), formulas (4.6) and (4.7) provide affine functions in z_t and y_t for $R(t, h)$ and $R_n(t, h)$.

The Kitagawa-Hamilton algorithm is used to compute the probabilities of being in the crisis regime. Figure 3 illustrates that the crisis periods are associated with increasing and highly-volatile sovereign spreads. The approach provides a good data fit (see Figure 4). The standard deviation of the yield pricing errors is of 18 bp and the model accounts for 98% of the yields' variances.

This framework allows us to compute historical probabilities of default. Figure 5 displays the estimated term structures of probabilities of default (PDs), at two dates of our sample. Note that most of the methods implemented by practitioners to extract market-perceived PDs implicitly assume that historical and risk-neutral probabilities are equal. However, based on the present methodology, Monfort, Renne (2011) show that the historical probabilities of default tend to be

significantly lower than their risk-neutral counterparts.

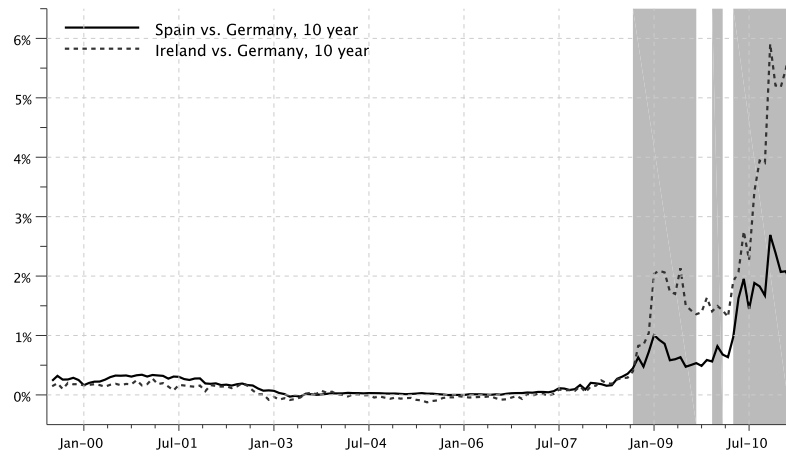


Figure 3: The crisis regime. The grey-shaded areas correspond to crisis periods (estimated as those periods for which the smoothed probabilities of the crisis regime are higher than 50%. The smoothed probabilities are based on the Kitagawa-Hamilton algorithm). The plot also displays the Spanish-German and the Irish-German 10-year sovereign spreads.

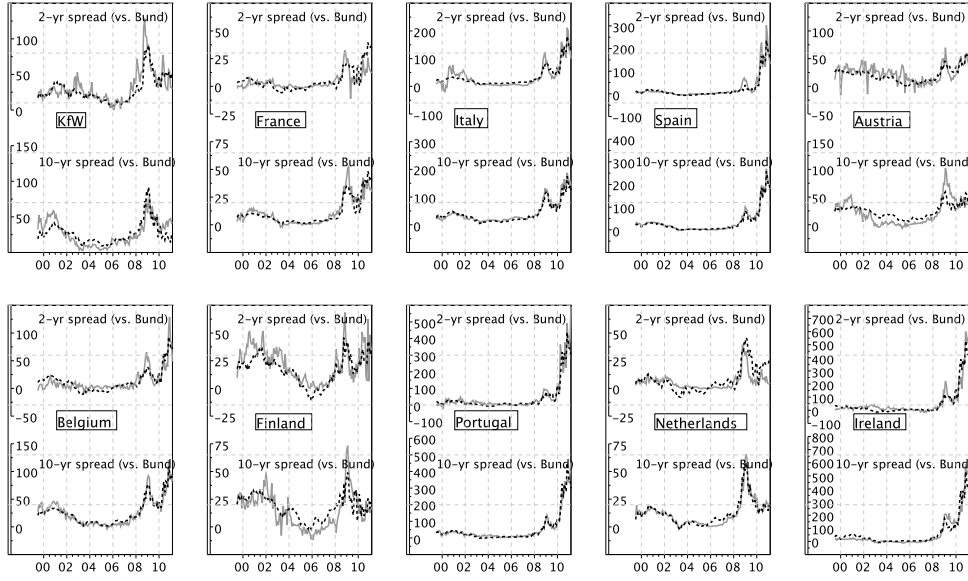


Figure 4: Model-implied vs. actual spreads. The black dotted lines (grey solid lines) correspond to model-implied (actual) spreads.

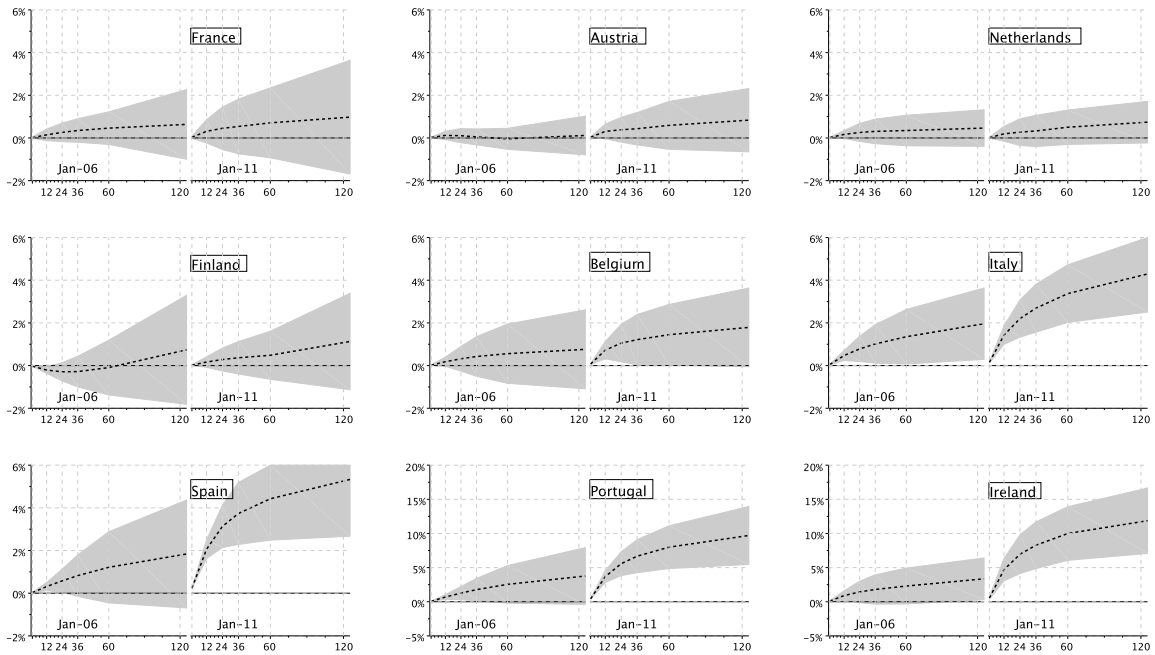


Figure 5: Term structures of (historical-world) probabilities of default. These plots show the term structures of default probabilities at two different dates for the different countries. For instance, for country n and for the five-year maturity (60 months on the x -axis), the plot reports the model-implied probability that country n defaults in the next five years. Note that these probabilities are historical ones, that is, they are based on the historical dynamics of the factor. 95% confidence intervals are reported.

