6 Compound auto-regressive processes and defaultable bond pricing

Alain Monfort and Jean-Paul Renne

6.1 Introduction

Various yield spreads have attracted a lot of attention since the onset of the current financial crisis. For instance, (a) the spreads between interbank unsecured rates and the overnight index swaps – the so-called LIBOR-OIS spreads – gauging market concerns regarding banks' solvency and liquidity, (b) the spreads between corporate bonds and their Treasuries counterparts and, more recently, (c) sovereign spreads can all be seen as thermometers for the intensity of the crisis developments.[1] These spreads reflect market-participants' assessment of the risks ahead and therefore contain information that is key for both policymakers and investors. In particular, meaningful information is embedded in the *term-structure* of those spreads and in its dynamics. In order to optimally extract this information, one has to rely on term-structure models. In several respect, the ongoing financial crisis has highlighted the limits of many dynamic term-structure models, notably those that are not able to accommodate non-linearities, stochastic volatilities or switching regimes.

The aim of the present chapter is to propose a general and tractable strategy to model the dynamics of the term structure of defaultable-bond yields. To achieve this, we rely on the properties of compound auto-regressive (Car) processes. The usefulness of these processes in the building of risk-free (non-defaultable) bond pricing models is now well documented (see Darolles, Gourieroux and Jasiak, 2006, Gourieroux and Monfort, 2007, Monfort and Pegoraro, 2007 or Le, Singleton and

We are grateful to Alain Durré and to an anonymous referee for helpful comments and suggestions. The views expressed in this chapter are ours and do not necessarily reflect the views of the Banque de France.

[1] For the LIBOR-OIS spread, see e.g. Taylor and Williams (2009) or Sengupta and Tam (2008); for corporate credit spreads, see e.g. Gilchrist and Zakrajšek (2011); for sovereign spreads see e.g. Borgy et al. (2011), Longstaff et al. (2011) or Monfort and Renne (2012).

Da, 2011). On the contrary, the number of papers using Car processes in defaultable-bond pricing models is small.[2]

Beyond the search for a good fit to the data, our approach is aimed at exploring potential credit-risk premia present in bond yields. Such risk premia are likely to enter bond prices as soon as the underlying credit risk – that reflects the risk that the issuer defaults – can not be diversified away and is correlated with investors' utility. Because of risk premia, extracting so-called "market expectations" from asset prices under the assumption that investors are risk neutral is misleading. As regards risk-free bond yields, risk premia account for the failure of the expectation hypothesis (EH). Under the latter, long-term yields should be equal to the expectation of future short-term ones (till bond maturity). Most empirical studies, however, suggest that this assumption does not hold (see e.g. Campbell and Shiller, 1991) and the difference between observed long-term bonds and the yields that would prevail under the EH are attributed to risk premia that are called *term premia*. The same kind of hypothesis tends to be rejected in the case of credit spreads: for instance, Huang and Huang (2002) show that only a small part of average credit spreads can be accounted for by expectations of default-related loss rates, pointing to the existence of risk premia associated with credit risk.

Our framework is consistent with the existence of such *credit risk premia* in credit spreads and further, it allows us to study the dynamics of these. These risk premia stem from the specification of a stochastic discount factor (sdf). In that context, the physical and the risk-neutral dynamics of the pricing factors – and notably the default process – do not coincide. The risk-neutral dynamics is the dynamics of the pricing factors that would be consistent with observed prices under the (potentially false) assumption that investors are risk-neutral. In our framework, we can assess the size of the (potential) errors that are implied by assuming that the historical and the risk-neutral dynamics coincide. A typical example lies in the computation of market-based probabilities of default (PDs). To get these, the vast majority of practitioners or market analysts resort to approaches ending up with risk-neutral PDs.[3] While risk-neutral PDs are relevant for pricing purposes, historical ones are

[2] See Gourieroux, Monfort and Polimenis (2006), Monfort and Renne (2012) and (2013), Gourieroux et al. (2012).

[3] Most of these methodologies build on Litterman and Iben (1991), see e.g. (amongst many others) Bank of England (2012), CMA Datavision (2011) and O'Kane and Turnbull (2003). Studies resorting to these methods are usually silent about this caveat. Notable recent exceptions include Blundell-Wignall and Slovik (2010), in an OECD study, who note: "In the real world, actual defaults are fewer than market-driven default probability calculations would indicate. That is because market participants demand

needed (a) if one wants to extract real-world investors' perception of the credit quality of the issuer, (b) for the sake of forecasting or more generally (c) for risk management purposes. Regarding the latter point, note for instance that value-at-risk measures (VaR) should be based on the real-world measure and not on the risk-neutral one (see Gourieroux and Jasiak, 2009).

Being able to identify risk premia is important from a policy perspective. In particular, in the context of the ongoing financial crisis, Longstaff et al. (2011) stress the key importance of a better understanding of the so-called sovereign risk. To that respect, Borri and Verdelhan (2011) develop a theoretical model that highlights the central place of risk premia in the derivation of optimal borrowing and default decisions by sovereign entities. Exploiting the properties of Car processes, Monfort and Renne (2012) show that the present is appropriate to model the joint dynamics of euro-area sovereign spreads. Their results point to the importance of the risk premia to account for the dynamics of these spreads. In particular, during stress periods, these premia translate into wide deviations between risk-neutral and physical probabilities of default. In the present chapter, an empirical section supports Monfort and Renne's (2012) result by presenting an estimated model of Spanish sovereign spreads (versus Germany). Such findings are of significant interest in the current context where regulators want banks to model the *actual* default risk of even high-rated government bonds.[4]

In order to emphasise the fact that the appealing properties of the Car processes are needed in the risk-neutral world, but not necessarily in the historical world, we adopt a back modelling approach (see Bertholon, Monfort and Pegoraro, 2008), in which the risk-neutral dynamics and the short rate are specified in the first place, and where the historical dynamics is obtained through the specification of the stochastic discount factor. We show that we can obtain quasi-explicit formulas for the bond prices and therefore for the yields, even if the whole state vector is not Car in the risk-neutral world. Only the subvector appearing in the specification of the short rate and of the default intensities has to be Car. In particular, if we assume that these variables are not directly impacted by the individual default events, we obtain appealing linear formulas for the risky yields, both in the cases of zero and non-zero recovery rates.

a risk premium – an excess return – compared to the risk-neutral rate, and that premium cannot be observed. This makes it difficult to use the above measure [the risk-neutral PDs] to imply the likelihood of actual defaults in the periphery of Europe or anywhere else."

[4] Regulators' views are expressed e.g. by Hannoun (2011) or Nouy (2011). As stressed by Carver (2012), these changes in regulation reveal the practitioners' lack of tools to extract actual default probabilities from market prices.

Once this modelling of the risk-neutral dynamics and of the short rate is done, and once possible internal consistency conditions are taken into account, the modelling of the historical dynamics through the stochastic discount rate is completely free and, hence, very flexible. In such a context, one can fully benefit from the discrete-time framework to ensure a good fitting of the data and to model potential interactions between the pricing factors. Modelling interactions between pricing factors is easier in discrete time than in continuous time since discrete-time models are much more flexible than continuous-time models (see e.g. Le et al., 2010 or Gourieroux et al., 2006). Working with discrete-time models necessitates choosing a time unit adapted to the objective of the model and to the data. A specification for a given time unit implies a dynamics for all the time units which are multiples of the basic time unit, but not for the other time units, in particular those which are smaller than the basic one, contrary to the continuous-time specification. However, the implicit assumption of the continuous-time approach, namely that the dynamics corresponding to all time units (from the minute to the year, for instance) can be derived from a unique specification, is highly questionable. Moreover, the discrete versions of the continuous-time models are in general intractable, except in simple cases without practical interest, and are replaced by approximations for which the time consistency is lost. Finally, the discrete-time models obtained from a discretisation of a continuous-time model are often poor compared to those which can be introduced directly, like the ones proposed in this chapter.

The remaining of this chapter is organised as follows. Section 6.2 gives the definition of compound autoregressive processes and develops a recursive algorithm to compute multi-horizon Laplace transforms of the processes (which is key to deriving the term structure of yields). Section 6.3 details the risk-neutral dynamics and its pricing implications. After having introduced the stochastic discount factor, Section 6.4 derives the implied historical dynamics of the processes. Section 6.5 provides examples of Car processes. Section 6.6 presents possible estimation strategies. Section 6.7 proposes an application to the modelling of the term structure of Spanish sovereign spreads and Section 6.8 concludes.

6.2 Compound autoregressive processes

6.2.1 Definition

Let us define the compound autoregressive (Car) processes. An n-dimensional process w_t is called Car(p) if its conditional log-Laplace transform $\psi_{t-1}(u) = \ln E_{t-1} e^{u' w_t}$ is of the form

$$\psi_{t-1}(u) = a_1(u)'w_{t-1} + \cdots + a_p(u)'w_{t-p} + b(u), \qquad u \in \mathbb{R}^n,$$

where the a_i and b are some $\mathbb{R}$-valued functions and where E_{t-1} denotes the conditional expectation, given the past values of w_t: $\{w_{t-1}, w_{t-2}, \ldots, w_1\}$. It is straightforward to show that if w_t is Car(p), then $(w_t', \ldots, w_{t-p+1}')'$ is Car(1). Therefore, without loss of generality, we will focus on Car(1) processes in the following.

As will be shown below (Section 6.5), this class includes a large number of processes, e.g. Markov-switching Gaussian vector autoregressive, autoregressive gamma processes or quadratic functions of Gaussian processes. The reason why Car processes are central in term-structure modelling is that there exist quasi-explicit formulas to compute multi-horizon Laplace transforms. These formulas are provided in the following subsection.

6.2.2 *Multi-horizon Laplace transform*

Let us consider a multivariate Car(1) process w_t and its conditional Laplace transform given by $u \mapsto \exp\left[a'(u)w_{t-1} + b(u)\right]$. Let us further denote by $L_{t,h}(U_h)$ its multi-horizon Laplace transform given by

$$L_{t,h}(U_h) = E_t\left[\exp\left(u_h' w_{t+1} + \cdots + u_1' w_{t+h}\right)\right], \; t = 1, \ldots, T,$$

where $U_h = (u_1', \ldots, u_h')$ is a given sequence of vectors.

