Online Appendix

Affine Modeling of Credit Risk, Pricing of Credit Events and Contagion

Alain Monfort, Fulvio Pegoraro, Guillaume Roussellet and Jean-Paul Renne

A.1 Proofs

A.1.1 Proof of Proposition 2.2

$$
\begin{aligned}
&\mathbb{E}\left[\exp(u_y'y_t + u_\delta'\delta_t)\,\middle|\,\mathcal{F}_{t-1}\right] = \mathbb{E}\left[\mathbb{E}\left[\exp(u_y'y_t + u_\delta'\delta_t)\,\middle|\,\mathcal{F}_t^*,\, \mathcal{D}_{t-1}\right]\,\middle|\,\mathcal{F}_{t-1}\right]\\
=\;& \mathbb{E}\left[\exp\left(u_y'y_t + \lambda_t^{\mathbb{P}\prime}\,\frac{u_\delta\odot\mu_\delta}{\mathbf{1}-u_\delta\odot\mu_\delta}\right)\,\middle|\,\mathcal{F}_{t-1}\right]\\
=\;& \exp\left(\left(\frac{u_\delta\odot\mu_\delta}{\mathbf{1}-u_\delta\odot\mu_\delta}\right)'(\alpha_\lambda + \mathbf{C}'\delta_{t-1})\right)\times\mathbb{E}\left[\exp\left(\left(u_y + \beta_\lambda^{(y)}\frac{u_\delta\odot\mu_\delta}{\mathbf{1}-u_\delta\odot\mu_\delta}\right)'y_t\right)\,\middle|\,\mathcal{F}_{t-1}\right]\\
=\;& \exp\left(\left(\frac{u_\delta\odot\mu_\delta}{\mathbf{1}-u_\delta\odot\mu_\delta}\right)'(\alpha_\lambda + \mathbf{C}'\delta_{t-1})\right)\times\varphi_{y_{t-1}}^{\mathbb{P}}\left(u_y + \beta_\lambda^{(y)}\frac{u_\delta\odot\mu_\delta}{\mathbf{1}-u_\delta\odot\mu_\delta}\right).
\end{aligned}
$$

Transforming the conditional Laplace transform of y_t using Assumption 4, we obtain the desired result. ■

A.1.2 Proof of Proposition 2.3

The fact that $\varphi_{y_{t-1}}^{\mathbb{P}}$ (defined in Assumption 4) is exponential affine in w_{t-1} directly stems from the knowledge of the Laplace transform of the non-central Gamma distribution (see Monfort et al. 2017). (Functions $A_y^{(y)}$, $A_y^{(\delta)}$ and B_y are deduced from the same Laplace transform.) Hence Assumption 4 is satisfied, and Proposition 2.2 therefore applies. ■

A.1.3 Proof of Propositions 2.7

Proposition 2.4 gives:

$$
\varphi_{w_t}^{\mathbb{Q}}(u_1) = \mathbb{E}^{\mathbb{Q}}\left[\exp(u_1'\,w_{t+1}\middle|\mathcal{F}_t\right] = \exp\left(A_w^{\mathbb{Q}(y)}(u_1)'\,y_t + A_w^{\mathbb{Q}(\delta)}(u_1)'\,\delta_t + B_w^{\mathbb{Q}}(u_1)\right),
$$

which shows that Equations (10) and (11) are satisfied for $h=1$.

Let us now assume that it holds for a given horizon h, with $h>0$. We then have

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[\exp\left(u_1'w_{t+1}+\cdots+u_{h+1}'w_{t+h+1}\right)\middle|\mathcal{F}_t\right]\\
=\;& \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\exp\left(u_1'w_{t+1}+\cdots+u_{h+1}'w_{t+h+1}\right)\middle|\mathcal{F}_{t+1}\right]\middle|\mathcal{F}_t\right]\\
&\text{(using the law of iterated expectations)}\\
=\;& \mathbb{E}^{\mathbb{Q}}\left[\exp(u_1'w_{t+1})\mathbb{E}^{\mathbb{Q}}\left[\exp\left(u_2'w_{t+2}+\cdots+u_{h+1}'w_{t+h+1}\right)\middle|\mathcal{F}_{t+1}\right]\middle|\mathcal{F}_t\right]\\
=\;& \mathbb{E}^{\mathbb{Q}}\left[\exp(u_1'w_{t+1})\exp\left[\mathcal{A}_h(u_2,\ldots,u_{h+1})'\,w_{t+1}+\mathcal{B}_h(u_2,\ldots,u_{h+1})\right]\middle|\mathcal{F}_t\right]\\
&\text{(using the induction assumption)}\\
=\;& \mathbb{E}^{\mathbb{Q}}\left(\exp\left[\{u_1+\mathcal{A}_h(u_2,\ldots,u_{h+1})\}'w_{t+1}+\mathcal{B}_h(u_2,\ldots,u_{h+1})\right]\middle|\mathcal{F}_t\right)\\
=\;& \exp\left[A_w^{\mathbb{Q}}(u_1+\mathcal{A}_h(u_2,\ldots,u_{h+1}))'\,w_{t+1}+B_w^{\mathbb{Q}}(u_1+\mathcal{A}_h(u_2,\ldots,u_{h+1}))+\mathcal{B}_h(u_2,\ldots,u_{h+1})\right],
\end{aligned}
$$

which implies that Equation (10) then also holds for $h+1$, leading to the result. ■