5 CONCLUDING REMARKS

In this paper we have stressed the role of the regime switching approach in various kinds of bond pricing models. We have seen that the regimes can capture a wide variety of underlying phenomena, and that the Regime Switching model are able to combine flexibility and tractability.

There are many related topics which have not been treated in this paper, in particular the inference methods adapted to these kinds of models, the treatment of contagion and the modeling of rating dynamics [see Monfort, Renne (2013)], the simultaneous modeling of nominal and real yield curves [see Ang, Bekaert, Wei (2008)], the joint modeling of yield curves of several countries and the associated exchange rates [see, among the others, Backus, Foresi and Telmer (2001), Brennan and Xia (2006), Leippold and Wu (2007), Gourieroux, Monfort, Sufana (2010) and Graveline, Joslin (2011) for an approach without switching regimes] or the ability of regime switching models to solve the bias problems appearing in the estimation of highly persistent models [see Jardet, Monfort and Pegoraro (2013)].

APPENDIX 1 - Proof of Proposition 1

For expository purpose we omit the exponent h in $\gamma_i^{(h)}$. The formula is true for $h = 1$ since :

$$E_t \exp(\gamma'_1 z_{t+1}) = e' \text{diag}[\exp(\gamma_1)] E_t(z_{t+1}) = e' \text{diag}[\exp(\gamma_1)] \Pi'_{t+1} z_t = e' P'_{t+1}(\gamma_1) z_t,$$

since $E_t(z_{t+1}) = (\pi_{i,1,t+1}, \dots, \pi_{i,J,t+1})' = \Pi'_{t+1} e_i$, if $z_t = e_i$. Assuming that the formula of Proposition 1 is true for $h - 1$, we get :

$$\begin{aligned} \varphi_{t,h}^{(z)} &= E_t[\exp(\gamma'_1 z_{t+1} + \dots + \gamma'_h z_{t+h})] = E_t[\exp(\gamma'_1 z_{t+1}) E_{t+1} \exp(\gamma'_2 z_{t+2} + \dots + \gamma'_h z_{t+h})] \\ &= E_t[\exp(\gamma'_1 z_{t+1}) e' P'_{t+h}(\gamma_h) \dots P'_{t+2}(\gamma_2) z_{t+1}] \\ &= E_t[e' P'_{t+h}(\gamma_h) \dots P'_{t+2}(\gamma_2) \text{diag}[\exp(\gamma_1)] z_{t+1}], \end{aligned}$$

where $\text{diag}[\exp(\gamma)]$ is the diagonal matrix whose diagonal terms are the exponential of the components of γ . Therefore we have :

$$\varphi_{t,h}^{(z)} = e' P'_{t+h}(\gamma_h) \dots P'_{t+2}(\gamma_2) \text{diag}[\exp(\gamma_1)] \Pi'_t z_t = e' P'_{t+h}(\gamma_h) \dots P'_{t+1}(\gamma_1) z_t.$$

APPENDIX 2 - Proof of Proposition 2

$$\begin{aligned} E_{t-1}[\exp(u' w_{2,t} + v' z_t)] &= E_{t-1}\{\exp(v' z_t) E[\exp(u' w_{2,t}) | w_{2,t-1}, \underline{z}_t, \Delta_t]\} \\ &= E_{t-1}\{\exp[v' z_t + a'_{t-1}(u) w_{2,t-1} + b_{t-1}^{(0)'}(u) \Delta_t z_t]\} \\ &= \exp\{a'_{t-1}(u) w_{2,t-1} + [A_1(u, v), \dots, A_J(u, v)] z_{t-1}\}, \end{aligned}$$

with :

$$A_i(u, v) = \log \sum_{j=1}^J \pi_{i,j,t} \exp[\psi_j(b_{t-1}^{(0)}(u))],$$

$\psi_j(\cdot)$ being the log-Laplace transform of Δ_t^j . Therefore $(w'_{2,t}, z'_t)'$ is Car (1). If Δ_t^j is non random we have $\psi_j(b_{t-1}^{(0)}(u)) = b_{t-1}^{(0)}(u)' \Delta_t^j$.

APPENDIX 3 - Proof of Proposition 3

Let us still omit the exponent h in the $\gamma_j^{(h)}$. For any $j = 1, \dots, h$ we have :

$$(i) \quad \varphi_{t,h}^{(w)} = E_t \left[\exp(\gamma_1' w_{t+1} + \dots + \gamma_j' w_{t+j} + A_{t,h-j}^{(h)'} w_{t+j} + B_{t,h-j}^{(h)}) \right],$$

where :

$$(ii) \quad \begin{cases} A_{t,h-j+1}^{(h)} = a_{t+j} \left(\gamma_j + A_{t,h-j}^{(h)} \right), \\ B_{t,h-j+1}^{(h)} = b_{t+j} \left(\gamma_j + A_{t,h-j}^{(h)} \right) + B_{t,h-j}^{(h)}, \\ A_{t,0}^{(h)} = 0, B_{t,0}^{(h)} = 0. \end{cases}$$

Indeed, we can prove formula (i) by recursion. Formula (i) is true for $j = h$, and, if this is true for j , we get :

$$\varphi_{t,h}^{(w)} = E_t \left[\exp \left(\gamma_1' w_{t+1} + \dots + \gamma_{j-1}' w_{t+j-1} + a_{t+j}' (\gamma_j + A_{t,h-j}^{(h)}) w_{t+j-1} + b_{t+j}' (\gamma_j + A_{t,h-j}^{(h)}) + B_{t,h-j}^{(h)} \right) \right].$$

Therefore formula (i) is true with $j - 1$, $A_{t,h-j+1}^{(h)}$, $B_{t,h-j+1}^{(h)}$ being given by formulas (ii) above. For $j = 1$ we get :

$$\varphi_{t,h}^{(w)} = E_t \exp(\gamma_1' w_{t+1} + A_{t,h-1}^{(h)'} w_{t+1} + B_{t,h-1}^{(h)}) = \exp(A_{t,h}' w_t + B_{t,h}).$$

Finally, if we put $h - j + 1 = i$, formula (i) becomes the formula of Proposition 3.