Proposition 6.1 *We have, for any t,*

$$L_{t,h}(U_h) = \exp\left(A_h' w_t + B_h\right),$$

where the sequences A_h, B_h are obtained recursively by

$$\begin{aligned} A_h &= a(u_h + A_{h-1}), \\ B_h &= b(u_h + A_{h-1}) + B_{h-1}, \end{aligned}$$

with the initial conditions $A_0 = 0$ and $B_0 = 0$.

Proof It is straightforward to show that the formula is true for $h = 1$. If it is true for $h - 1$, we get:

$$\begin{aligned} L_{t,h}(U_h) &= E_t\left[\exp\left(u_h' w_{t+1}\right) E_{t+1}\left(\exp\left(u_{h-1}' w_{t+2} + \cdots + u_1' w_{t+h}\right)\right)\right] \\ &= E_t\left[\exp\left(u_h' w_{t+1}\right) L_{t+1,h-1}(U_{h-1})\right] \\ &= \exp\left[a(u_h + A_{h-1})' w_t + b(u_h + A_{h-1}) + B_{h-1}\right] \end{aligned}$$

and the result follows. □

6.3 Risk-neutral modelling

6.3.1 Assumptions

The global new information in the economy at date t is $w_{G,t} = (w'_{c,t}, w'_{s,t}, d'_t)'$, where $w_{c,t}$ is a vector of common factors (for instance macroeconomic variables), $w_{s,t}$ is a vector $\left(w^{1'}_{s,t}, \dots w^{n'}_{s,t}, \dots w^{N'}_{s,t}\right)$ of specific variables, $w^n_{s,t}$ being associated with debtor n (for instance a firm or a country), and $d_t = \left(d^1_t, \dots d^n_t, \dots d^N_t\right)'$ is a vector of binary variables, d^n_t indicating whether entity n is in default ($d^n_t = 1$) or not ($d^n_t = 0$) at time t.

We assume that $\left(w'_{s,t}, d'_t\right)'$ does not Granger-cause $w_{c,t}$ in the risk-neutral ($\mathbb{Q}$) dynamics, that d_t does not Granger-$\mathbb{Q}$-cause $\left(w'_{c,t}, w'_{s,t}\right)'$ and that the $\left(w^{n'}_{s,t}, d^n_t\right)'$, $n = 1, \dots, N$ are independent conditionally on $(w_{c,t}, \underline{w}_{G,t-1})$.

We introduce the notations $w_t = \left(w'_{c,t}, w'_{s,t}\right)'$ and $w^n_t = \left(w'_{c,t}, w^{n'}_{s,t}\right)'$. Further, we assume that, for any n, w^n_t is $\mathbb{Q}$-Car(1), which implies that w_t is also $\mathbb{Q}$-Car(1), and which includes the case where w_t is Car(p) as mentioned in Section 6.2. Therefore we have

$$E^{\mathbb{Q}}\left(\exp\left(u'w_t\right)\middle|\,\underline{w}_{G,t-1}\right) = E^{\mathbb{Q}}\left(\exp\left(u'w_t\right)\middle|\,\underline{w}_{t-1}\right) = \exp\left(a^{\mathbb{Q}}(u)'w_{t-1} + b^{\mathbb{Q}}(u)\right). \tag{6.1}$$

We also assume that

$$\mathbb{Q}\left(d^n_t = 0\middle|\,d^n_{t-1} = 0, \underline{w}_t\right) = \exp\left(-\lambda^{\mathbb{Q}}_{n,t}\right), \tag{6.2}$$

where

$$\lambda^{\mathbb{Q}}_{n,t} = \alpha_{0,n} + \alpha'_{1,n} w^n_t. \tag{6.3}$$

The risk-neutral default intensity $\lambda^{\mathbb{Q}}_{n,t}$ is close to the conditional default probability $\mathbb{Q}\left(d^n_t = 1\middle|\,d^n_{t-1} = 0, \underline{w}_t\right)$ if it is small. We also assume that the default state is absorbing.

Finally, we assume that the risk-free short-term rate between t and $t+1$, which is known at t, only depends on $w_{c,t}$ and is given by

$$r_t = \beta_0 + \beta'_1 w_{c,t}. \tag{6.4}$$

6.3.2 Computation of the defaultable bond prices

Assuming first that the recovery rate is zero, the price at time t of a zero-coupon bond issued by entity n, with a residual maturity (at t) of h, is

$$B_n(t,h) = E_t^{\mathbb{Q}}\left(\exp\left(-r_t - \cdots - r_{t+h-1}\right)\left(1 - d_{t+h}^n\right)\right). \tag{6.5}$$

We have assumed that w_t is $\mathbb{Q}$-Car(1). However, it is easily seen that $\left(w_t', d_t^n\right)'$ is not $\mathbb{Q}$-Car.[5] Nevertheless, the causality structure described above implies that the computation of $B_n(t,h)$, for any h, boils down to the computation of a multi-horizon Laplace transform of w_t in which the coefficients are ordered backward and is therefore easily obtained (see Proposition 6.1 above).

Proposition 6.2 *$B_n(t,h)$ is obtained, for any h, from a multi-horizon $\mathbb{Q}$-Laplace transform of $w_{t+1},\dots,w_{t+h}$ in which the coefficients are ordered backward.*

Proof We proceed under the assumption that entity n is alive at date t, i.e. $d_t^n = 0$. We have

$$\begin{aligned} B_n(t,h) &= E_t^{\mathbb{Q}}\left(\exp\left(-r_t - \cdots - r_{t+h-1}\right)\left(1 - d_{t+h}^n\right)\right) \\ &= \exp\left(-r_t\right) E_t^{\mathbb{Q}}\left(\exp\left(-r_{t+1} - \cdots - r_{t+h-1}\right)\left(1 - d_{t+h}^n\right)\right). \end{aligned}$$

Conditioning with respect to $\underline{w}_{t+h}$, we get

$$\begin{aligned} B_n(t,h) = \exp\left(-r_t\right) E_t^{\mathbb{Q}}\Big(&\exp\left(-r_{t+1} - \cdots - r_{t+h-1}\right) \\ &\times \prod_{j=1}^{h} \mathbb{Q}\left(d_{t+j}^n = 0 \,\middle|\, d_{t+j-1}^n = 0, \underline{w}_{t+h}\right)\Big). \end{aligned}$$

Since d_t does not Granger-$\mathbb{Q}$-cause w_t and Granger non-causality is equivalent to Sims non-causality, we can replace $\underline{w}_{t+h}$ by $\underline{w}_{t+j}$ in the generic term of the product above and we get

$$\begin{aligned} &B_n(t,h) \\ &\quad= \exp\left(-r_t\right) E_t^{\mathbb{Q}}\left(\exp\left(-r_{t+1} - \cdots - r_{t+h-1} - \lambda_{n,t+1}^{\mathbb{Q}} - \cdots - \lambda_{n,t+h}^{\mathbb{Q}}\right)\right) \\ &\quad= \exp\left(-h(\beta_0 + \alpha_{0,n}) - \tilde{\beta}_1' w_t^n\right) E_t^{\mathbb{Q}}\Big(\exp\Big(-(\tilde{\beta}_1 + \alpha_{1,n})' w_{t+1}^n - \cdots \\ &\qquad -(\tilde{\beta}_1 + \alpha_{1,n})' w_{t+h-1}^n - \alpha_{1,n}' w_{t+h}^n\Big)\Big), \end{aligned} \tag{6.6}$$

[5] See Monfort and Renne (2013).

where $\tilde{\beta}_1 = (\beta_1', 0)'$. So the expectation term in the previous equation is a multi-horizon Laplace transform of $w_{t+1}^n, \ldots, w_{t+h}^n$, with a sequence of coefficients $U_h = \{u_1, \ldots, u_h\}$ (see Proposition 6.1 for the definition of the u_i) defined by

$$u_1 = -\alpha_{1,n}, \quad u_j = -(\tilde{\beta}_1 + \alpha_{1,n}), \quad \forall j \geq 2. \tag{6.7}$$

□

Given Proposition 6.1, the previous proposition implies that the yield $R_n(t,h)$ of a zero-coupon bond issued by debtor n with a residual maturity of h is of the form

$$\begin{aligned} R_n(t,h) &= -\frac{1}{h} \log B_n(t,h) \\ &= c_n(h)' w_t^n + f_n(h), \end{aligned} \tag{6.8}$$

where the $c_n(h), f_n(h)$, $h = 1, \ldots, H$ are computed from a unique recursive scheme.[6] We get an affine term structure of interest rates. It is important to stress that the $c_n(h)$ and the $f_n(h)$ have to be computed only once (they do not depend on t). Since these pricing formulas do not require the use of time-demanding simulations, this framework turns out to be very tractable and, hence, amenable to empirical estimation.

In that framework, it is easily seen that credit spreads, i.e. yields differentials between defaultable bonds and their risk-free counterpart (of the same maturity), are also affine functions of the factors w_t^n. Specifically, let us denote by $R^*(t,h)$ the yield of the risk-free bond with a residual maturity of h. By definition, the risk-free issuer is characterised by zero default intensity, that is, $\alpha_0^* = 0$ and $\alpha_1^* = 0$. Using the recursive algorithm, we get $c^*(h)$ and $f^*(h)$ coefficients that are such that

$$R^*(t,h) = c^*(h)' w_t^n + f^*(h),$$

where only the components of $c^*(h)$ that correspond to $w_{c,t}$ are non-zero. Thus, the credit spreads associated with entity n are simply given by

$$R_n(t,h) - R^*(t,h) = \left[c_n(h) - c^*(h)\right]' w_t^n + \left[f_n(h) - f^*(h)\right]. \tag{6.9}$$

6.3.3 Non-zero recovery rate

Building on the so-called recovery of market value assumption introduced by Duffie and Singleton (1999), Monfort and Renne (2013) show that, when the recovery rate is non-zero, the previous pricing machinery

[6] By "unique" recursive scheme, we mean that only H recursions are needed to compute the $c_n(h)$ and the $f_n(h)$ for any $h \leq H$. This is obtained thanks to the fact that the sequence U_h defined by (6.7) corresponds to the beginning of a longer sequence.

is still valid if λ_t^n is replaced by a recovery-adjusted default intensity $\tilde{\lambda}_t^n$. Specifically, the recovery of market value assumption can be stated in the following manner: if the issuer defaults between the dates $t-1$ and t, the recovery payoff is equal to a fraction $\zeta_{n,t}$ – that can be a function of w_t^n – of the price that would have prevailed, absent the default of the issuer. In this context, the recovery adjusted default intensity is given by[7]

$$\exp(-\tilde{\lambda}_t^n) = \exp(-\lambda_{n,t}^{\mathbb{Q}}) + \left(1 - \exp(-\lambda_{n,t}^{\mathbb{Q}})\right)\zeta_{n,t}. \tag{6.10}$$

Obviously, if $\zeta_{n,t} \equiv 0$, the recovery-adjusted intensity is equal to the default intensity and if $\zeta_{n,t} \equiv 1$, the bond turns out to be equivalent to a risk-free bond.