A.1.4 Proof of Lemma 3.1

For a given u_1 and $u_2 \geq 0$

$$\lim_{u_2 \to +\infty} \mathbb{E}\left[\exp(u_1' Z_1 - u_2\, Z_2)\right] = \mathbb{E}\left[\exp(u_1' Z_1)\, \mathbb{1}_{\{Z_2=0\}}\right] + \lim_{u_2 \to +\infty} \mathbb{E}\left[\exp(u_1' Z_1 - u_2\, Z_2)\, \mathbb{1}_{\{Z_2>0\}}\right],$$

and since in the second term on the right-hand side $\exp(-u_2\, Z_2)\, \mathbb{1}_{\{Z_2>0\}} \to 0$ when $u_2 \to +\infty$, relation (3.1) is a consequence of the Lebesgue theorem. ■

A.1.5 Proof of Proposition 3.2 (Bond Pricing under the RMV Convention)

Consider the case of a one-period bond, on date t. According to Definition 3.1, the recovery value of date $t+1$ is the price of the bond if there had been no credit event. For this one-period bond, $t+1$ is also the maturity date, and the recovery value is therefore 1. As a result, under the RMV convention (Definition 3.1) and with the recovery rate assumption of Equation (15), the price of a one-period bond is given by:

$$B_i(t,1) = \exp(-r_t)\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{d_{i,t+1}=0\}} \times 1 + \exp(-\delta_{i,t+1})\mathbb{1}_{\{d_{i,t+1}=1\}} \times 1|\mathcal{F}_t],$$

where the first "1" on the right-hand side stands for the price of the bond in the case of no default and the second "1" stands for the recovery value. Using that $\mathbb{1}_{\{d_{i,t+1}=0\}} = \mathbb{1}_{\{d_{i,t+1}=0\}} \exp(-\delta_{i,t+1})$, the previous equation becomes:

$$B_i(t,1) = \mathbb{E}^{\mathbb{Q}}[\exp(-r_t - \delta_{i,t+1})|\mathcal{F}_t], \tag{a.1}$$

which proves Equation (18) for $h = 1$.

Consider now the pricing of a two-period bond, as of date t. The definitions of the recovery value and of the recovery rate in the RMV case – Definition 3.1 and Equation (15), respectively – imply that if a default occurs on date $t+1$, the payoff is $\exp(-\delta_{i,t+1})\widetilde{B}_i(t+1,1)$. In the previous expression, according to Definition 3.1, $\widetilde{B}_i(t+1,1)$ is the price of the bond at time $t+1$ "if there had been no credit event on this date" (Definition 3.1); this recovery value $\widetilde{B}_i(t+1,1)$ is therefore equal to $B_i(t+1,1)$, whose expression is given by Equation (a.1). This implies that:

$$\begin{aligned}
B_i(t,2) &= \exp(-r_t)\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{d_{i,t+1}=0\}} \times B_i(t+1,1) + \exp(-\delta_{i,t+1})\mathbb{1}_{\{d_{i,t+1}=1\}} \times B_i(t+1,1)|\mathcal{F}_t] \\
&= \exp(-r_t)\mathbb{E}^{\mathbb{Q}}[\exp(-\delta_{i,t+1})\mathbb{1}_{\{d_{i,t+1}=0\}}B_i(t+1,1) + \exp(-\delta_{i,t+1})\mathbb{1}_{\{d_{i,t+1}=1\}}B_i(t+1,1)|\mathcal{F}_t] \\
&\quad \text{(using again } \mathbb{1}_{\{d_{i,t+1}=0\}} = \mathbb{1}_{\{d_{i,t+1}=0\}} \exp(-\delta_{i,t+1})) \\
&= \exp(-r_t)\mathbb{E}^{\mathbb{Q}}[\exp(-\delta_{i,t+1})B_i(t+1,1)|\mathcal{F}_t] \\
&= \exp(-r_t)\mathbb{E}^{\mathbb{Q}}[\exp(-\delta_{i,t+1})\exp(-r_{t+1} - \delta_{i,t+2})|\mathcal{F}_t],
\end{aligned}$$

where the last equality uses (a.1). This proves Equation (18) for $h = 2$. Iterating on the previous arguments clearly proves Equation (18) for any $h \in \mathbb{N}$.