APPENDIX 4 - Section 2.4 - Tables

	Mean	Std. Dev.	Max.	Min.	Skewness	Kurtosis	$\rho(1)$	$\rho(2)$	$\rho(3)$	$\rho(5)$	$\rho^*(1)$	$\rho^*(2)$	$\rho^*(3)$	$\rho^*(5)$
r_t	0.060	0.028	0.162	0.008	0.913	4.337	0.968	0.933	0.902	0.850	0.950	0.899	0.858	0.795
S_t	0.017	0.014	0.050	-0.033	-0.559	3.065	0.909	0.828	0.756	0.640	0.861	0.742	0.665	0.572
$S_t^{(b)}$	0.012	0.011	0.052	-0.028	-0.296	4.007	0.856	0.750	0.652	0.496	0.769	0.558	0.412	0.326

Table 1: Summary Statistics for the short rate r_t , long-term spread S_t and the butterfly spread $S_t^{(b)}$. $\rho(k)$ and $\rho^*(k)$ denote the empirical autocorrelation at lag k of any given variable and its square, respectively. Yields are in annual basis and observed at monthly frequency from January 1970 to December 2003.

VAR(1) Model											
ν	Φ_1			$(\Omega' \Omega) \times 1000$							
0.00005	0.97735	-0.04220	0.11687	0.00029							
(0.41178)	(61.74365)	(-0.78166)	(1.92042)	(14.26534)							
0.00004	0.01578	1.03261	-0.16021	-0.00022			0.00024				
(0.38346)	(1.10184)	(21.14217)	(-2.90967)	(-12.92216)			(14.26534)				
-0.00014	0.03854	0.19231	0.67026	-0.00016			0.00018			0.00021	
(-1.50857)	(2.84944)	(4.16864)	(12.88779)	(-10.97944)			(12.70913)			(14.26534)	
<i>JB</i> test				$LB(\varepsilon_{1,t})$			$LB(\varepsilon_{2,t})$			$LB(\varepsilon_{3,t})$	
skewness				skewness			skewness			$LB(\varepsilon_{1,t}^2)$	
kurtosis				kurtosis			kurtosis			$LB(\varepsilon_{2,t}^2)$	
joint				joint			joint			$LB(\varepsilon_{3,t}^2)$	
99.81				2661.97			2761.78			26.59	
5.24**				5.24**			5.24**			0.02**	
5.68**				5.68**			5.68**			26.10	
\mathcal{L}	<i>AIC</i>	<i>BIC</i>	<i>HQ</i>	(6.25)	(6.25)	(10.64)	5.68**	0.59**	14.86	50.21	14.09
8036	-39.44	-39.35	-39.41	[7.18]	[7.18]	[12.59]	6.42**	0.59**	16.37	64.49	15.22
VAR(2) Model											
ν	Φ_1			Φ_2			$(\Omega' \Omega) \times 1000$				
0.00007	1.30676	0.08811	0.33884	-0.33239	-0.11569	-0.24873	0.00027				
(0.66392)	(14.96499)	(0.70123)	(3.51229)	(-3.78647)	(-0.90985)	(-2.56166)	(14.24781)				
0.00002	-0.23831	0.90652	-0.28948	0.25702	0.11952	0.14082	-0.00020			0.00023	
(0.17453)	(-2.97444)	(7.86349)	(-3.27036)	(3.19119)	(1.02447)	(1.58062)	(-12.84512)			(14.24781)	
-0.00011	-0.06940	0.35757	0.42061	0.10309	-0.21264	0.30830	-0.00015			0.00017	
(-1.18896)	(-0.92002)	(3.29431)	(5.04689)	(1.35946)	(-1.93589)	(3.67544)	(-10.90475)			(12.73285)	
<i>JB</i> test				$LB(\varepsilon_{1,t})$			$LB(\varepsilon_{2,t})$			$LB(\varepsilon_{3,t})$	
skewness				skewness			skewness			$LB(\varepsilon_{1,t}^2)$	
kurtosis				kurtosis			kurtosis			$LB(\varepsilon_{2,t}^2)$	
joint				joint			joint			$LB(\varepsilon_{3,t}^2)$	
144.51				2922.64			3067.14			9.48	
1.41**				1.41**			1.41**			0.17**	
2.68**				2.68**			2.68**			20.49	
\mathcal{L}	<i>AIC</i>	<i>BIC</i>	<i>HQ</i>	(6.25)	(6.25)	(10.64)	2.68**	0.41**	1.76**	25.58	13.48
8047	-39.55	-39.37	-39.48	[7.18]	[7.18]	[12.59]	4.64**	0.45**	5.83**	37.80	14.89
2-state RS-VAR(1) Model with Homogeneous Markov chain											
ν	Φ_1			$\Omega(e_1)$			$\Omega(e_2)$				
0.000004	0.98285	-0.06197	0.15770	0.00024			0			0.00100	
(0.067835)	(80.70039)	(-1.92512)	(3.86876)	(18.37654)			(12.38928)			0	
0.00006	0.01195	1.04869	-0.18295	-0.00017			0.00025			0	
(0.81148)	(0.97590)	(29.46257)	(-4.19969)	(-8.31523)			(23.45753)			(-11.04684)	
-0.00009	0.03222	0.16583	0.68737	-0.00009			0.00022			0.00019	
(-1.25308)	(2.68108)	(4.60918)	(15.87541)	(-4.06101)			(14.52050)			(21.72434)	
<i>JB</i> test				$LB(\varepsilon_{1,t})$			$LB(\varepsilon_{2,t})$			$LB(\varepsilon_{3,t})$	
skewness				skewness			skewness			$LB(\varepsilon_{1,t}^2)$	
kurtosis				kurtosis			kurtosis			$LB(\varepsilon_{2,t}^2)$	
joint				joint			joint			$LB(\varepsilon_{3,t}^2)$	
3.37056				1.39**			0.01**			9.39	
19.38				19.38			19.38			1.16**	
50.79				50.79			50.79			0.42**	
\mathcal{L}	<i>AIC</i>	<i>BIC</i>	<i>HQ</i>	(6.25)	(6.25)	(10.64)	3.82**	0.47**	16.00	4.69**	12.26
8617	-42.30	-42.21	-42.26	[7.18]	[7.18]	[12.59]	9.97**	0.64**	18.39	5.96**	12.41
2-state RS-VAR(1) Model with Non-Homogeneous Markov chain											
ν	Φ_1			$\Omega(e_1)$			$\Omega(e_2)$				
-0.000004	0.98483	-0.06386	0.16418	0.00024			0			0.00100	
(-0.056919)	(79.83213)	(-1.96155)	(3.92406)	(18.26592)			(12.04753)			0	
0.00007	0.01038	1.05026	-0.18679	-0.00016			0.00024			0	
(0.87560)	(0.83867)	(29.35676)	(-4.19305)	(-8.37592)			(19.38769)			(-10.72024)	
-0.00009	0.03133	0.16686	0.68424	-0.00009			0.00022			0.00019	
(-1.18068)	(2.56405)	(4.60318)	(15.50013)	(-4.17512)			(13.24740)			(20.58206)	
<i>JB</i> test				$LB(\varepsilon_{1,t})$			$LB(\varepsilon_{2,t})$			$LB(\varepsilon_{3,t})$	
skewness				skewness			skewness			$LB(\varepsilon_{1,t}^2)$	
kurtosis				kurtosis			kurtosis			$LB(\varepsilon_{2,t}^2)$	
joint				joint			joint			$LB(\varepsilon_{3,t}^2)$	
6.22445				1.22**			0.00**			8.82	
14.89				14.89			14.89			0.68**	
47.48				47.48			47.48			0.53**	
\mathcal{L}	<i>AIC</i>	<i>BIC</i>	<i>HQ</i>	(6.25)	(6.25)	(10.64)	3.65**	0.57**	15.61	4.78**	11.69
8623	-42.33	-42.24	-42.30	[7.18]	[7.18]	[12.59]	11.69*	0.92**	17.81	6.70**	11.73*