6.3.4 *Internal-consistency conditions*

The general framework of Section 6.3.1 contains the particular situation in which some components of w_t^n are yields $R_n(t, h_i)$, $i = 1, \dots, I$. In this case, the coefficients $c_n(h)$ and $f_n(h)$ appearing in formula (6.8) must satisfy the conditions

$$c_n(h_i) = e_i,\ f_n(h_i) = 0,$$

where e_i is the selection vector picking the entry of w_t^n that correspond to $R_n(t, h_i)$.

6.4 Back to the historical world

6.4.1 *$\mathbb{Q}$-dynamics, $\mathbb{P}$-dynamics, sdf and the short rate*

The historical and risk-neutral conditional distributions of $w_{G,t}$ given $\underline{w}_{G,t-1}$ have probability density functions (pdf) with respect to a same measure, these pdf's are denoted respectively by $f^{\mathbb{P}}(w_{G,t} | \underline{w}_{G,t-1})$ and $f^{\mathbb{Q}}(w_{G,t} | \underline{w}_{G,t-1})$. Denoting by $M_{t-1,t}(\underline{w}_{G,t})$ the stochastic discount factor between $t-1$ and t, and by $r_{t-1}(\underline{w}_{G,t-1})$ the riskfree short rate between $t-1$ and t, we know that the four mathematical objects $f^{\mathbb{P}}$, $f^{\mathbb{Q}}$, $M_{t-1,t}$, r_{t-1} are linked by the relation[8]

[7] See Monfort and Renne (2013) for the proof.

[8] The existence, the unicity and the positivity of the sdf are consequences of the assumptions of existence, linearity and continuity of the pricing function and of the assumption of absence of arbitrage opportunity (see Bertholon, Monfort and Pegoraro, 2007). In the particular case where $w_{G,t}$ is a vector of prices of basic assets, this implies that the other prices are functions of the basic prices. This can be viewed as a completeness property. However, as usual in discrete-time models where the reallocation of portfolios is only allowed at discrete dates, exact replicating portfolios do not exist in general.

$$f^{\mathbb{Q}}(w_{G,t}|\,\underline{w}_{G,t-1}) = f^{\mathbb{P}}(w_{G,t}|\,\underline{w}_{G,t-1})M_{t-1,t}(\underline{w}_{G,t})\exp(r_{t-1}(\underline{w}_{G,t-1})), \tag{6.11}$$

Since $f^{\mathbb{Q}}(w_{G,t}|\,\underline{w}_{G,t-1})$ integrates to one, we have

$$E^{\mathbb{P}}_{t-1}(M_{t-1,t}) = \exp(-r_{t-1}(\underline{w}_{G,t-1})), \tag{6.12}$$

Equation (6.11) can also be written

$$f^{\mathbb{P}}(w_{G,t}|\,\underline{w}_{G,t-1}) = f^{\mathbb{Q}}(w_{G,t}|\,\underline{w}_{G,t-1})M^{-1}_{t-1,t}(\underline{w}_{G,t})\exp(-r_{t-1}(\underline{w}_{G,t-1})), \tag{6.13}$$

which implies

$$E^{\mathbb{Q}}_{t-1}M^{-1}_{t-1,t}(\underline{w}_{G,t}) = \exp(r_{t-1}(\underline{w}_{G,t-1})) \tag{6.14}$$

or

$$M_{t-1,t}(\underline{w}_{G,t}) = \frac{f^{\mathbb{Q}}(w_{G,t}|\,\underline{w}_{G,t-1})}{f^{\mathbb{P}}(w_{G,t}|\,\underline{w}_{G,t-1})}\exp(-r_{t-1}(\underline{w}_{G,t-1})). \tag{6.15}$$

In particular, (6.15) shows that once $f^{\mathbb{Q}}$ and r_{t-1} are specified, as we did above, $f^{\mathbb{P}}$ can be chosen arbitrarily and (6.15) gives the stochastic discount factor.

In this chapter, we assume that the sdf $M_{t-1,t}$ only depends on the common variables $\underline{w}_{c,t}$ or, equivalently, that the specific variables $\underline{w}_{s,t}$ and the individual default variables d_t have no direct impact on $M_{t-1,t}$. In this case $f^{\mathbb{P}}$ given by (6.13) is no longer arbitrary. More precisely, since we have assumed that $f^{\mathbb{Q}}(w_{G,t}|\,\underline{w}_{G,t-1})$ can be factorized as

$$f^{\mathbb{Q}}(w_{c,t}, w_{s,t}, d_t|\,\underline{w}_{G,t-1}) = f^{\mathbb{Q}}_{c}(w_{c,t}|\,\underline{w}_{c,t-1})f^{\mathbb{Q}}_{sd}(w_{s,t}, d_t|\,w_{c,t}, \underline{w}_{G,t-1}). \tag{6.16}$$

We see, by integrating both sides of (6.13) with respect to $(w_{s,t}, d_t)$ that

$$f^{\mathbb{P}}_{c}(w_{c,t}|\,\underline{w}_{G,t-1}) = f^{\mathbb{Q}}_{c}(w_{c,t}|\,\underline{w}_{c,t-1})M^{-1}_{t-1,t}(\underline{w}_{c,t})\exp(-r_{t-1}(\underline{w}_{c,t-1})). \tag{6.17}$$

Therefore, $(w_{s,t}, d_t)$ does not cause $w_{c,t}$ in the historical world and, moreover,

$$\begin{aligned} f^{\mathbb{P}}_{sd}(w_{s,t}, d_t|\,\underline{w}_{c,t}, \underline{w}_{G,t-1}) &= \frac{f^{\mathbb{P}}(w_{c,t}, w_{s,t}, d_t|\,\underline{w}_{G,t-1})}{f^{\mathbb{P}}_{c}(w_{c,t}|\,\underline{w}_{G,t-1})} \\ &= f^{\mathbb{Q}}_{sd}(w_{s,t}, d_t|\,\underline{w}_{c,t}, \underline{w}_{G,t-1}). \end{aligned}$$

Hence the following proposition.

Proposition 6.3 *Under the causality assumptions defined in Section 6.3.1, and if $M_{t-1,t}$ only depends on the common variables $\underline{w}_{c,t}$:*

- *$(w_{s,t}, d_t)$ does not cause $w_{c,t}$ in the historical world;*
- *the conditional distribution of $(w_{s,t}, d_t)$ given $(w_{c,t}, \underline{w}_{G,t-1})$ are the same in both worlds.*

The assumptions made in Section 6.3.1 and the second part of the previous proposition imply that the historical conditional distribution of $w_{s,t}$ given $(w_{c,t}, \underline{w}_{G,t-1})$ is the same as in the risk-neutral world and that the functional forms of the risk-neutral default intensities given in (6.2) and (6.3) are also valid in the historical world:

$$\lambda^{\mathbb{P}}_{n,t} = \lambda^{\mathbb{Q}}_{n,t} = \alpha_{0,n} + \alpha'_{1,n} w^n_t. \tag{6.18}$$

However, it is important to stress that the risk-neutral and the historical dynamics of $\lambda^{\mathbb{P}}_{n,t}$ (and $\lambda^{\mathbb{Q}}_{n,t}$) are different because those of w_t are different. The fact that $\lambda^{\mathbb{P}}_{n,t} = \lambda^{\mathbb{Q}}_{n,t}$ is very important. Indeed, it means that we can compute historical – or real-world – default probabilities as soon as we know the historical dynamics of w_t and the parameterisations of the vectors $\alpha_{0,n}$ and $\alpha_{1,n}$. Given that (a) we observe the historical dynamics of the yields $R_n(t, h)$ and that (b) these yields are related to the factors w_t through the coefficients $c_n(h)$ and $f_n(h)$ that depend themselves on the $\alpha_{i,n}$s, it is possible to estimate the vectors $\alpha_{0,n}$ and $\alpha_{1,n}$ and the historical dynamics of the factors w_t (inference will be discussed in Section 6.6). This is exploited by Borgy et al. (2011) and Monfort and Renne (2012) who derive historical term structures of probabilities of default after having estimated some affine-term structure models.

6.4.2 *Specification of the sdf*

From formula (6.17) we see that once the risk-neutral dynamics of the common factors $w_{c,t}$ is specified as well as the function $r_{t-1}(w_{c,t-1})$, we can choose the historical dynamics of $w_{c,t}$ completely freely, and the sdf $M_{t-1,t}(\underline{w}_{c,t})$ is obtained as a by-product. However, in order to have a specification of $M_{t-1,t}(\underline{w}_{c,t})$ which is easily interpretable, it is usual to choose a particular form, for instance the exponential affine form:[9]

$$M_{t-1,t}(\underline{w}_{c,t}) = \exp\left(-r_{t-1}(\underline{w}_{c,t-1}) + \gamma(\underline{w}_{c,t-1})' w_{c,t} + \Psi^{\mathbb{Q}}_{t-1}\left(-\gamma(\underline{w}_{c,t-1})\right)\right), \tag{6.19}$$

where $\Psi^{\mathbb{Q}}_{t-1}(u)$ is the risk-neutral conditional log-Laplace transform of $w_{c,t}$ defined by

$$\Psi^{\mathbb{Q}}_{t-1}(u) = \ln E^{\mathbb{Q}}_{t-1}\left[\exp\left(u' w_{c,t}\right)\right].$$

[9] See Monfort and Pegoraro (2012) for a generalisation to the exponential quadratic case.

Note that condition (6.14) is automatically satisfied by this specification.

The coefficients $\gamma(\underline{w}_{c,t-1})$, also denoted by γ_{t-1}, are interpreted as the risk sensitivities associated with the risk factors, i.e. the components of $w_{c,t}$.