Let us now prove relation (19). Given $r_t = \xi_0 + \xi' w_t$ and $\delta_{i,t} = \mathrm{e}_{\delta_i}' w_t$, we can write:

$$\begin{aligned}
B_i(t,h) &= \mathbb{E}^{\mathbb{Q}}\left\{\exp\left[-\sum_{\ell=0}^{h-1}(r_{t+\ell} + \delta_{i,t+\ell+1})\right] | \mathcal{F}_t\right\} \\
&= \exp\left[-\xi_0 - \xi' w_t\right] \mathbb{E}^{\mathbb{Q}}\left\{\exp\left[-\sum_{\ell=1}^{h-1}(\xi_0 + \xi' w_{t+\ell}) - \sum_{\ell=1}^{h}\mathrm{e}_{\delta_i}' w_{t+\ell}\right] | \mathcal{F}_t\right\} \\
&= \exp\left[-\xi_0\, h - \xi' w_t\right] \mathbb{E}^{\mathbb{Q}}\Big\{\exp\big[-(\xi + \mathrm{e}_{\delta_i})' w_{t+1} - \ldots - (\xi + \mathrm{e}_{\delta_i})' w_{t+h-1} - \mathrm{e}_{\delta_i}' w_{t+h}\big] \,\big|\, \mathcal{F}_t\Big\} \\
&= \exp\left[-\xi_0\, h - \xi' w_t\right] \varphi^{\mathbb{Q}}_{w_t(h)}\left(-\xi - \mathrm{e}_{\delta_i}, -\mathrm{e}_{\delta_i}\right),
\end{aligned}$$

which leads to the result. ■

A.1.6 Proof of Proposition 3.3 (Bond Pricing under the RFV Convention)

Given relation (17), as well as the recovery assumption (14) and $\Pi_{i,t}(h) = 1$, the price of the defaultable zero-coupon bond of interest can be written as:

$$\begin{aligned} B_i(t,h) &= \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell}\right) \exp\left(-\omega_{i,0} - \omega_i^{(w)'} w_{t+k}\right) \mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}} \mid \mathcal{F}_t \right\} \\ &\quad - \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell}\right) \exp\left(-\omega_{i,0} - \omega_i^{(w)'} w_{t+k}\right) \mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}} \mid \mathcal{F}_t \right\} \\ &\quad + \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=0}^{h-1} r_{t+\ell}\right) \mathbb{1}_{\{\delta_{i,t:t+h}=\mathbf{0}\}} \mid \mathcal{F}_t \right\}. \end{aligned} \tag{a.2}$$

Using Lemma 3.1, the previous relation becomes:

$$\begin{aligned} B_i(t,h) &= \lim_{u\to+\infty} \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell} - u\sum_{\ell=0}^{k-1}\delta_{i,t+\ell}\right) \exp\left(-\omega_{i,0} - \omega_i^{(w)'} w_{t+k}\right) \mid \mathcal{F}_t \right\} \\ &\quad - \lim_{u\to+\infty} \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell} - u\sum_{\ell=0}^{k}\delta_{i,t+\ell}\right) \exp\left(-\omega_{i,0} - \omega_i^{(w)'} w_{t+k}\right) \mid \mathcal{F}_t \right\} \\ &\quad + \lim_{u\to+\infty} \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=0}^{h-1} r_{t+\ell} - u\sum_{\ell=0}^{h}\delta_{i,t+\ell}\right) \mid \mathcal{F}_t \right\}, \end{aligned}$$

which gives

$$\begin{aligned} & B_i(t,h) \\ &= \lim_{u\to+\infty} e^{-\omega_{i,0}} \sum_{k=1}^{h} e^{-k\xi_0 - (\xi + ue_{\delta_i})' w_t} \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=1}^{k-1} (\xi + ue_{\delta_i})' w_{t+\ell} - \omega_i^{(w)'} w_{t+k}\right) \mid \mathcal{F}_t \right\} \\ &\quad - \lim_{u\to+\infty} e^{-\omega_{i,0}} \sum_{k=1}^{h} e^{-k\xi_0 - (\xi + ue_{\delta_i})' w_t} \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=1}^{k-1} (\xi + ue_{\delta_i})' w_{t+\ell} - \left(\omega_i^{(w)} + ue_{\delta_i}\right)' w_{t+k}\right) \mid \mathcal{F}_t \right\} \\ &\quad + \lim_{u\to+\infty} e^{-h\xi_0 - (\xi + ue_{\delta_i})' w_t} \mathbb{E}^{\mathbb{Q}} \left\{ \exp\left(-\sum_{\ell=1}^{h-1} (\xi + ue_{\delta_i})' w_{t+\ell} - ue_{\delta_i}' w_{t+h}\right) \mid \mathcal{F}_t \right\}, \end{aligned}$$

which leads to Equation (20) using the definition of the multi-horizon Laplace transform $\varphi^{\mathbb{Q}}_{w_t}$ (see Proposition 2.7). ■