Table 2: Parameter estimates (t -values in parenthesis), Likelihood-based selection criteria and residual tests of single-regime Gaussian VAR(1) and VAR(2) models and of 2-state Gaussian RS-VAR(1) models with homogeneous and non-homogeneous Markov chain. \mathcal{L} denotes the maximum value of the log-likelihood function with associated Akaike (*AIC*), Bayesian (*BIC*) and Hannan-Quin (*HQ*) selection criteria. *JB* stands for Jarque-Bera test for Gaussianity of model residuals (0.10 and 0.05 critical values are in parenthesis and square brackets, respectively). $LB(\varepsilon_{j,t})$ and $LB(\varepsilon_{j,t}^2)$ denote, respectively, the Ljung-Box test on the j^{th} equation model residuals and squared model residuals, for lags $k \in \{1, 2, 3, 5\}$; (**) and (*) denote the null hypothesis accepted at 0.05 and 0.01, respectively.

2-state RS-VAR(2) Model with Homogeneous Markov chain															
ν	Φ_1			Φ_2			$\Omega(e_1)$			$\Omega(e_2)$					
0.00002	1.18254	0.04852	0.23170	-0.19639	-0.09760	-0.11900	0.00024	0	0	0.00097	0	0			
(0.28079)	(16.22693)	(0.52302)	(3.24316)	(-2.66996)	(-1.04213)	(-1.63344)	(18.60039)			(12.17575)					
0.00004	-0.14453	0.93395	-0.19680	0.15573	0.10996	0.04501	-0.00017	0.00024	0	-0.00075	0.00034	0			
(0.47156)	(-1.98779)	(9.64940)	(-2.61763)	(2.11619)	(1.12641)	(0.59604)	(-8.53996)	(20.55545)		(-10.64627)	(12.03146)				
-0.00007	0.05064	0.33891	0.54728	-0.02669	-0.19994	0.18784	-0.00010	0.00022	0.00019	-0.00057	0.00025	0.00041			
(-0.84623)	(0.70463)	(3.51235)	(7.19597)	(-0.37032)	(-2.06380)	(2.42474)	(-4.76163)	(14.28459)	(21.40311)	(-8.46649)	(5.12067)	(12.58897)			
a_1	a_2	<i>JB</i> test					$LB(\varepsilon_{1,t})$	$LB(\varepsilon_{2,t})$	$LB(\varepsilon_{3,t})$	$LB(\varepsilon_{1,t}^2)$	$LB(\varepsilon_{2,t}^2)$	$LB(\varepsilon_{3,t}^2)$			
3.32890	2.09815	skewness kurtosis joint					0.00**	0.00**	0.95**	0.21**	0.39**	2.62**			
(8.73863)	(5.38068)	20.98 38.15 59.12					2.43**	0.28**	1.43**	1.38**	3.18**	4.46**			
\mathcal{L}	<i>AIC</i>	<i>BIC</i>	<i>HQ</i>	(6.25)	(6.25)	(10.64)	2.46**	0.35**	4.07**	2.77**	7.02**	4.67**			
8639	-42.47	-42.29	-42.40	[7.18]	[7.18]	[12.59]	9.64**	0.60**	7.01**	4.28**	7.24**	8.58**			
2-state RS-VAR(2) Model with Non-Homogeneous Markov chain															
ν	Φ_1			Φ_2			$\Omega(e_1)$			$\Omega(e_2)$					
0.00002	1.19022	0.05291	0.23299	-0.20312	-0.10208	-0.11967	0.00024	0	0	0.00098	0	0			
(0.23394)	(16.53269)	(0.59458)	(3.33973)	(-2.80716)	(-1.14396)	(-1.70635)	(18.52565)			(12.29823)					
0.00004	-0.15025	0.92834	-0.19554	0.16059	0.11513	0.04594	-0.00016	0.00024	0	-0.00076	0.00035	0			
(0.49414)	(-2.10294)	(9.95441)	(-2.65125)	(2.23388)	(1.23537)	(0.62634)	(-8.73494)	(19.96848)		(-10.60078)	(12.18160)				
-0.00006	0.04173	0.33621	0.54450	-0.01844	-0.19772	0.19130	-0.00010	0.00021	0.00019	-0.00057	0.00025	0.00041			
(-0.80678)	(0.58855)	(3.55339)	(7.24376)	(-0.26010)	(-2.09319)	(2.52881)	(-4.80968)	(13.80405)	(21.10199)	(-8.42305)	(5.24151)	(12.65863)			
a_1	a_2	b_1	b_2	<i>JB</i> test			$LB(\varepsilon_{1,t})$	$LB(\varepsilon_{2,t})$	$LB(\varepsilon_{3,t})$	$LB(\varepsilon_{1,t}^2)$	$LB(\varepsilon_{2,t}^2)$	$LB(\varepsilon_{3,t}^2)$			
5.86646	-0.82837	-553.75759	412.75813	skewness kurtosis joint			0.00**	0.00**	1.06**	0.21**	0.52**	3.14**			
(3.80661)	(-0.57001)	(-2.13749)	(1.72261)	19.34 31.40 50.75			2.48**	0.23**	1.51**	1.32**	1.61**	4.72**			
\mathcal{L}	<i>AIC</i>	<i>BIC</i>	<i>HQ</i>	(6.25)	(6.25)	(10.64)	2.51**	0.29**	4.02**	2.85**	7.32**	4.98**			
8647	-42.51	-42.33	-42.44	[7.18]	[7.18]	[12.59]	9.42**	0.63**	6.22**	4.38**	7.75**	9.56**			
3-state RS-VAR(1) Model with Homogeneous Markov chain															
ν	Φ_1			Φ_2			$\Omega(e_1)$			$\Omega(e_2)$			$\Omega(e_3)$		
0.00002	0.98510	-0.05050	0.12593				0.00014	0	0	0.00036	0	0	0.00114	0	0
(0.33526)	(85.39781)	(-1.67324)	(3.20251)				(6.46570)			(9.12401)			(10.64022)		
0.00004	0.01066	1.04387	-0.16384				-0.00007	0.00021	0	-0.00028	0.00028	0	-0.00086	0.00033	0
(0.58081)	(0.89304)	(30.59122)	(-3.83727)				(-2.13760)	(11.42970)		(-7.26357)	(12.35689)		(-9.61827)	(10.66865)	
-0.00009	0.03032	0.14811	0.71517				-0.00002	0.00021	0.00015	-0.00019	0.00021	0.00024	-0.00065	0.00027	0.00045
(-1.22716)	(2.50378)	(4.15092)	(16.48013)				(-0.44157)	(9.76395)	(11.58300)	(-5.60653)	(8.72646)	(11.56677)	(-7.27891)	(4.24445)	(10.98500)
π_{11}	π_{21}	π_{22}	π_{33}	<i>JB</i> test			$LB(\varepsilon_{1,t})$	$LB(\varepsilon_{2,t})$	$LB(\varepsilon_{3,t})$	$LB(\varepsilon_{1,t}^2)$	$LB(\varepsilon_{2,t}^2)$	$LB(\varepsilon_{3,t}^2)$			
0.87076	0.10273	0.86876	0.91341	skewness kurtosis joint			0.98**	0.05**	8.77	2.03**	0.43**	0.82**			
(17.14685)	(1.02079)	(10.63185)	(22.78022)	2.22 8.69 10.91			1.83**	1.04**	9.41	2.37**	6.11*	1.21**			
\mathcal{L}	<i>AIC</i>	<i>BIC</i>	<i>HQ</i>	(6.25)	(6.25)	(10.64)	1.86**	1.28**	12.40	2.63**	9.88*	1.93**			
8646	-42.44	-42.35	-42.41	[7.18]	[7.18]	[12.59]	8.93**	1.62**	16.46	3.79**	10.36**	4.45**			
3-state RS-VAR(2) Model with Homogeneous Markov chain															
ν	Φ_1			Φ_2			$\Omega(e_1)$			$\Omega(e_2)$			$\Omega(e_3)$		
0.00004	1.15427	0.05727	0.20007	-0.17123	-0.11377	-0.07671	0.00018	0	0	0.00040	0	0	0.00121	0	0
(0.65449)	(18.12211)	(0.71462)	(2.82210)	(-2.68462)	(-1.41497)	(-1.11456)	(9.16655)			(7.89416)			(8.43522)		
0.00002	-0.11814	0.93051	-0.17956	0.13132	0.11870	0.01632	-0.00010	0.00022	0	-0.00033	0.00029	0	-0.00091	0.00034	0
(0.30549)	(-1.76092)	(10.28574)	(-2.34372)	(1.95909)	(1.31421)	(0.22003)	(-3.72936)	(11.86146)		(-6.73151)	(11.92797)		(-7.87560)	(9.23665)	
-0.00007	0.07301	0.31790	0.57723	-0.04752	-0.18814	0.16154	-0.00006	0.00021	0.00015	-0.00022	0.00021	0.00028	-0.00070	0.00032	0.00043
(-0.85942)	(1.03184)	(3.43449)	(7.41662)	(-0.67417)	(-2.02197)	(2.06510)	(-2.19272)	(10.82625)	(13.57273)	(-5.03637)	(5.93078)	(8.42813)	(-6.20111)	(3.88267)	(8.59047)
π_{11}	π_{21}	π_{22}	π_{33}	<i>JB</i> test			$LB(\varepsilon_{1,t})$	$LB(\varepsilon_{2,t})$	$LB(\varepsilon_{3,t})$	$LB(\varepsilon_{1,t}^2)$	$LB(\varepsilon_{2,t}^2)$	$LB(\varepsilon_{3,t}^2)$			
0.89203	0.10274	0.83085	0.92723	skewness kurtosis joint			0.40**	0.03**	1.07**	1.01**	0.59**	3.52**			
(21.31069)	(0.83291)	(5.76688)	(21.15464)	1.95 5.98 7.93			0.42**	0.65**	1.14**	2.41**	4.90**	5.79**			
\mathcal{L}	<i>AIC</i>	<i>BIC</i>	<i>HQ</i>	(6.25)	(6.25)	(10.64)	0.42**	0.88**	6.01**	3.28**	7.90*	8.35*			
8667	-42.60	-42.43	-42.53	[7.18]	[7.18]	[12.59]	10.84**	1.33**	12.28**	4.03**	8.75**	9.93**			