In this case, the historical conditional log-Laplace transform of $w_{c,t}$ is

$$\begin{aligned}\Psi^{\mathbb{P}}_{t-1}(u) &= \log E^{\mathbb{P}}_{t-1}\left[\exp\left(u'w_{c,t}\right)\right] \\ &= \log E^{\mathbb{Q}}_{t-1}\left[M^{-1}_{t-1,t}(\underline{w}_{G,t})\exp(-r_t(\underline{w}_{G,t-1}))\exp\left(u'w_{c,t}\right)\right] \\ &= \log E^{\mathbb{Q}}_{t-1}\left[\exp\left((u-\gamma_{t-1})'w_{c,t} - \Psi^{\mathbb{Q}}_{t-1}(-\gamma_{t-1})\right)\right] \\ &= \Psi^{\mathbb{Q}}_{t-1}(u-\gamma_{t-1}) - \Psi^{\mathbb{Q}}_{t-1}(-\gamma_{t-1}). \end{aligned} \tag{6.20}$$

Hence the following proposition.

Proposition 6.4 *If the sdf has the exponential affine form (6.19), the historical dynamics of $w_{c,t}$ is easily deduced from the risk-neutral one, by the formula*

$$\Psi^{\mathbb{P}}_{t-1}(u) = \Psi^{\mathbb{Q}}_{t-1}(u-\gamma_{t-1}) - \Psi^{\mathbb{Q}}_{t-1}(-\gamma_{t-1}).$$

Also note that setting $u = \gamma_{t-1}$ in (6.20) results in $\Psi^{\mathbb{P}}_{t-1}(\gamma_{t-1}) = -\Psi^{\mathbb{Q}}_{t-1}(-\gamma_{t-1})$ and replacing u by $u+\gamma_{t-1}$, we get the reverse relation:

$$\Psi^{\mathbb{Q}}_{t-1}(u) = \Psi^{\mathbb{P}}_{t-1}(u+\gamma_{t-1}) - \Psi^{\mathbb{P}}_{t-1}(\gamma_{t-1}). \tag{6.21}$$

In Section 6.2.2, we have assumed that $w_t = (w'_{c,t}, w'_{s,t})'$ is Car(1). Therefore, its log-Laplace transform is of the form (6.1):

$$\begin{aligned}&E^{\mathbb{Q}}_{t-1}\left[\exp\left(u'w_{c,t} + v'w_{s,t}\right)\right] \\ &\quad = \exp\left[a^{\mathbb{Q}}_1(u,v)'w_{c,t-1} + a^{\mathbb{Q}}_2(u,v)'w_{s,t-1} + b^{\mathbb{Q}}(u,v)\right].\end{aligned} \tag{6.22}$$

Since $w_{s,t}$ does not cause $w_{c,t}$, it turns out that $w_{c,t}$ is also Car(1). Indeed, the conditional Laplace transform of $w_{c,t}$ given $\underline{w}_{t-1} = (\underline{w}'_{c,t}, \underline{w}'_{s,t})'$ is obtained by putting $v = 0$ in (6.22), and since the result does not depend on $w_{s,t-1}$, we have $a^{\mathbb{Q}}_2(u,0) = 0$ and finally

$$\begin{aligned}E^{\mathbb{Q}}_{t-1}\left[\exp\left(u'w_{c,t}\right)\right] &= \exp\left[a^{\mathbb{Q}}_1(u,0)'w_{c,t-1} + b^{\mathbb{Q}}(u,0)\right] \\ &= \exp\left[a^{\mathbb{Q}}_c(u)'w_{c,t-1} + b^{\mathbb{Q}}_c(u)\right], \quad \text{(say)}.\end{aligned}$$

Therefore, $\Psi^{\mathbb{Q}}_{t-1}(u) = a^{\mathbb{Q}}_c(u)'w_{c,t} + b^{\mathbb{Q}}_c(u)$ and

$$\Psi^{\mathbb{P}}_{t-1}(u) = \left[a^{\mathbb{Q}}_c(u-\gamma_{t-1}) - a^{\mathbb{Q}}_c(-\gamma_{t-1})\right]'w_{c,t-1} + b^{\mathbb{Q}}_c(u-\gamma_{t-1}) - b^{\mathbb{Q}}_c(-\gamma_{t-1}).$$

This shows that, in general, the historical dynamics of $w_{c,t}$ is not Car, except when the previous function is affine in $w_{c,t-1}$. A sufficient but not necessary condition is that γ_{t-1} is constant.

6.5 Examples

6.5.1 Autoregressive gamma latent factors

Autoregressive gamma processes are investigated by Gourieroux and Jasiak (2006). A process y_t that follows an autoregressive gamma process of order one, or ARG(1), can be defined in the following way:

$$\begin{aligned}
\left.\frac{y_t}{\mu_y}\right|\tilde{y}_t &\sim \gamma(\nu_y + \tilde{y}_t), \quad \nu_y > 0,\\
\tilde{y}_t | y_{t-1} &\sim \mathcal{P}(\rho_y y_{t-1}/\mu_y), \quad \rho_y > 0, \mu_y > 0,
\end{aligned}$$

where γ and $\mathcal{P}$ denote respectively the gamma and the Poisson distributions, μ_y is the scale parameter, ν_y is the degree of freedom, ρ_y is the correlation parameter and $\tilde{y}_t$ is the mixing variable. As shown by Gourieroux and Jasiak (2006), the ARG(1) process is the discrete-time equivalent of the square-root (CIR) diffusion process. It can be shown that

$$y_t = \nu_y \mu_y + \rho_y y_{t-1} + \eta_{y,t},$$

where $\eta_{y,t}$ is a martingale difference sequence whose conditional variance (at date $t-1$) is given by $\nu_y \mu_y^2 + 2\rho_y \mu_y y_{t-1}$. Figure 6.1 (upper right plot) shows the simulated path of an ARG(1) process.

Let us assume that $w_{c,t}$ is a univariate $\mathbb{Q}$ autoregressive gamma (ARG(1)), as well as $w_{s,t}^n$ $(n = 1, \dots, N)$, and that they are $\mathbb{Q}$ (and therefore $\mathbb{P}$) independent. The extension to the multivariate case is straightforward. The conditional log-Laplace transforms of $w_{c,t}$ and $w_{s,t}^n$ in the risk-neutral world are respectively:

$$\begin{aligned}
&\frac{\rho_c u}{1 - u\mu_c} w_{c,t-1} - \nu_c \log(1 - u\mu_c),\\
&\frac{\rho_s^n u}{1 - u\mu_s^n} w_{s,t-1}^n - \nu_s^n \log(1 - u\mu_s^n).
\end{aligned}$$

If $w_{c,t}$ and $w_{s,t}$ are latent, we can assume that the scale parameters μ_c and μ_s^n are equal to one. It is well-known (see Gourieroux and Jasiak, 2006) that these processes are positive and that ρ_c (respectively ρ_s^n) is a (positive) correlation parameter whereas ν_c (respectively ν_s^n) is a shape parameter. Further, we assume that:

$$\begin{aligned}
\lambda_{n,t} &= \alpha_{0,n} + \alpha_{1,n}^{c} w_{c,t} + \alpha_{1,n}^{s} w_{s,t}^n,\\
r_t &= \beta_0 + \beta_1' w_{c,t-1}.
\end{aligned}$$

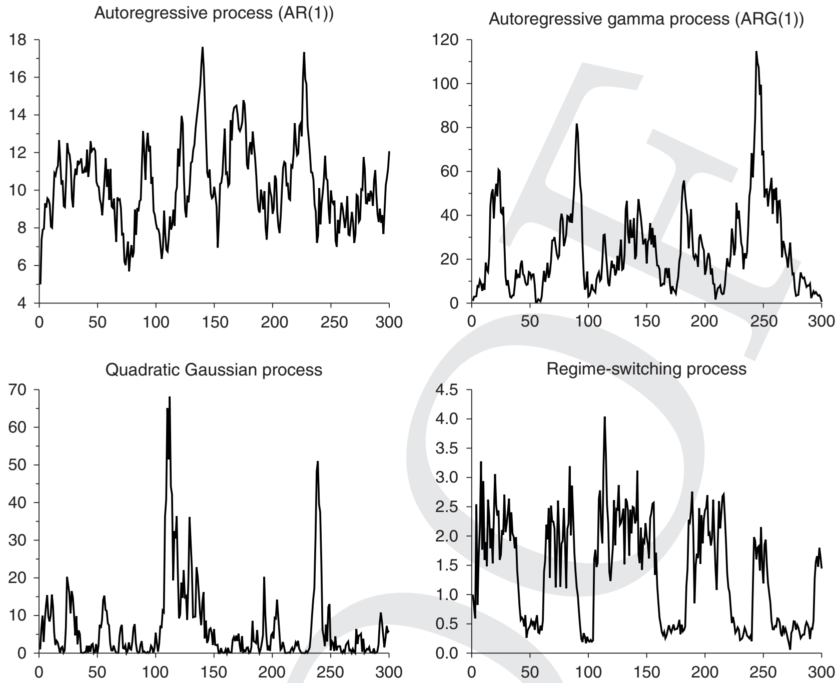

Figure 6.1 Examples of Car processes
The upper left panel shows a simulation of a Gaussian process parameterised by $\varphi = 1$, $\phi = 0.9$ and $\Sigma = 1$ (see Section 6.5.2). The upper right panel displays the simulated path of an autoregressive gamma process with $\mu = 1$, $\rho = 0.95$, and $\nu = 1$ (see Section 6.5.1). The lower left panel shows the square of a Gaussian AR process (therefore a quadratic Gaussian process), with $\varphi = 0$, $\phi = 0.9$ and $\Sigma = 1$ (see Section 6.5.3). The lower right panel displays the simulated path of a switching Gaussian AR process: there are two regimes, the probability of staying in each of these two regimes being 95% (i.e. the $\pi_{i,i}$s are equal to 0.95), $\varphi(z) = [0.1 \quad 1]z$, $\Phi(z) = [0.7 \quad 0.5]z$ and $\Sigma(z) = [0.01 \quad 0.25]z$

Since $(w_{c,t}, w^n_{s,t})$ is obviously $\mathbb{Q}$-Car(1), we can easily obtain the yields $R_n(t, h)$ as affine functions of $w_{c,t}$ and $w^n_{s,t}$ (using the results of Section 6.3.2 and applying the recursive algorithm proposed in Section 6.2.2). We know, given the independence of $w_{c,t}$ and $w^n_{s,t}$ and the assumption on $M_{t-1,t}$, that the historical dynamics of $w^n_{s,t}$ is the same as the risk-neutral one (Proposition 6.3), and that, if we adopt an exponential affine sdf, the historical conditional log-Laplace transform of $w_{c,t}$ is (using Equation (6.20))

$$\rho_c\left(\frac{u-\gamma_{t-1}}{1-u+\gamma_{t-1}}+\frac{\gamma_{t-1}}{1+\gamma_{t-1}}\right)w_{c,t-1}-\nu_c\{\log(1-(u-\gamma_{t-1}))$$
$$-\log(1+\gamma_{t-1})\}=\frac{\rho_{t-1}u}{1-u\mu_{t-1}}w_{c,t-1}-\nu_c\log(1-u\mu_{t-1})$$

with $\rho_{t-1}=\frac{\rho_c}{(1+\gamma_{t-1})^2}$ and $\mu_{t-1}=\frac{1}{1+\gamma_{t-1}}$.