A.1.7 Proof of Proposition 3.4 (CDS pricing)

Using Lemma 3.1, relation (21) can be written as:

$$\begin{aligned} \mathrm{PB}_i(t,h) &= \mathcal{S}_i(t,h) \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}} \left[\exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell}\right) \mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}} \,\middle|\, \mathcal{F}_t \right] \\ &= \mathcal{S}_i(t,h) \lim_{u\to+\infty} \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}} \left[\exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell} - \sum_{\ell=0}^{k} u\delta_{i,t+\ell}\right) \,\middle|\, \mathcal{F}_t \right] \\ &= \mathcal{S}_i(t,h) \lim_{u\to+\infty} \sum_{k=1}^{h} e^{-k\xi_0 - (\xi + ue_{\delta_i})' w_t} \times \varphi^{\mathbb{Q}}_{w_t(k)}\left(-\xi - ue_{\delta_i}, -ue_{\delta_i}\right). \end{aligned} \tag{a.3}$$

Let us then split relation (22) as:

$$\begin{aligned}\mathrm{PS}_i(t,h) &= \sum_{k=1}^{h}\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{\ell=0}^{k-1}r_{t+\ell}\right)(1-\varrho_{i,t+k})\left[\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}}-\mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}\right]\big|\,\mathcal{F}_t\right]\\ &= \sum_{k=1}^{h}\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{\ell=0}^{k-1}r_{t+\ell}\right)\left[\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}}-\mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}\right]\big|\,\mathcal{F}_t\right]\\ &\quad-\sum_{k=1}^{h}\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{\ell=0}^{k-1}r_{t+\ell}\right)\varrho_{i,t+k}\left[\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}}-\mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}\right]\big|\,\mathcal{F}_t\right].\end{aligned} \tag{a.4}$$

Then, let us rewrite the RFV pricing formula (a.2) for different values of the recovery rate. Using the notation

$$\begin{aligned}&B_i^{\mathrm{RFV}}\left(t,h;\,\omega_{i,0},\,\omega_i^{(w)}\right)\\ &= \sum_{k=1}^{h}\mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{k-1}r_{t+\ell}\right)\exp\left(-\omega_{i,0}-\omega_i^{(w)\prime}w_{t+k}\right)\left[\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}}-\mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}\right]\,|\,\mathcal{F}_t\right\}\\ &\quad+\mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{h-1}r_{t+\ell}\right)\mathbb{1}_{\{\delta_{i,t:t+h}=\mathbf{0}\}}\,|\,\mathcal{F}_t\right\},\end{aligned} \tag{a.5}$$

we obtain:

$$\begin{aligned}B_i^{\mathrm{RFV}}(t,h;\,0,\,\mathbf{0}) &= \sum_{k=1}^{h}\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{\ell=0}^{k-1}r_{t+\ell}\right)\left[\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}}-\mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}\right]\big|\,\mathcal{F}_t\right]\\ &\quad+\mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{h-1}r_{t+\ell}\right)\mathbb{1}_{\{\delta_{i,t:t+h}=\mathbf{0}\}}\,|\,\mathcal{F}_t\right\},\end{aligned} \tag{a.6}$$

such that:

$$\mathrm{PS}_i(t,h) = B_i^{\mathrm{RFV}}(t,h;\,0,\,\mathbf{0}) - B_i^{\mathrm{RFV}}\left(t,h;\,\omega_{i,0},\,\omega_i^{(w)}\right). \tag{a.7}$$

The price of default protection (23) is easily obtained by imposing (a.3) = (a.7), thus proving Proposition 3.4. ■

A.1.8 Multi Currency Credit Default Swap Pricing

In this subsection, we extend the CDS pricing formula provided by Proposition 3.4 (Subsection 3.3) to the case where the currency of denomination of the CDS is not the domestic one (that is the currency in which the assets of the reference entity are denominated). Typically, a CDS protection on sovereign bonds is frequently available in a foreign and in the domestic currency. The reason behind the development of foreign-currency-denominated CDS is the protection they provide not only against the credit event but also against the associated potential depreciation of the domestic currency with respect to the foreign one (see Section 4.7).

Consider a CDS denominated in a foreign currency. We denote by s_t the log of the exchange rate between the domestic and the foreign currency, where the exchange rate is defined as the price in the domestic currency of one unit of foreign currency. That is, an increase in s_t corresponds to a depreciation of the domestic currency. Let us denote by $\mathcal{S}_i^f(t,h)$ the foreign-currency-denominated CDS spread, set at date t with maturity $t+h$.