Table 3: Parameter estimates (t -values in parenthesis), Likelihood-based selection criteria and residual tests of 2-state Gaussian RS-VAR(2) models with homogeneous and non-homogeneous Markov chain and of 3-state Gaussian RS-VAR(1) and RS-VAR(2) models with homogeneous Markov chain. \mathcal{L} denotes the maximum value of the log-likelihood function with associated Akaike (*AIC*), Bayesian (*BIC*) and Hannan-Quin (*HQ*) selection criteria. *JB* stands for Jarque-Bera test for Gaussianity of model residuals (0.10 and 0.05 critical values are in parenthesis and square brackets, respectively). $LB(\varepsilon_{j,t})$ and $LB(\varepsilon_{j,t}^2)$ denote, respectively, the Ljung-Box test on the j^{th} equation model residuals and squared model residuals, for lags $k \in \{1, 2, 3, 5\}$; (**) and (*) denote the null hypothesis accepted at 0.05 and 0.01, respectively.

APPENDIX 5 - Proof of Proposition 4

$$\begin{aligned}
B(t, h) &= \exp(-r_t) E^{\mathbb{Q}}[\exp(-r_{t+1} \dots - r_{t+h-1}) | \underline{r}_t, \underline{y}_t] \\
&= \exp(-r_t) E^{\mathbb{Q}}\{E^{\mathbb{Q}}[\exp(-r_{t+1} \dots - r_{t+h-1}) | \underline{z}_{t+h-1}, \underline{r}_{t+h-2}, \underline{y}_{t+h-2}] | \underline{r}_t, \underline{y}_t\} \\
&= \exp(-r_t) E^{\mathbb{Q}}\{\exp(\tilde{\gamma}'_{t+h-1} z_{t+h-1} - r_{t+1} \dots - r_{t+h-2}) | \underline{r}_t, \underline{z}_t\} \\
&= \exp(-r_t) E^{\mathbb{Q}}\{\exp(\tilde{\gamma}'_{t+h-1} z_{t+h-1}) E^{\mathbb{Q}}[\exp(-r_{t+1} \dots - r_{t+h-2}) | \underline{z}_{t+h-1}, \underline{r}_{t+h-3}] | \underline{r}_t, \underline{y}_t\}
\end{aligned}$$

Using the non causality from (r_t, y_t) to z_t , we can replace \underline{z}_{t+h-1} by \underline{z}_{t+h-2} in the conditioning and get :

$$B(t, h) = \exp(-r_t) E^{\mathbb{Q}}\{\exp(\tilde{\gamma}'_{t+h-1} z_{t+h-1} + \tilde{\gamma}'_{t+h-2} z_{t+h-2} - r_{t+1} \dots - r_{t+h-3}) | \underline{r}_t, \underline{y}_t\},$$

and, by recursion :

$$B(t, h) = \exp(-r_t) E^{\mathbb{Q}}\{\exp(\tilde{\gamma}'_{t+1} z_{t+1} + \dots + \tilde{\gamma}'_{t+h-1} z_{t+h-1}) | \underline{r}_t, \underline{y}_t\}.$$