Therefore, if γ_{t-1} is constant, the historical dynamics of $w_{c,t}$ is also ARG(1) with modified parameters $\rho_c^*=\frac{\rho_c}{(1+\gamma)^2}$, $\mu_c^*=\frac{1}{1+\gamma}$, $\nu_c^*=\nu_c$. It is important to note that since the processes $w_{c,t}$ and $w_{s,t}$ are positive, the same is true for r_t and λ_t if we take positive coefficients. Moreover, since

$$B_n(t,h)=E_t^{\mathbb{Q}}[-r_{t+1}-\cdots-r_{t+h}-\lambda_{t+1}-\cdots-\lambda_{t+h}],$$

the function under the expectation is always smaller than one and, therefore, all the yields $R_n(t,h)=-1/h\log B_n(t,h)$ are positive. The ability of the ARG processes to model positive yields and spreads is a substantial advantage of these processes.

6.5.2 Gaussian VAR factors and affine term structures

We assume that the risk-neutral dynamics of $(w'_{c,t},w'_{s,t})'$ is defined by

$$\begin{aligned}w_{c,t}&=\varphi_c+\Phi_c w_{c,t-1}+\varepsilon_{c,t},\\ w_{s,t}^n&=\varphi_s^n+\Phi_{sc}^n w_{c,t-1}+\Phi_s^n w_{s,t-1}^n+\varepsilon_{s,t}^n,\end{aligned}$$

where $\varepsilon_{c,t}$ and the $\varepsilon_{s,t}^n$s are independent zero-mean Gaussian white noises, with respective variance–covariance matrices Σ_c and Σ_s^n.

This implies that the vectors $(w'_{c,t},w^{n'}_{s,t})'$ and $(w'_{c,t},w'_{s,t})'$ are Gaussian VAR(1). Again, assuming that r_t is an affine function of $w_{c,t}$ and that $\lambda_{n,t}$ is an affine function of $w_{c,t}$ and $w_{s,t}^n$, the yields $R_n(t,h)$ are affine functions of $w_{c,t}$ and $w_{s,t}^n$.

Let us use an sdf of the form

$$M_{t-1,t}=\exp\left(-r_{t-1}+\nu'_{t-1}\varepsilon_{c,t}+\frac{1}{2}\nu'_{t-1}\Sigma_c\nu_{t-1}\right)$$

with $\nu_{t-1}=\nu_0(z_{t-1})+\nu_1(z_{t-1})w_{c,t-1}$. This sdf is obviously exponential affine in $w_{c,t}$ and satisfies conditions (6.14), i.e. $E_{t-1}^{\mathbb{Q}}(M_{t-1,t}^{-1})=\exp(r_{t-1})$.

The conditional risk-neutral log-Laplace transform of $\varepsilon_{c,t}$ is

$$\psi_{t-1}^{\mathbb{Q}}=\frac{1}{2}u'\Sigma_c u.$$

Using (6.20), it is easily seen that the conditional historical log-Laplace transform of $\varepsilon_{c,t}$ is

$$\begin{aligned}\Psi^{\mathbb{P}}_{t-1}(u) &= \frac{1}{2}(u-\nu_{t-1})'\Sigma_c(u-\nu_{t-1}) - \frac{1}{2}\nu'_{t-1}\Sigma_c\nu_{t-1}\\ &= -\nu'_{t-1}\Sigma_c u + \frac{1}{2}u'\Sigma_c u.\end{aligned}$$

Therefore, the conditional historical distribution of $\varepsilon_{c,t}$ is $\mathcal{N}(-\Sigma_c\nu_{t-1}, \Sigma_c)$ or, in other words, $\varepsilon_{c,t}$ is equal to $-\Sigma_c\nu_{t-1} + \eta_{c,t}$, where $\eta_{c,t}$ is a white noise $\mathcal{N}(0,\Sigma_c)$ in the historical world and we get:

$$\begin{aligned}w_{c,t} &= \varphi_c + \Phi_c w_{c,t-1} - \Sigma_c\nu_{t-1} + \eta_{c,t}\\ &= (\varphi_c - \Sigma_c\nu_0) + (\Phi_c - \Sigma_c\nu_1)w_{c,t-1} + \eta_{c,t}.\end{aligned}$$

So, with the specification of $M_{t-1,t}$ chosen above, $w_{c,t}$ is also a Gaussian VAR(1) in the historical world, with arbitrarily modified vector of constants and autoregressive matrix but with the same conditional variance–covariance matrix. This model has been extensively used in the affine term-structure literature.[10]

6.5.3 Gaussian VAR factors and quadratic term-structures

Let us consider the same risk-neutral dynamics of $w_t = (w'_{c,t}, w'_{s,t})'$ and the same sdf as in the previous subsection. The historical dynamics of w_t is also the same as before. Let us now assume that the short rate r_t and the default intensities are some quadratic functions of the factors

$$\begin{aligned}r_t &= \beta_0 + \beta'_1 w_{c,t} + w'_{c,t}\beta_2 w_{c,t},\\ \lambda_{n,t} &= \alpha_{0,n} + \alpha_{1,n}w^n_t + w^{n'}_t\alpha_{2,n}w^n_t\end{aligned}$$

with $w^n_t = (w'_{c,t}, w^{n'}_{s,t})'$. It will prove convenient to rewrite r_t in the following way:

$$r_t = \beta_0 + \tilde{\beta}'_1 w^n_t + w^{n'}_t\tilde{\beta}_2 w^n_t,$$

where $\tilde{\beta}_1 = (\beta'_1, 0)'$ and $\tilde{\beta}_2 = \begin{pmatrix}\beta_2 & 0\\ 0 & 0\end{pmatrix}$. Indeed, r_t and the $\lambda_{n,t}$s are then both quadratic forms in w^n_t and they can also be written

$$\begin{aligned}r_t &= \beta_0 + \beta'_1 w_{c,t} + \mathrm{Trace}(\beta_2 w_{c,t}w'_{c,t}),\\ \lambda_{n,t} &= \alpha_{0,n} + \alpha_{1,n}w^n_t + \mathrm{Trace}(\alpha_{2,n}w^n_t w^{n'}_t).\end{aligned}$$

Since w^n_t follows a Gaussian VAR(1), it can be shown that $(w^{n'}_t, \mathrm{vech}(w^n_t w^{n'}_t))'$ is a Car(1) process (see Gourieroux and Sufana, 2011),

[10] Notably by Ang and Piazzesi (2003) and Joslin, Singleton and Zhu (2011).

therefore any yield $R_n(t,h)$ will be easily computed recursively as an affine function of $(w_t^{n'}, \text{vech}(w_t^n w_t^{n'}))'$, that is to say as a quadratic form in w_t^n.[11]

6.5.4 Switching Gaussian VAR factors

We now introduce a Markov chain z_t valued in $(e_1, \ldots, e_J)$, where e_i is the J-vector whose entries are equal to zero, except the ith one, which is equal to one. We assume that, under $\mathbb{Q}$, z_t has a transition matrix $\Pi = \{\pi_{i,j}\}_{i,j\in\{1,\ldots,J\}}$, where $\pi_{i,j} = \mathbb{Q}(z_t = e_j \,|\, z_{t-1} = e_i)$.[12] We also assume that

$$\begin{aligned} w_{c,t} &= \varphi_c(z_{t-1}) + \Phi_c w_{c,t-1} + \varepsilon_{c,t}, \\ w_{s,t}^n &= \varphi_c^n(z_{t-1}) + \Phi_{s,c}^n w_{c,t} + \Phi_s^n w_{s,t-1}^n + \varepsilon_{s,t}^n, \end{aligned} \tag{6.23}$$

where, conditionally to the past, $\varepsilon_{c,t}$ and $\varepsilon_{s,t}^n$ are independent zero-mean Gaussian with respective variance–covariance matrices $\Sigma_c(z_{t-1})$ and $\Sigma_s^n(z_{t-1})$ in the risk-neutral world. (For the sake of illustration, Figure 6.1 (bottom-right plot) shows the simulated path of an AR(1) switching process.)

We can see this new model as an augmented common factor model with $\tilde{w}_{c,t} = (w_{c,t}', z_t')'$. It is easily checked that $\tilde{w}_{c,t}$ and $\tilde{w}_t^n = (\tilde{w}_{c,t}', w_{s,t}^{n'})'$ are $\mathbb{Q}$-Car(1) and that $R_n(t,h)$ is affine in $(w_{c,t}', z_t', w_{s,t}^{n'})'$. Let us consider an exponential affine sdf:

$$M_{t-1,t} = \exp\left(-r_{t-1} + \nu_{t-1}'\varepsilon_{c,t} + \frac{1}{2}\nu_{t-1}'\Sigma_c(z_{t-1})\nu_{t-1} + \delta_{t-1}' z_t\right),$$

where $\nu_{t-1} = \nu_0(z_{t-1}) + \nu_1(z_{t-1})w_{c,t-1}$ and δ_{t-1} is the function of $w_{c,t-1}$ and z_{t-1} whose jth component is equal to $\log(\pi_{i,j}/\tilde{\pi}_{i,j,t})$ when $z_{t-1} = e_i$, where the $\tilde{\pi}_{i,j,t}$s are functions of $w_{c,t-1}$ satisfying $\Sigma_j \tilde{\pi}_{i,j,t} = 1$ for any i and any $w_{c,t-1}$. Note that $M_{t-1,t}$ automatically satisfies condition (6.14), i.e. $E_{t-1}^{\mathbb{Q}}(M_{t-1,t}^{-1}) = \exp(r_{t-1})$.

If $z_{t-1} = e_i$, the risk-neutral conditional Laplace transform of $(\varepsilon_{c,t}, z_t)$ is

$$E_{t-1}^{\mathbb{Q}}(\exp(u'\varepsilon_{c,t} + v'z_t)) = \exp\left(\frac{1}{2}u'\Sigma_c(e_i)u\right) . \sum_{j=1}^{J} \pi_{i,j}\exp(v_j).$$

[11] For any matrix M of dimension $q \times q$, the half-vectorisation of M, denoted by vech(M), is the $q(q+1)/2 \times 1$ column vector obtained by vectorising only the lower triangular part of M.