Both the notional and the premium payments of a CDS are expressed in the currency of denomination. We assume in the following that the notional of the CDS is equal to one unit of the foreign currency (i.e. to $\exp(s_t)$ units of the domestic currency). The CDS spread is such that the present value of the payments made by the protection buyer (the fixed leg) is equal to the present value of the payment made by the protection seller in case of default (the floating leg).

As far as the fixed leg is concerned, if entity i has not defaulted at date $t+k$ ($\leq t+h$), the cash flow on this date, expressed in the domestic currency, is: $\mathcal{S}_i^f(t,h)\exp(s_{t+k})$. The present value of the fixed-leg payments, expressed in the domestic currency ($PB_i^f(t,h)$, say), is thus given by:

$$PB_i^f(t,h) = \mathcal{S}_i^f(t,h) \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}}\left[\exp\left(s_{t+k} - \sum_{\ell=0}^{k-1} r_{t+\ell}\right) \mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}} \mid \underline{w_t}\right]. \qquad \text{(a.8)}$$

Under the RFV convention, the protection seller will make a payment of $(1-\varrho_{i,t+k})\exp(s_{t+k})$ (the Loss-Given-Default) at date $t+k$ in case of default over the time interval $]t+k-1,t+k]$. (Observe that the recovery rate $\varrho_{i,t+k}$ is the same as for a CDS denominated in the domestic currency.) The present value of this promised payment, expressed in the domestic currency, is:

$$PS_i^f(t,h) = \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}}\left[\exp\left(s_{t+k} - \sum_{\ell=0}^{k-1} r_{t+\ell}\right)(1-\varrho_{i,t+k})\left(\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}} - \mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}\right) \mid \underline{w_t}\right], \qquad \text{(a.9)}$$

and the date-t CDS spread $\mathcal{S}_i^f(t,h)$ is such that $PB_i^f(t,h) = PS_i^f(t,h)$.

Assume, consistently with Assumption 7 (Equation 14), that $\varrho_{i,t} = \exp\left(-\omega_{i,0} - \omega_i^{(w)\prime}\, w_t\right)$ and:

$$\Delta s_t = \chi + u_s' w_t. \qquad \text{(a.10)}$$

We then have the following:

Proposition a.1 *The no-arbitrage CDS spread $\mathcal{S}_i^f(t,h)$, negotiated at date t and associated to a maturity-h CDS denominated in the foreign currency whose exchange rate (w.r.t. the domestic currency) is defined by Equation (a.10), is given by:*

$$\mathcal{S}_i^f(t,h) = \frac{B_{i,f}^{\mathrm{RFV}}(t,h;\,0,\,\mathbf{0}) - B_{i,f}^{\mathrm{RFV}}\left(t,h;\,\omega_{i,0},\,\omega_i^{(w)}\right)}{\lim\limits_{u\to+\infty}\sum_{k=1}^{h} e^{-k\xi_0-(k+1)\chi-(\xi+u\mathrm{e}_{\delta_i}+u_s)'w_t} \times \varphi_{w_t(k)}^{\mathbb{Q}}\left(-\xi - u\mathrm{e}_{\delta_i} - u_s, -u\mathrm{e}_{\delta_i} - u_s\right)}, \qquad \text{(a.11)}$$

where:

$$\begin{aligned}
& B_{i,f}^{\mathrm{RFV}}\left(t,h;\,\omega_{i,0},\,\omega_i^{(w)}\right) \\
= & \lim_{u\to+\infty}\left[\sum_{k=1}^{h} e^{-\omega_{i,0}-k\xi_0-(k+1)\chi-(\xi+u\mathrm{e}_{\delta_i}+u_s)'w_t}\left(\varphi_{w_t(k)}^{\mathbb{Q}}\left[-\xi - u\mathrm{e}_{\delta_i} - u_s,\, -\omega_i^{(w)} - u_s\right]\right.\right. \\
& \left. - \varphi_{w_t(k)}^{\mathbb{Q}}\left[-\xi - u\mathrm{e}_{\delta_i} - u_s,\, -u\mathrm{e}_{\delta_i} - \omega_i^{(w)} - u_s\right]\right) \\
& \left. + \; e^{-h\xi_0-(h+1)\chi-(\xi+u\mathrm{e}_{\delta_i}+u_s)'w_t}\varphi_{w_t(h)}^{\mathbb{Q}}\left[-\xi - u\mathrm{e}_{\delta_i} - u_s,\, -u\mathrm{e}_{\delta_i} - u_s\right]\right],
\end{aligned} \qquad \text{(a.12)}$$

is the date-t price of a foreign-currency-denominated bond paying (in domestic currency) $\exp(s_t)\,\varrho_{i,t+k}$ *at* $t+k$ *if* $\tau_i \in]t+k-1, t+k]$ *and paying* $\exp(s_{t+h})$ *at* $t+h$ *if the default does not happen during the bond lifetime.*

Proof *Straightforward generalization of Proposition 3.4.* ■

It can be noted that the CDS spread $\mathcal{S}_i(t,h)$ of a CDS denominated in the domestic currency (given by Proposition 3.4) coincides with the one resulting from Proposition a.1 when $\chi = 0$ and $u_s = 0$.