Conditioning first by $z_t, \underline{r}_t, \underline{y}_t$ and using Proposition 1, we get :

$$B(t, h) = \exp(-r_t) E^{\mathbb{Q}}\{e' P'_{t+h-1}(\tilde{\gamma}_{t+h-1}) \dots P'_{t+1}(\tilde{\gamma}_{t+1}) z_t | \underline{r}_t, \underline{y}_t\}$$

with $P_t(\gamma) = \Pi_t \text{diag}[\exp(\gamma)]$. Finally :

$$B(t, h) = e' P'_{t+h-1}(\tilde{\gamma}_{t+h-1}) \dots P'_{t+1}(\tilde{\gamma}_{t+1}) \hat{z}_t \exp(-r_t).$$

APPENDIX 6 - Proof of Proposition 5

Let us consider the case where the conditional distribution of $r_{t+1} = e_1$ is the point mass at zero, and define the p.d.f. of $(z_{t+1}, z_t, r_{t+1} y_{t+1})$ given $\underline{r}_t, \underline{z}_t$, with respect to the measure $(\sum_{j=1}^J \delta_j)^{\otimes 2} \otimes (\delta_0 +$

$\lambda_1) \otimes \lambda_1^K$, J being the number of states in the Markov chain z_t , K the size of y_t , $\delta_j, j = 1, \dots, J$ the unit point mass at e_j and δ_0 the unit point mass at 0. This p.d.f can be factorized as :

$$q_t(z_{t+1}|z_t)f_t(r_{t+1}|z_{t+1})g_t(y_{t+1}|z_{t+1}, \underline{r_{t+1}}, \underline{y_t})p_t(z_t|\underline{r_t}, \underline{y_t}),$$

where $q_t(z_{t+1}|z_t)$, $f_t(r_{t+1}|z_{t+1})$, $g_t(y_{t+1}|z_{t+1}, \underline{r_{t+1}}, \underline{y_t})$, $p_t(z_t|\underline{r_t}, \underline{y_t})$ denote the conditional p.d.f. with respect to the appropriate measure. In particular, we have $f_{1,t}(0) = 1$ and $f_{1,t}(r) = 0, \forall r \neq 0$. Therefore, we get :

$$p_{t+1}(z_{t+1}|\underline{r_{t+1}}, \underline{y_{t+1}}) = \frac{\sum_{z_t} q_t(z_{t+1}|z_t)f_t(r_{t+1}|z_{t+1})g_t(y_{t+1}|z_{t+1}, \underline{r_{t+1}}, \underline{y_t})}{\sum_{z_{t+1}} \sum_{z_t} q_t(z_{t+1}|z_t)f_t(r_{t+1}|z_{t+1})g_t(y_{t+1}|z_{t+1}, \underline{r_{t+1}}, \underline{y_t})p_t(z_t|\underline{r_t}, \underline{y_t})}.$$

Stacking the different value of $p_{t+1}(e_j|\underline{r_{t+1}}, \underline{y_{t+1}}) = \hat{z}_{j,t+1}$:

$$\hat{z}_{t+1} = \frac{\text{diag}(f_t g_t) \Pi_t' \hat{z}_t}{e' \text{diag}(f_t g_t) \Pi_t' \hat{z}_t},$$

where $\text{diag}(f_t g_t)$ is the diagonal matrix whose k^{th} diagonal element is the product of $f_{k,t}(r_{t+1}) = f_t(r_{t+1}|e_k)$ by $g_{k,t}(y_{t+1}|\underline{r_{t+1}}, \underline{y_t}) = g_t(y_{t+1}|e_k, \underline{r_{t+1}}, \underline{y_t})$.

APPENDIX 7

Proof of Proposition 6

$$\pi_{i,j,t} = \mathbb{Q}(z_t = e_j | z_{t-1} = e_i) = \int \mathbb{Q}(z_t = e_j | r_t, y_t, z_{t-1} = e_i) \alpha_{i,t} \alpha_{0,t} d\mu = E_{0,t}(\alpha_{i,t} \beta_{j,t})$$

$$B(t, h) = \exp(-r_t) E_t^{\mathbb{Q}}[\exp(-r_{t+1} - \dots - r_{t+h-1})].$$

We have to show that, for $h \geq 2$:

$$E_t^{\mathbb{Q}}[\exp(-r_{t+1} - \dots - r_{t+h-1})] = e' \tilde{P}'_{t+h-1} \dots \tilde{P}'_{t+1} \beta_t, \forall t. \quad (a)$$

The formula is true for $h = 2$, since :

$$E_t^{\mathbb{Q}}[\exp(-r_{t+1})] = E_{0,t+1}[\exp(-r_{t+1})\alpha'_{t+1}]\beta_t = e' E_{0,t+1}[\exp(-r_{t+1})\beta_{t+1}, \alpha'_{t+1}]\beta_t = e' \tilde{P}'_{t+1}\beta_t,$$

(since $e'\beta_{t+1} = 1$).

Let us assume that formula (a) is valid for $h - 1$, we get :

$$\begin{aligned} E_t^{\mathbb{Q}}[\exp(-r_{t+1} - \dots - r_{t+h-1})] &= E_t^{\mathbb{Q}}[\exp(-r_{t+1})e' \tilde{P}'_{t+h-1} \dots \tilde{P}'_{t+2}\beta_{t+1}] \\ &= e' \tilde{P}'_{t+h-1} \dots \tilde{P}'_{t+2} E_t^{\mathbb{Q}}[\exp(-r_{t+1})\beta_{t+1}] \\ &= e' \tilde{P}'_{t+h-1} \dots \tilde{P}'_{t+2} E_{0,t+1}[\exp(-r_{t+1})\beta_{t+1}\alpha'_{t+1}\beta_t] \\ &= e' \tilde{P}'_{t+h-1} \dots \tilde{P}'_{t+1}\beta_t. \end{aligned}$$

APPENDIX 8 - Proof of Proposition 7

Let us consider the FDD historical dynamics defined by the conditional p.d.f. $\alpha_{0,t}(w_t)\tilde{\alpha}'_t(w_t)\tilde{\beta}_{t-1}(w_{t-1})$.