[12] Therefore, the rows of Π sum to one.

Therefore, for any z_{t-1}, the conditional log-Laplace transform of ($\varepsilon_{c,t}$, z_t) is

$$\Psi^{\mathbb{Q}}_{t-1}(u,v) = \frac{1}{2}u'\Sigma_c(z_{t-1})u + \left[A^{\mathbb{Q}}_1(v), \ldots, A^{\mathbb{Q}}_{\mathcal{J}}(v)\right] z_{t-1}$$

with $A^{\mathbb{Q}}_i(v) = \ln\left(\sum_{j=1}^{\mathcal{J}} \pi_{i,j}\exp(v_j)\right)$.

Using the results of Proposition 6.4, the historical conditional log-Laplace transform of $(\varepsilon_{c,t}, z_t)$ is

$$\begin{aligned}
\Psi^{\mathbb{P}}_{t-1}(u,v) &= \Psi^{\mathbb{Q}}_{t-1}(u-\nu_{t-1}, v-\delta_{t-1}) - \Psi^{\mathbb{Q}}_{t-1}(-\nu_{t-1}, -\delta_{t-1}) \\
&= \frac{1}{2}(u-\nu_{t-1})'\Sigma_c(z_{t-1})(u-\nu_{t-1}) - \frac{1}{2}\nu'_{t-1}\Sigma_c(z_{t-1})\nu_{t-1} \\
&\quad + \frac{1}{2}\Big[\ A^{\mathbb{Q}}_1(v-\delta_{t-1}) - A^{\mathbb{Q}}_1(-\delta_{t-1}),\ \ldots,\ A^{\mathbb{Q}}_{\mathcal{J}}(v-\delta_{t-1}) \\
&\quad - A^{\mathbb{Q}}_{\mathcal{J}}(-\delta_{t-1})\Big] z_{t-1}.
\end{aligned}$$

It is straightforward to show that

$$\begin{aligned}
A^{\mathbb{Q}}_i(v-\delta_{t-1}) - A^{\mathbb{Q}}_i(-\delta_{t-1}) &= \ln\left(\sum_{j=1}^{\mathcal{J}} \pi_{i,j,t}\exp(v_j)\right) \\
&= A^{\mathbb{P}}_i(v) \quad \text{(say)},
\end{aligned}$$

which leads to

$$\begin{aligned}
\Psi^{\mathbb{P}}_{t-1}(u,v) = &-\nu'_{t-1}\Sigma_c(z_{t-1})u + \frac{1}{2}u'\Sigma_c(z_{t-1})u \\
&+ \left[A^{\mathbb{P}}_1(v), \ldots, A^{\mathbb{P}}_{\mathcal{J}}(v)\right] z_{t-1}.
\end{aligned}$$

Therefore, in the historical world, $\varepsilon_{c,t}$ and z_t are conditionally independent, the marginal distribution of $\varepsilon_{c,t}$ is $\mathcal{N}(-\Sigma_c(z_{t-1})\nu_{t-1}, \Sigma_c(z_{t-1}))$ and the distribution of z_t is defined by the mass points $\pi_{i,j,t}$, $j = 1, \ldots, \mathcal{J}$ if $z_{t-1} = e_i$.

Using the form of $\nu_{t-1} = \nu_0(z_{t-1}) + \nu_1(z_{t-1})w_{c,t-1}$, we conclude that the historical distribution of $(w_{c,t}, z_t)$ is given by

$$w_{c,t} = [\varphi_c(z_{t-1}) - \Sigma_c(z_{t-1})\nu_0(z_{t-1})] + [\Phi_c - \Sigma_c(z_{t-1})\nu_1(z_{t-1})]\, w_{c,t-1} + \eta_{c,t},$$

where the historical conditional distribution of $\eta_{c,t}$ is $\mathcal{N}(0, \Sigma_c(z_{t-1}))$ and z_t is the non-homogenous Markov chain with transition probabilities $\pi_{i,j,t}$ which may be functions of $w_{c,t-1}$.

The previous equation also shows that, in the historical world, the autoregressive matrix may also depend on the regime z_{t-1}. Finally, as mentioned in Proposition 6.3, the historical conditional distribution of

$w^n_{s,t}$ given $(\tilde{w}_{c,t}, \underline{\tilde{w}}^n_{t-1})$ is the same as the risk-neutral one and is given by the second equation of (6.23).

6.6 Inference

The various mathematical objects introduced above are specified parametrically and the parameters must be estimated. The estimation method crucially depends on the observability of the components of $w_t = (w'_{c,t}, w'_{s,t})'$. Regarding the yields, we usually observe several of them: $R_n(t, h_i), i = 1, \dots, I_n, n = 1, \dots, N, t = 1, \dots, T$, which are affine functions of the factors w_t. Various kinds of interest-rate term structures can be thought of: sovereign, supra-national, agency, corporate bonds' yields or swap yields.[13]

If the w_ts are observable we can, adding measurement error terms in Equations (6.8), compute the likelihood of the observations $\{w_t, R_n(t, h_i), i = 1, \dots, I_n, n = 1, \dots, N\}$ for $t = 1, \dots, T$ and derive the maximum likelihood estimator (MLE) of the unknown parameters.[14] Note that analytical formulas are readily available to compute the likelihood associated with the processes presented in Section 6.5.

If some components of w_t are latent, some filters have to be applied to compute the likelihood. For instance, if some of the components of w_t are unobserved Gaussian autoregressive processes, the Kalman filter is the appropriate tool (see de Jong, 2000, among many others). When the specifications of the intensities include quadratic functions of Gaussian autoregressive factors (as in Section 6.5.3), the yields are quadratic functions of the factors. In that case, the standard Kalman filter has to be replaced by its extended or unscented versions (see respectively Kim and Singleton, 2012 or Chen et al., 2008), or by the particle filter (see Andreasen and Meldrum, 2011). The next section also shows how inversion techniques *à la* Chen and Scott (1993) can be resorted to in that case. If the unobserved components of w_t follow Markov-switching Gaussian VAR (as in Section 6.5.4), the Kitagawa–Hamilton filter can be applied (see Monfort and Pegoraro, 2007). When the latent factors follow Markov-switching VAR processes, one can use Kim's (1994) filter (see Monfort and Renne, 2012). Besides, Monfort and Renne (2013) show how inversion techniques can be

[13] Using approaches consistent with the present framework, Monfort and Renne (2012) and Borgy et al. (2011) model the term structures of euro-area sovereign bond yields. Monfort and Renne (2013) and Mueller (2008) model the term structures of yields associated with different credit-rating classes.

[14] The error terms are usually supposed to be mutually and serially independent and identically normally distributed.

mixed with other techniques, notably the Kitagawa–Hamilton filter, to simultaneously handle different forms of latency in the dynamics of the pricing factors. This methodology is applied by Renne (2012).

6.7 Application: credit-risk premia in Italian and Spanish sovereign yields

6.7.1 Outline

In this section, we use our framework to investigate the dynamics of sovereign spreads. Specifically, we model the term-structure of spreads between Spanish government bond yields and their German counterpart. As in many studies on European sovereign bonds, we consider German bonds (the so-called *Bunds*) as risk-free benchmarks. We make use of the quadratic framework (see Section 6.5.3), the pricing factors w_t are latent and the model is estimated using maximum likelihood techniques. Once we have estimated both the historical and the risk-neutral dynamics, we compute credit-risk premia. Our results indicate that an important share of the spreads are accounted for by credit-risk premia, which is consistent with the existence of an undiversifiable sovereign risk (see e.g. Longstaff et al., 2011). As in Monfort and Renne (2012), we finally assess the influence of these risk premia on the deviation between risk-neutral and physical probabilities of default.

6.7.2 Model

Since we consider a single country, we drop the n index in the following. In order to keep this illustration simple, as e.g. Pan and Singleton (2008), Longstaff et al. (2011) or Monfort and Renne (2012), we assume that the short-term rate is exogenous. In that context, taking into account non-zero recovery rates (see Section 6.3.3) and using Equation (6.6), it is easily shown that

$$B(t,h) = \exp(-hR^*(t,h))E_t^{\mathbb{Q}}\left(\exp\left(-\tilde{\lambda}_{t+1} - \cdots - \tilde{\lambda}_{t+h}\right)\right)$$

(where $R^*(t,h)$ is the risk-free yield of maturity h and $\tilde{\lambda}_t$ is the recovery-adjusted default intensity), which leads to

$$R^*(t,h) - R(t,h) = -\frac{1}{h}\log E_t^{\mathbb{Q}}\left(\exp\left(-\tilde{\lambda}_{t+1} - \cdots - \tilde{\lambda}_{t+h}\right)\right). \quad (6.24)$$

Therefore, the spread versus the risk-free yield depends on the risk-neutral dynamics of the default intensities only. Therefore, we do not have to specify the dynamics of the risk-free short-term rate. The recovery-adjusted default intensity is given by

$$\tilde{\lambda}_t = \alpha_0 + w_t'\alpha_2 w_t,$$

where w_t is a 2×1 vector of (latent) variables. An important advantage of this setting is that one can easily ensure that the intensity remains positive (which is consistent with its interpretation in terms of probabilities, see Equation (6.2)). For instance, this is obtained if α_2 is a positive-definite matrix and if $\alpha_0 > 0$.

The historical and risk-neutral dynamics of the pricing factors w_t respectively read

$$\begin{aligned} w_t &= \Phi w_{t-1} + \varepsilon_t, \quad \text{where } \varepsilon_t \sim \mathcal{N}^{\mathbb{P}}(0, I), \\ w_t &= \varphi^* + \Phi^* w_{t-1} + \varepsilon_t^*, \quad \text{where } \varepsilon_t^* \sim \mathcal{N}^{\mathbb{Q}}(0, I). \end{aligned}$$

Let us denote by S_t the $k \times 1$ vectors of modelled spreads (for k different maturities). In Section 6.5.3, it is shown that in this framework, the spreads are quadratic functions of the factors. Considering the kth spread, that we denote by $S_{k,t}$, we have

$$\underbrace{S_{k,t}}_{\text{observed spread}} = \underbrace{f_k + c_k w_t + w_t' C_k w_t}_{\text{modelled spread}} + \underbrace{\eta_{k,t}}_{\text{pricing error}},$$

where the $\eta_{k,t}$s are Gaussian iid error terms.