A.1.9 Defaultable Bonds Pricing under Recovery of Treasury (RT)

The recovery of Treasury (RT), introduced by Jarrow and Turnbull (1995) and Longstaff and Schwartz (1995), states that, upon issuer default, the creditor receives a fraction (corresponding to the recovery rate) of the present value of the principal. This means that, in case of default at date $\tau_i = t+k$, the payoff is:

$$\varrho_{i,t+k} \times \Pi_{t+k}(h-k) = \exp\left(-\omega_{i,0} - \omega_i^{(w)\prime} w_{t+k}\right) \times B^*(t+k, h-k)\,, \tag{a.13}$$

where $B^*(t,h) = \mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{h-1} r_{t+\ell}\right) \mid \mathcal{F}_t\right\}$ is the date-t market price of an otherwise equivalent default-free zero-coupon bond maturing at $t+h$. Using Proposition 3.1, we have:

$$\begin{aligned} B^*(t,h) &= \exp\left\{-h\xi_0 + \left[\mathcal{A}_{h-1}(-\xi, -\xi) - \xi\right]' w_t + \mathcal{B}_{h-1}(-\xi, -\xi)\right\} \\ &=: \exp\left(\mathcal{A}_h^{*\prime} w_t + \mathcal{B}_h^*\right). \end{aligned}$$

In this case, we have the following proposition.

Proposition a.2 *Under the RT convention, the no-arbitrage price at date* $t < \tau_i$ *of a defaultable zero-coupon bond issued by an entity* i *and maturing in* h *periods is given by:*

$$\begin{aligned} & B_i(t,h) \\ = & \lim_{u \to +\infty} e^{-r_t} \Bigg[e^{-\omega_{i,0}} \sum_{k=1}^{h} e^{\mathcal{B}_k^* - (k-1)\xi_0} \left(\varphi^{\mathbb{Q}}_{w_t(k)}\left[-\xi - u e_{\delta_i}, \mathcal{A}_k^* - \omega_i^{(w)}\right] - \varphi^{\mathbb{Q}}_{w_t(k)}\left[-\xi - u e_{\delta_i}, \mathcal{A}_k^* - u e_{\delta_i} - \omega_i^{(w)}\right]\right) \\ & + \; e^{-(h-1)\xi_0} \varphi^{\mathbb{Q}}_{w_t(h)}\left[-\xi - u e_{\delta_i}, -u e_{\delta_i}\right] \Bigg]. \end{aligned} \tag{a.14}$$

Proof The price $B_i(t,h)$ is equal to:

$$\begin{aligned} & \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell}\right) \exp\left(-\omega_{i,0} - \omega_i^{(w)\prime} w_{t+k}\right) \times B^*(t+k, h-k)\, \mathbb{1}_{\{\delta_{i,t:t+k-1} = \mathbf{0}\}} \mid \mathcal{F}_t\right\} \\ & - \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell}\right) \exp\left(-\omega_{i,0} - \omega_i^{(w)\prime} w_{t+k}\right) \times B^*(t+k, h-k)\, \mathbb{1}_{\{\delta_{i,t:t+k} = \mathbf{0}\}} \mid \mathcal{F}_t\right\} \\ & + \mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{h-1} r_{t+\ell}\right) \mathbb{1}_{\{\delta_{i,t:t+h} = \mathbf{0}\}} \mid \mathcal{F}_t\right\}. \end{aligned}$$

Let us slightly adapt the notation introduced in Equation (a.12) as follows:

$$\begin{aligned}
& B_i^{\mathrm{RFV}}\left(t,h;\,\{\omega_{i,0,1:h}\},\,\left\{\omega_{i,1:h}^{(w)}\right\}\right)\\
= & \sum_{k=1}^{h}\mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{k-1}r_{t+\ell}\right)\exp\left(-\omega_{i,0,k}-\omega_{i,k}^{(w)\prime}w_{t+k}\right)\left[\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}}-\mathbb{1}_{\{\delta_{i,t:t+k}\}}\right]\mid\mathcal{F}_t\right\}\\
& +\mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{h-1}r_{t+\ell}\right)\mathbb{1}_{\{\delta_{i,t:t+h}=\mathbf{0}\}}\mid\mathcal{F}_t\right\}.
\end{aligned}$$