In this case the s.d.f. is :

$$M_{t-1,t} = \frac{\alpha'_t(w_t)\beta_{t-1}(w_{t-1})}{\tilde{\alpha}'_t(w_t)\tilde{\beta}_{t-1}(w_{t-1})} \exp(-r_{t-1}).$$

Conversely, let us consider a s.d.f. of the form :

$$M_{t-1,t}(w_t, w_{t-1}) = M_{1,t-1,t}(w_t) M_{2,t-1,t}(w_{t-1})$$

satisfying $E_{t-1}^{\mathbb{Q}}(M_{t-1,t}^{-1}) = \exp(r_{t-1})$. The historical conditional p.d.f. of w_t given w_{t-1} is given by :

$$M_{t-1,t}^{-1} \alpha_{0,t} \alpha'_t \beta_{t-1} \exp(-r_{t-1})$$

Let us define the p.d.f., w.r.t. the distribution $\alpha_{0,t}$:

$$\tilde{\alpha}_{j,t} = \frac{\alpha_{j,t} M_{1,t-1,t}^{-1}}{E_{0,t}(\alpha_{j,t} M_{1,t-1,t}^{-1})} \quad (4.8)$$

and the probabilities :

$$\tilde{\beta}_{j,t-1} = \beta_{j,t-1} M_{2,t-1,t}^{-1} E_{0,t}(\alpha_{j,t} M_{1,t-1,t}^{-1}) \exp(-r_{t-1}) \quad (4.9)$$

which are summing to one since :

$$\sum_{j=1}^J \tilde{\beta}_{j,t-1} = \exp(-r_{t-1}) E_{0,t}(\alpha'_t \beta_{t-1} M_{t-1,t}^{-1}) = \exp(-r_{t-1}) E_{t-1}^{\mathbb{Q}}(M_{t-1,t}^{-1}) = 1.$$

We have $M_{t-1,t}^{-1} \alpha_{0,t} \alpha'_t \beta_{t-1} \exp(-r_{t-1}) = \alpha_{0,t} \tilde{\alpha}'_t \tilde{\beta}_{t-1}$. \square

APPENDIX 9 - Proof of Proposition 8

i) By definition the price of the defaultable zero-coupon bond with zero recovery rate is $B_n(t, h) = E_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{t+h-1})(1 - d_{t+h}^n)]$. Conditioning with respect to $\underline{\tilde{w}}_{t+h}$ and using Bayes formula, we get :

$$B_n(t, h) = \exp(-r_t) E_t^{\mathbb{Q}}\{\exp(-r_{t+1} - \dots - r_{t+h-1}) \times \prod_{j=1}^h \mathbb{Q}(d_{t+j}^n = 0 \mid d_{t+j-1}^n = 0, \underline{\tilde{w}}_{t+h})\}.$$

Since $\{d_t\}$ does not cause $\{\tilde{w}_t\}$ we can replace $\underline{\tilde{w}}_{t+h}$ by $\underline{\tilde{w}}_{t+j}$ in the generic term of the product.

Finally, we get:

$$B_n(t, h) = \exp(-r_t) E_t^{\mathbb{Q}}[\exp(-r_{t+1} - \dots - r_{t+h-1} - \lambda_{n,t+1}^{\mathbb{Q}} - \dots - \lambda_{n,t+h}^{\mathbb{Q}})].$$

ii) Formula *i)* is obtained by replacing r_t and $\lambda_{n,t}^{\mathbb{Q}}$ by their expressions given in Assumptions A.4 and A.5.

APPENDIX 10 - Term structure of recovery adjusted defaultable bonds

If the recovery payoff, when issuer n defaults between $t-1$ and t , is equal to a fraction $F_{n,t}$ (function of \tilde{w}_t^n) of the price that would have prevailed without default, $B_n(t, h)$ can still be computed in

the same way as in Proposition 8, provided that the R.N. default intensity $\lambda_{nt}^{\mathbb{Q}}$ is replaced by a R.N. "recovery adjusted" default intensity $\tilde{\lambda}_{n,t}^{\mathbb{Q}}$ defined by :

$$\exp(-\tilde{\lambda}_{n,t}^{\mathbb{Q}}) = \exp(-\lambda_{n,t}^{\mathbb{Q}}) + [1 - \exp(-\lambda_{n,t}^{\mathbb{Q}})] F_{n,t}.$$

The quantity $\exp(-\tilde{\lambda}_{n,t}^{\mathbb{Q}})$ represents the short term R.N. expected gain. If there is no expected default, the recovery rate is equal to 1, which corresponds to the first component. If there is an expected default, with probability $1 - \exp(-\lambda_{n,t}^{\mathbb{Q}})$, the recovery rate is the contractual market value $F_{n,t}$. If $F_{n,t} = 0$, we get the previous model $\tilde{\lambda}_{n,t}^{\mathbb{Q}} = \lambda_{n,t}^{\mathbb{Q}}$ and, if $F_{n,t} = 1$, we get $\tilde{\lambda}_{n,t}^{\mathbb{Q}} = 0$, that is, the default-free case. If $\lambda_{n,t}^{\mathbb{Q}}$ is small, we get $\tilde{\lambda}_{n,t}^{\mathbb{Q}} \simeq \lambda_{n,t}^{\mathbb{Q}}(1 - F_{n,t})$.

REFERENCES

- Ang, A., and G., Bekaert (2002)a: "International Asset Allocation with Regime Shifts", *Review of Financial Studies*, 15, 1137-1187.
- Ang, A., and G., Bekaert (2002)b: "Regime Switches in Interest Rates", *Journal of Business and Economic Statistics*, 20, 163-182.
- Ang, A., and G., Bekaert (2002)c:"Short Rate Nonlinearities and Regime Switches", *Journal of Economic Dynamics and Control*, 26, 1243-1274.
- Ang, A., and G., Bekaert (2004): "How do Regimes Affect Asset Allocations", *Financial Analysts Journal*, 60, 86-99.
- Ang, A., Bekaert, G., and M., Wei (2008): "The Term Structure of Real Rates and Expected Inflation", *Journal of Finance*, 63, 797-849.
- Ang, A., and J., Chen (2002): "Asymmetric Correlations of Equity Portfolios", *Journal of Financial Economics*, 63, 443-434.
- Ang, A., and A., Timmerman (2011): "Regime Changes and Financial Markets", CEPR Discussion Paper, 8480.
- Backus, D.K., S. Foresi and C.I. Telmer (2001): "Affine Term Structure Models and the Forward Premium Anomaly" , *Journal of Finance*, 51, 279-304.
- Balduzzi, P., Bertola, G., and S. Foresi (1997): "A model of target changes and the term structure of interest rates," *Journal of Monetary Economics*, 39 (2), 223-249.
- Bansal, R., and H., Zhou (2002): "Term Structure of Interest Rates with Regime Shifts", *Journal of Finance*, 57, 1997-2043.
- Bekaert, G., and R., Hodrick (1993): "On Biases in the Measurement of Foreign Exchange Risk