6.7.3 *Estimation*

As mentioned in Section 6.6, different filtering techniques can be implemented in order to estimate the model by MLE when the w_ts are latent. Here, we resort to inversion techniques *à la* Chen and Scott (1993). If the model features M (say) latent variables, this technique consists of assuming that M of the observed yields (spreads) or, more generally, that M combinations of observed yields (or spreads), are priced without error by the model. Then, using the M corresponding pricing formula, one can solve for the M latent variables from the M perfectly-priced yields. The remaining yields are then priced with error (the $\eta_{k,t}$s). Appendix A details the computation of the log-likelihood.

The data are weekly and cover the period July 2008–October 2012 (224 dates). The upper plots of Figure 6.2 show the fit of the model. Obviously, because of the estimation approach, two spreads are perfectly modelled (the 2-year and the 10-year spreads). Regarding the other two spreads that are used in the estimation – the 3-year and the 5-year spreads – pricing errors are relatively small (with a standard deviation

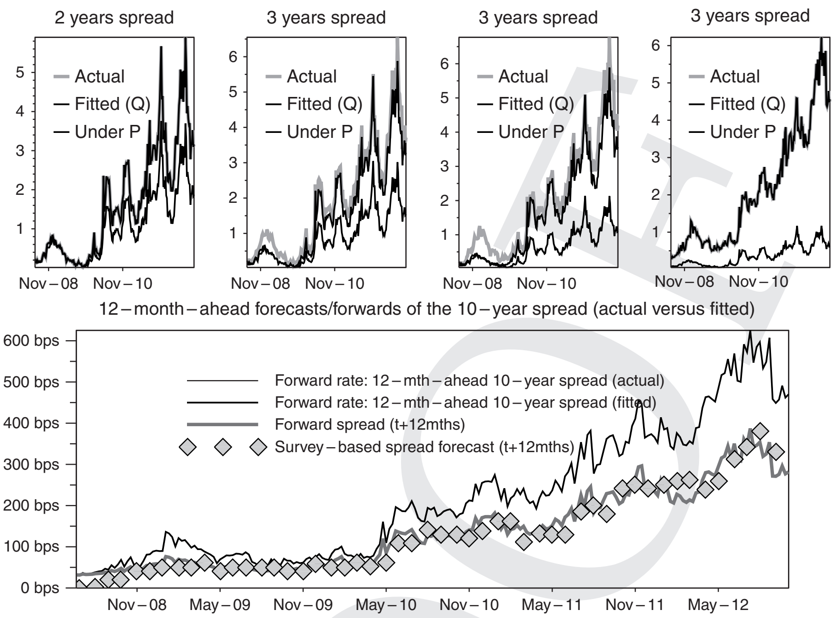

Figure 6.2 Modelling the term-structure of sovereign spreads (Spain vs. Germany): model fit and credit risk premia
The four upper plots show the model fit and also illustrate the influence of credit-risk premia. Indeed, the plots report three kinds of spread: the actual (observed) ones, the (model-based) fitted ones (under $\mathbb{Q}$) and the "physical" ones (under $\mathbb{P}$). The fitted ones are obtained by applying Equation (6.24). The physical ones are the ones that would be observed if the investor were risk-neutral; they are computed in the same way as the fitted ($\mathbb{Q}$) ones, except that the expectation $E^{\mathbb{Q}}$ appearing in Equation (6.24) is replaced by $E^{\mathbb{P}}$ (this amounts to setting the risk sensitivities to zero). The deviation between fitted and physical spreads corresponds to credit risk premia. The lower plot compares 12-month-ahead forwards of the 10-year spread with 12-month-ahead expectation of the 10-year spread (under $\mathbb{P}$). Two kinds of ($\mathbb{P}$) expectation are reported: the model-based ones (solid line) and survey-based ones (diamonds)

of about 20 basis points). The upper plots in Figure 6.2 also illustrate the influence of credit-risk premia. These are discussed in the next subsection.

6.7.4 Credit-risk premia

Credit-risk premia are defined as the deviation between actual spreads and the ones that would prevail if the investor were risk neutral, that is, if the risk sensitivities were nil (see Section 6.4.2 for a formal definition of the risk sensitivities). The upper plots in Figure 6.2 show that these risk premia are sizeable, especially for the longest maturities. This has various important implications in terms of spread analysis.

First, the fact that the historical and the risk-neutral dynamics of the pricing factors w_t are not the same implies that credit-spread forwards cannot be interpreted as market forecasts of future spreads.[15] This is illustrated by the lower plot of Figure 6.2. This plot displays 12-month ahead survey-based forecasts of the Spanish–German spread (source: *Consensus Forecasts*) together with 1-year-forward spreads between Spanish and German 10-year yields.[16] The deviation between the two series is substantial. This suggests that using forwards of spreads to assess market expectations regarding the evolution of future spreads is misleading. In addition, the same plot illustrates the ability of the model to capture this phenomenon. Indeed, it shows that the model-implied 12-month-ahead forecast of the spread (thick solid line) is able to reproduce the level and the main fluctuations in the survey-based forecasts.

Second, the existence of credit-risk premia implies that risk-neutral probabilities of default – for instance those that are backed out from CDS quotes following e.g. O'Kane and Turnbull (2003) – do not coincide with the real-world PDs. Formally, it is easily shown that, in our framework, risk-neutral and historical probabilities of default are given by (see Monfort and Renne, 2012)

$$\begin{aligned}
\mathbb{Q}\left(d_{t+h}=1 \middle| d_t=0,\underline{w}_t\right) &= 1-E^{\mathbb{Q}}\left[\exp\left(-\lambda_{t+1}-\ldots-\lambda_{t+h}\right)\middle|\underline{w}_t\right],\\
\mathbb{P}\left(d_{t+h}=1 \middle| d_t=0,\underline{w}_t\right) &= 1-E^{\mathbb{P}}\left[\exp\left(-\lambda_{t+1}-\ldots-\lambda_{t+h}\right)\middle|\underline{w}_t\right].
\end{aligned} \tag{6.25}$$

The linearization of Equation (6.10) leads to $\lambda_t \approx \tilde{\lambda}_t/(1-\zeta_t)$. Then, assuming that the recovery rate is constant and equal to ζ, the previous probabilities are easily computed using Proposition 6.1 (the expectation terms in Equation (6.25) being multi-horizon Laplace transforms of w_t). Based on a constant recovery rate of 50%, Figure 6.3 compares both kinds of default probability.[17] These computations suggest that real-world probabilities of default are far lower than their risk-neutral counterpart, consistently with the findings of Monfort and Renne (2012).

[15] See e.g. Cochrane and Piazzesi (2005) for an investigation of this phenomenon in the non-defaultable case.

[16] This forward rate (in continuously-compounded terms) is simply given by $(11S(t,11)-S(t,1))/10$.

[17] This recovery rate roughly corresponds to the average of the recovery rates observed for sovereign defaults over the last decade (see Moody's, 2010).

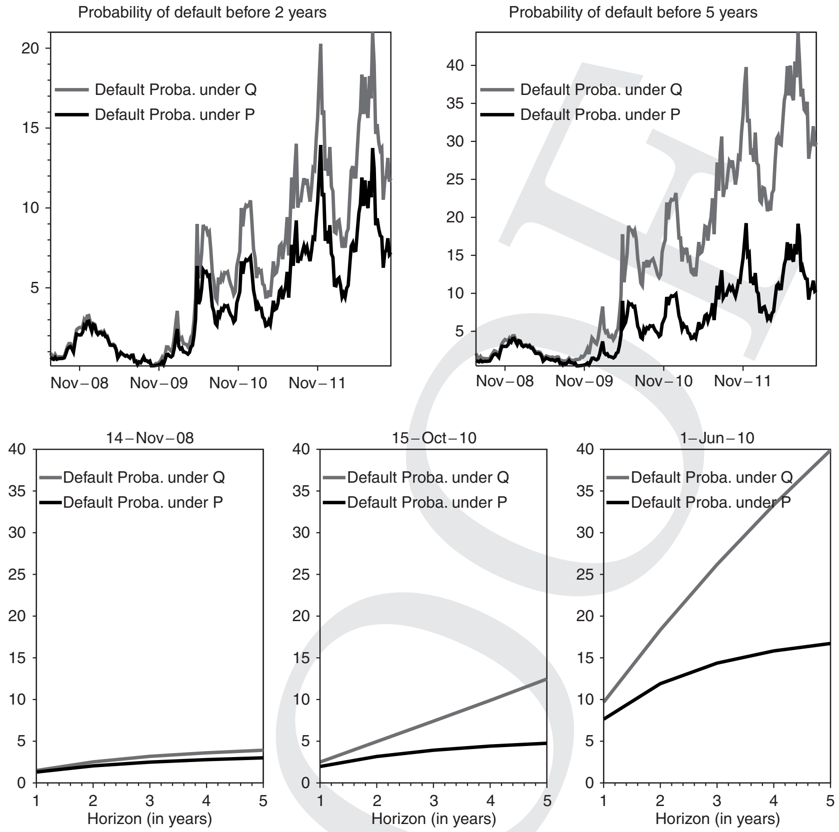

Figure 6.3 Historical vs. risk-neutral probabilities of default
The upper plots present estimates of (market-perceived) probabilities of a Spanish-government default. Risk-neutral probabilities (solid lines) are compared with physical ones (dotted lines). The lower plots present, for three differed dates, the term-structures of risk-neutral and physical probabilities of default (example: for maturity *m*, the plots show the probability of a Spanish-government default before *m* years).

6.8 Conclusion

In this chapter, we investigated the potential of Car processes to model the dynamics of defaultable-bond prices in a no-arbitrage framework. We showed that these processes make it possible to account for sophisticated dynamics of yields and spreads while remaining tractable. In this intensity-based framework, bond prices and yields are obtained in a quasi-explicit form, thanks to a simple recursive algorithm. This ensures that the models building on this framework are amenable to empirical estimation. Several examples of Car processes were provided (regime-switching, autoregressive gamma and quadratic Gaussian processes). These processes can reproduce various key features of observed yields such as non-linearities or stochastic volatilities. In addition, some of these processes can ensure positivity of yields (and/or spreads), which is crucial in the (current) context of extremely low short-term rates.