- Premium”, *Journal of International Money and Finance*, 12, 115-138.
- Bertholon, H., Monfort, A., and F., Pegoraro (2008): ”Econometric Asset Pricing Modelling”, *Journal of Financial Econometrics*, 4, 407-458.
- Billio, M., Getmansky, M., Lo, A, and L., Pelizzon (2012): ”Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors”, *Journal of Financial Economics*, 104, 535-559.
- Billio, M. and L., Pelizzon (2000): ”Value-at-Risk: a Multivariate Switching Regime Approach”, *Journal of Empirical Finance*, 7, 531-554.
- Brennan, M.J., and Y. Xia (2006): ”International Capital Markets and Foreign Exchange Risk”, *Review of Financial Studies*, 19, 753-795.
- Chib, S. and K., Kang (2013): ”Change Point in Affine Arbitrage-Free Term Structure Models”, *Journal of Financial Econometrics*, 11, 302-334.
- Dai, Q., Singleton, K., and W., Yang (2007): ”Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields”, *Review of Financial Studies*, 20, 1669-1706.
- Darolles, S., Gouriou, C., and J., Jasiak (2006): ”Structural Laplace Transform and Compound Autoregressive Models”, *Journal of Time Series Analysis*, 27, 477-503.
- Duffie, D., Pan, J., and K., Singleton (2000): ”Transform Analysis and Asset Pricing for Affine Jump-Diffusions”, *Econometrica*, 68, 6, 1343-1376.
- Duffie, D., and K., Singleton (1999): ”Modeling Term Structure of Defaultable Bonds”, *Review of Financial Studies*, 12, 687-720.
- Engel, C., and J., Hamilton (1990): ”Long Swings in the Dollar : Are they in the Data and Do Markets Know it ?”, *American Economic Review*, 80, 689-713.

- Fama, E. and Bliss, R. (1987): "The Information in Long-Maturity Forward Rates", *American Economic Review*, 77, 680-692.
- Feldhütter, P., and D., Lando (2008): "Decomposing Swap Rates", *Journal of Financial Economics*, 88, 375-405.
- Fontaine, J.-S. (2009): "Fed funds futures and the federal reserve", Working Papers, Bank of Canada.
- Fontaine, J.-S., and Garcia, R., (2012): "Bond Liquidity Premia", *The Review of Financial Studies*, 25 (4), 1207-1254.
- Garcia, R., and P., Perron (1996): "An Analysis of the Real Interest Rate Under Regime Shifts", *The Review of Economics and Statistics*, 78, 111-125.
- Gourieroux, C., and J., Jasiak (2000): "State Space Model with Finite Dimensional Dependence", *Journal of Time Series Analysis*, 22, 665-678.
- Gourieroux, C., and A., Monfort (2007): "Econometric Specification of Stochastic Discount Factor Models", *Journal of Econometrics*, 136, 509-530.
- Gourieroux, C., Monfort, A., and V., Polimenis (2006): "Affine Model for Credit Risk Analysis", *Journal of Financial Econometrics*, 5, 103-151.
- Gourieroux, C., Monfort, A., and Sufana, R., (2010): "International Money and Stock Market Contingent Claim", *Journal of International Money and Finance*, 29 (8), 1727-1751.
- Gray, S. (1996): "Modelling the Conditional Distribution of Interest Rates as Regime Switching Process", *Journal of Financial Economics*, 42, 27-308.
- Graveline, J., and Joslin, S., (2011): "G10 and Exchange Rates", Working Paper, University of Minnesota Carlson School of Management, and MIT Sloan School of Management.

- Guidolin, M., and A., Timmermann (2008): "Size and Value Anomalies Under Regime Shifts", *Journal of Financial Econometrics*, 6, 1-48.
- Hamilton, J. (1988): "Rational Expectations Econometrics Analysis of Changes in Regime : An Investigation of the Term Structure of Interest Rates", *Journal of Economic Dynamics and Control*, 12, 385-423.
- Hamilton, J., and R., Susmel (1994): "Autoregressive Conditional Heteroscedasticity and Changes in Regime", *Journal of Econometrics*, 64, 307-333.
- Hansen, L., and S., Richard (1987): "The Role of Conditioning Information in Deducing Testable Restriction Implied by Dynamic Asset Pricing Models", *Econometrica*, 55, 587-613.
- Harrison ,M. and D., Kreps (1979): "Martingales and Arbitrage in Multiperiod Securities Markets", *Journal of Economic Theory*, 20, 381-408.
- Huisman, R. and R., Mahieu (2003): "Regime Jumps in Electricity Prices", *Energy Economics*, 25 (5), 425-424.
- Jardet, C., Monfort, A., and F., Pegoraro (2013): "No-Arbitrage Near-Cointegrated VAR(p) Term Structure Models, Term Premia and GDP Growth", *Journal of Banking and Finance*, 37, 389-402.
- Jarrow, R., Lando, D. and S., Turnbull (1997): "A Markov Model for the Term Structure of Credit Risk Spreads", *Review of Financial Studies*, 102, 481-523.
- Le, A., Singleton, K., and Q. Dai (2010): "Discrete-Time Affine^Q Term Structure Models with Generalized Market Prices of Risk", *Review of Financial Studies*, 23, 2184-2227.
- Leippold, M. and L.Wu (2007): "Design and Estimation of Multi-Currency Quadratic Models", *Review of Finance*, 11, 167-207.
- Liu, J., Longstaff, F. and R. Mandell (2006): "The Market Price of Risk in Interest Rate Swaps : The Roles of Default and Liquidity Risks", *Journal of Business*, 79 (5), 2337-2360.

- Monfort, A. and O., Féron (2012): "Joint Econometric Modeling of Spot Electricity Prices, Forwards and Options", *Review of Derivatives Research*, 15, 217-256.
- Monfort, A., and F., Pegoraro (2007): "Switching VARMA Term Structure Models", *Journal of Financial Econometrics*, 5, 105-153.
- Monfort, A., and F., Pegoraro (2012): "Asset Pricing with Second-Order Esscher Transforms", *Journal of Banking and Finance*, 36, 1678-1687.
- Monfort, A., and J.P., Renne (2011): "Credit and Liquidity Risks in Euro-Area Sovereign Yield Curves", *CREST-DP*, 2011-26.
- Monfort, A., and J.P., Renne (2013): "Default, Liquidity, and Crises : an Econometric Framework", *Journal of Financial Econometrics*, 11 (2), 221-262.
- Mount, T., Ning, Y. and X., Cai (2005): "Predicting Price Spikes in Electricity Markets using a Regime Switching Model with Time-Varying Parameters", *Energy Economics*, 28, 2-80.
- Piazzesi, M. (2005): "Bond Yields and the Federal Reserve", *The Journal of Political Economy*, 113 (2), 311-344.
- Renne, J.P. (2012): "A Model of the Euro-Area Yield Curve with Discrete Policy Rates", *Banque de France*, DP 395.
- Rudebusch, G. (1995): "Federal Reserve interest rate targeting, rational expectations, and the term structure", *Journal of Monetary Economics*, 35 (2), 245-274.
- Tu, J. (2010): "Is Regime Switching in Stock Returns Important in Portfolio Decisions ?", *Management Science*, 56, 1198-1215.