As an illustration, we exploited our framework to investigate the dynamics of the term structure of Spanish sovereign spreads (vs. Germany). After having estimated the model, we exhibited credit-risk premia, that are defined as those differences between observed credit spreads and the ones that would prevail if the investor were risk neutral. The results suggest that these risk premia are sizeable. This has important implications. Notably, it results in the fact that risk-neutral probabilities of defaults (backed out from spreads under the assumptions that investors are risk-neutral) overestimate physical, or real-world, probabilities of default.

Appendix A Computation of the likelihood (model presented in Section 6.7)

Let $S_{1,t}$ be a 2×1 subvector of S_t which is modelled without pricing error and let $S_{2,t}$ be the vector of the remaining spreads in S_t. In other terms, $\eta_{k,t}$ is equal to 0 if k is one of the two maturities corresponding to $S_{1,t}$ and an iid Gaussian pricing error otherwise. The two equations associated with $S_{1,t}$ can be inverted in order to recover the two latent factors w_t.[18] Let us denote by $q_\theta(w_t, \theta)$ the function that assigns to the latent variables w_t the perfectly-priced spreads $S_{1,t}$. The general term of the likelihood function is the conditional distribution of S_t given $\underline{S}_{t-1}$, which is

$$\begin{aligned}
f_S\left(S_t \middle| \underline{S}_{t-1}\right) &= \left|\det \frac{\partial q_\theta^{-1}(S_{1,t})}{\partial S_{1,t}}\right| f_{S_2,w}\left(S_{2,t}, w_t \middle| \underline{S}_{t-1}\right) \\
&= \left|\det \frac{\partial q_\theta^{-1}(S_{1,t})}{\partial S_{1,t}}\right| f_{S_2}\left(S_{2,t} \middle| w_t, \underline{S}_{t-1}\right) f_w\left(w_t \middle| \underline{S}_{t-1}\right) \\
&= \left|\det \frac{\partial q_\theta(w_t)}{\partial w_t}\right|^{-1} f_{S_2}\left(S_{2,t} \middle| w_t\right) f_w\left(w_t \middle| w_{t-1}\right). \qquad \text{(A.26)}
\end{aligned}$$

[18] At each iteration (new date t), the numerical procedure uses the previously obtained (date $t-1$) factors as initial conditions. This rules out the possibility of obtaining some jumps in the estimated factors that would be due to the existence of several solutions to this system of equations.

The computation of the three terms of the right-hand side of Equation (A.26) is straightforward. The last two are Gaussian distributions. The first one is the inverse of the determinant of a multivariate quadratic function.

Appendix B Estimated model

Parameter estimates are obtained by maximising the log-likelihood computed as detailed in Appendix A. The model is estimated in two steps. In the first step, all parameters in α_0, α_2, Φ, φ^* and Φ^* are estimated. Then, the statistical significance of the parameters is assessed. Those parameters that are not statistically different from zero are then set to zero and excluded from the second step of estimation. The resulting estimated model reads

$$\lambda_t = 10^{-4}\left(\underset{(0.44)}{4.83} + w_t'\begin{bmatrix} \underset{(0.022)}{4.71} & \underset{(0.012)}{3.65} \\ \underset{(0.012)}{3.65} & \underset{(0.014)}{2.83} \end{bmatrix} w_t\right),$$

$$w_t = \begin{bmatrix} \underset{(0.0001)}{0.995} & \underset{(-)}{0} \\ \underset{(-)}{0} & \underset{(0.0000)}{0.992} \end{bmatrix} w_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}^{\mathbb{P}}(0, I),$$

$$w_t = \begin{bmatrix} \underset{(-)}{0} \\ \underset{(0.0001)}{-0.028} \end{bmatrix} + \begin{bmatrix} \underset{(0.0000)}{1.0} & \underset{(0.00005)}{-0.007} \\ \underset{(-)}{0} & \underset{(0.00001)}{0.995} \end{bmatrix} w_{t-1} + \varepsilon_t^*, \quad \varepsilon_t^* \sim \mathcal{N}^{\mathbb{Q}}(0, I).$$

The standard deviations of the parameter estimates are reported in parentheses. The standard-deviation estimates are based on the outer-product estimate of the Fisher information matrix.

References

Andreasen, M. and Meldrum, A. (2011). Likelihood inference in non-linear term structure models: The importance of the zero lower bound. BOE working papers, Bank of England.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, **50**(4), 745–787.

Bank of England (2012). Financial stability report, issue no. 31. Technical report, Bank of England.

Bertholon, H., Monfort, A. and Pegoraro, F. (2008). Econometric asset pricing modelling. *Journal of Financial Econometrics*, **6**(4), 407–456.

Blundell-Wignall, A. and Slovik, P. (2010). A market perspective on the European sovereign debt and banking crisis. *OECD Journal: Financial Market Trends*, no. 2, 9–36.

Borgy, V., Laubach, T., Mesonnier, J.-S. and Renne, J.-P. (2011). Fiscal sustainability, default risk and euro area sovereign bond spreads markets. Working paper series 350, Banque de France.

Borri, N. and Verdelhan, A. (2011). Sovereign risk premia. Atlanta 2010 meetings papers, AFA 2010 Atlanta Meetings Paper.

Campbell, J. Y. and Shiller, R. J. (1991). Yield spreads and interest rate movements: a bird's eye view. *Review of Economic Studies*, **58**(3), 495–514.

Carver, L. (2012). Irksome irc. *Risk Magazine*, **25**(6), 17–20.

Chen, R.-R. and Scott, L. (1993). Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *Journal of Fixed Income*, **3**, 14–31.

Chen, R.-R., Cheng, X., Fabozzi, F. J. and Liu, B. (2008). An explicit, multifactor credit default swap pricing model with correlated factors. *Journal of Financial and Quantitative Analysis*, **43**(01), 123–160.

CMA Datavision (2011). CMA global sovereign debt credit risk report, 1st quarter 2011. Technical report.

Cochrane, J. H. and Piazzesi, M. (2005). Bond risk premia. *American Economic Review*, **95**(1), 138–160.

Darolles, S., Gourieroux, C. and Jasiak, J. (2006). Structural Laplace transform and compound autoregressive models. *Journal of Time Series Analysis*, **27**(4), 477–503.

de Jong, F. (2000). Time series and cross-section information in affine term-structure models. *Journal of Business & Economic Statistics*, **18**(3), 300–314.

Duffie, D. and Singleton, K. J. (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies*, **12**(4), 687–720.

Gilchrist, S. and Zakrajšek, E. (2011). Credit spreads and business cycle fluctuations. Working Paper 17021, National Bureau of Economic Research.

Gourieroux, C. and Jasiak, J. (2006). Autoregressive gamma processes. *Journal of Forecasting*, **25**, 129–152.

Gourieroux, C. and Jasiak, J. (2009). Value at risk. In *Handbook of Financial Econometrics* (Y. Ait-Sahalia and L. P. Hansen, eds.), pp. 553–609. Amsterdam, The Netherlands: Elsevier.

Gourieroux, C. and Monfort, A. (2007). Econometric specification of stochastic discount factor models. *Journal of Econometrics*, **136**(2), 509–530.

Gourieroux, C. and Sufana, R. (2011). Discrete time Wishart term structure models. *Journal of Economic Dynamics and Control*, **35**(6), 815–824.

Gourieroux, C., Monfort, A. and Polimenis, V. (2006). Affine models for credit risk analysis. *Journal of Financial Econometrics*, **4**(3), 494–530.

Gourieroux, C., Monfort, A., Pegoraro, F. and Renne, J.-P. (2012). Regime-switching and bond pricing. Journal of Financial Econometrics lecture, Sofie annual meeting.

Hannoun, H. (2011). Sovereign risk in bank regulation and supervision: where do we stand? Speech by the BIS Deputy general manager at the Financial Stability Institute high-level meeting, Abu Dhabi, 26 October 2011, Bank for International Settlements.

Huang, J. Z. and Huang, M. (2003). How much of the corporate treasury yield spread is due to credit risk? Working paper, Penn State and Stanford Universities.

Joslin, S., Singleton, K. J., and Zhu, H. (2011). A new perspective on Gaussian dynamic term structure models. *Review of Financial Studies*, **24**(3), 926–970.

Kim, C.-J. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics*, **60**(1-2), 1–22.

Kim, D. H. and Singleton, K. J. (2012). Term structure models and the zero-lower bound: an empirical investigation of Japanese yields. Technical report.

Le, A., Singleton, K. J. and Dai, Q. (2010). Discrete-time affine-q term structure models with generalized market prices of risk. *Review of Financial Studies*, **23**(5), 2184–2227.

Litterman, R. and Iben, T. (1991). Corporate bond valuation and the term structure of credit spreads. *The Journal of Portfolio Management*, pp. 52–64.

Longstaff, F., Pan, J., Pedersen, L. and Singleton, K. (2011). How sovereign is sovereign credit risk? *Amercian Economic Journal: Macroeconomics*, **3**, 75–103.

Monfort, A. and Pegoraro, F. (2007). Switching Varma term structure models. *Journal of Financial Econometrics*, **5**(1), 105–153.

Monfort, A. and Pegoraro, F. (2012). Asset pricing with second order Esscher transforms. *Journal of Banking and Finance*, **36**(6), 1678–1687.

Monfort, A. and Renne, J.-P. (2012). Decomposing euro-area sovereign spreads: credit and liquidity risks. c.r.e.d.i.t. conference.

Monfort, A. and Renne, J.-P. (2013). Default, liquidity and crises: an econometric framework. *Journal of Financial Econometrics*, **11**(2).

Moody's (2010). Sovereign default and recovery rates, 1983–2009. Global credit policy, Moody's Investors Service.

Mueller, P. (2008). Credit spread and real activity. Technical report, Columbia Business School.

Nouy, D. (2011). Is sovereign risk properly addressed by financial regulation? *Financial Stability Review*, **16**, 95–106.

O'Kane, D. and Turnbull, S. (2003). Valuation of credit default swaps. Fixed Income – Quantitative credit research, Lehman Brothers.

Pan, J. and Singleton, K. (2008). Default and recovery implicit in the term structure of sovereign CDS spreads. *Journal of Finance*, **63**(5), 2345–2384.

Renne, J.-P. (2012). A model of the euro-area yield curve with discrete policy rates. Working paper series 395, Banque de France.

Sengupta, R. and Tam, Y. M. (2008). The LIBOR-OIS spread as a summary indicator. *Monetary Trends*, November.

Taylor, J. B. and Williams, J. C. (2009). A black swan in the money market. *American Economic Journal: Macroeconomics*, **1**(1), 58–83.