(which would be the price of a defaultable bond under RFV if the loadings of the recovery rate depended on the horizon at which the entity defaults). The price of a defaultable bond under the RT convention then writes:

$$\begin{aligned}
B_i(t,h) = & \sum_{k=1}^{h}\mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{k-1}r_{t+\ell}\right)\exp\left(\mathcal{B}_k^*-\omega_{i,0}+\left(\mathcal{A}_k^*-\omega_i^{(w)}\right)'w_{t+k}\right)\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}}\mid\mathcal{F}_t\right\}\\
& -\sum_{k=1}^{h}\mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{k-1}r_{t+\ell}\right)\exp\left(\mathcal{B}_k^*-\omega_{i,0}+\left(\mathcal{A}_k^*-\omega_i^{(w)}\right)'w_{t+k}\right)\mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}\mid\mathcal{F}_t\right\}\\
& +\mathbb{E}^{\mathbb{Q}}\left\{\exp\left(-\sum_{\ell=0}^{h-1}r_{t+\ell}\right)\mathbb{1}_{\{\delta_{i,t:t+h}=\mathbf{0}\}}\mid\mathcal{F}_t\right\}\\
= & B_i^{\mathrm{RFV}}\left(t,h;\,\{\omega_{i,0}-\mathcal{B}_{1:h}^*\},\,\left\{\omega_{i,1:h}^{(w)}-\mathcal{A}_{1:h}^*\right\}\right).
\end{aligned}$$

Using the formulation of the multi-horizon Laplace transform leads to the result. ■

A.2 Semi-Strong VAR Representation of the Model

The model described by Assumptions 2, 3 and 5 can be written as follows:

$$\begin{aligned}
\mathsf{P}_{y_j,t}\big|\mathcal{F}_{t-1} & \overset{\mathbb{P}}{\sim} \mathcal{P}\left(\alpha_{y_j}+\beta_{y_j}^{(y)\prime}y_{t-1}+\mathbf{I}_j'\delta_{t-1}\right)\\
y_{j,t}\big|\mathcal{F}_{t-1},\mathsf{P}_{y_j,t} & \overset{\mathbb{P}}{\sim} \Gamma_{\nu_{y_j}+\mathsf{P}_{y_j,t}}\left(\mu_{y_j}\right)\\
\mathsf{P}_{\delta_j,t}\big|\mathcal{F}_{t-1},y_t & \overset{\mathbb{P}}{\sim} \mathcal{P}\left(\alpha_{\lambda_j}+\beta_{\lambda_j}^{(y)\prime}y_t+\mathbf{C}_j'\delta_{t-1}\right)\\
\delta_{j,t}\big|\mathsf{P}_{\delta_j,t},\mathcal{F}_{t-1},y_t & \overset{\mathbb{P}}{\sim} \Gamma_{\mathsf{P}_{\delta_j,t}}\left(\mu_{\delta_j}\right),
\end{aligned}$$

where the $y_{j,t}$ are independent conditional to $\mathcal{F}_{t-1}$ and the $\delta_{j,t}$'s are independent conditionally to $(\mathcal{F}_{t-1},y_t)$.

Proposition a.3 *The dynamics of the state vector $w_t=(y_t,\delta_t)$, described by the four previous equations admits a semi-strong VAR representation. Specifically, we have:*

$$w_t=M_0+M_1w_{t-1}+\Sigma(w_{t-1})\xi_t, \tag{a.15}$$

where process $\{\xi_t\}$ is a martingale difference sequence whose covariance matrix, conditional on $\mathcal{F}_{t-1}$, is the identity matrix and where the conditional covariance matrix $\mathbb{V}ar(w_t|\mathcal{F}_{t-1})=\Sigma(w_{t-1})\Sigma(w_{t-1})'$ is of the form:

$$\left[\begin{array}{cc}
diag\left(M_2+M_3w_{t-1}\right) & diag\left(M_2+M_3w_{t-1}\right)M_4'\\
M_4 diag\left(M_2+M_3w_{t-1}\right) & M_4\,diag\left(M_2+M_3w_{t-1}\right)M_4'+diag\left(M_5+M_6w_{t-1}\right)
\end{array}\right],$$

matrices M_0, M_1, M_2, M_3, M_4, M_5 and M_6 being defined below in the proof. (If u is a n_u-dimensional vector, diag(u) denotes the diagonal matrix whose diagonal entries are the components of vector u.)

Proof Computing M_0 and M_1 amounts to computing $\mathbb{E}\left(w_t|\mathcal{F}_{t-1}\right)$:

$$\begin{aligned}
&\mathbb{E}\left(\left[\begin{array}{c} y_t \\ \delta_t \end{array}\right]\Big|\mathcal{F}_{t-1}\right) = \mathbb{E}\left(\mathbb{E}\left(\left[\begin{array}{c} y_t \\ \delta_t \end{array}\right]\Big|\mathcal{F}_{t-1}, y_t\right)|\mathcal{F}_{t-1}\right) \\
= \; & \mathbb{E}\left(\left[\begin{array}{c} y_t \\ \mu_\delta \odot \left(\alpha_\lambda + \beta_\lambda^{(y)'} y_t + \mathbf{C}'\delta_{t-1}\right) \end{array}\right]|\mathcal{F}_{t-1}\right) \\
= \; & \left[\begin{array}{c} 0 \\ \mu_\delta \odot \alpha_\lambda \end{array}\right] + \left[\begin{array}{c} \mu_y \odot \left(\nu_y + \alpha_y + \beta_y^{(y)'} y_{t-1} + \mathbf{I}'\delta_{t-1}\right) \\ \mu_\delta \odot \left(\beta_\lambda^{(y)'}\left[\mu_y \odot \left(\nu_y + \alpha_y + \beta_y^{(y)'} y_{t-1} + \mathbf{I}'\delta_{t-1}\right)\right] + \mathbf{C}'\delta_{t-1}\right) \end{array}\right] \\
= \; & \left[\begin{array}{c} \mu_y \odot (\nu_y + \alpha_y) \\ \mu_\delta \odot \alpha_\lambda + \{(\mu_\delta \mathbf{1}') \odot \beta_\lambda^{(y)'}\}\{\mu_y \odot (\nu_y + \alpha_y)\} \end{array}\right] + \\
& \left[\begin{array}{cc} (\mu_y \mathbf{1}') \odot \beta_y^{(y)'} & (\mu_y \mathbf{1}') \odot \mathbf{I}' \\ \{(\mu_\delta \mathbf{1}') \odot \beta_\lambda^{(y)'}\}\{(\mu_y \mathbf{1}') \odot \beta_y^{(y)'}\} & \{(\mu_\delta \mathbf{1}') \odot \beta_\lambda^{(y)'}\}\{(\mu_y \mathbf{1}') \odot \mathbf{I}'\} + \{(\mu_\delta \mathbf{1}') \odot \mathbf{C}'\} \end{array}\right]\left[\begin{array}{c} y_{t-1} \\ \delta_{t-1} \end{array}\right] \\
= \; & \underbrace{\left[\begin{array}{c} \mu_y \\ \mu_\delta \end{array}\right] \odot \left[\begin{array}{c} \nu_y + \alpha_y \\ \alpha_\lambda + \beta_\lambda^{(y)'}\{\mu_y \odot (\nu_y + \alpha_y)\} \end{array}\right]}_{=M_0} + \\
& \underbrace{\text{diag}\left(\left[\begin{array}{c} \mu_y \\ \mu_\delta \end{array}\right]\right)\left[\begin{array}{cc} \beta_y^{(y)'} & \mathbf{I}' \\ \beta_\lambda^{(y)'}\{(\mu_y \mathbf{1}') \odot \beta_y^{(y)'}\} & \beta_\lambda^{(y)'}\{(\mu_y \mathbf{1}') \odot \mathbf{I}'\} + \mathbf{C}' \end{array}\right]}_{=M_1}\left[\begin{array}{c} y_{t-1} \\ \delta_{t-1} \end{array}\right].
\end{aligned}$$

The computation of $\Sigma(w_{t-1})\Sigma(w_{t-1})' = \mathbb{V}ar\left(w_t|\mathcal{F}_{t-1}\right)$ is decomposed into the computation of $\mathbb{V}ar(y_t|\mathcal{F}_{t-1})$, $\mathbb{V}ar(\delta_t|\mathcal{F}_{t-1})$ and $\mathbb{C}ov(y_t, \delta_t|\mathcal{F}_{t-1})$.

Let us start with $\mathbb{V}ar(y_t|\mathcal{F}_{t-1})$. Because the $y_{i,t}$'s are conditionally independent, matrix $\mathbb{V}ar(y_t|\mathcal{F}_{t-1})$ is diagonal. Using Proposition 2.3 of Monfort, Pegoraro, Renne, and Roussellet (2017), the diagonal entries of this matrix are the components of the following vector:

$$\mu_y \odot \mu_y \odot (\nu_y + 2\alpha_y) + 2\text{diag}(\mu_y \odot \mu_y)\left(\beta_y^{(y)'} y_{t-1} + \mathbf{I}'\delta_{t-1}\right) =: M_2 + M_3 w_{t-1}.$$

In order to compute $\mathbb{V}ar(\delta_t|\mathcal{F}_{t-1})$, we use the law of total variance:

$$\mathbb{V}ar(\delta_t|\mathcal{F}_{t-1}) \;\; = \;\; \underbrace{\mathbb{V}ar(\mathbb{E}(\delta_t|y_t, \mathcal{F}_{t-1})|\mathcal{F}_{t-1})}_{=A} + \underbrace{\mathbb{E}(\mathbb{V}ar(\delta_t|y_t, \mathcal{F}_{t-1})|\mathcal{F}_{t-1})}_{=B}.$$

$$
\begin{aligned}
A &= \mathbb{V}ar(\mathbb{E}(\delta_t \big| y_t, \mathcal{F}_{t-1}) \big| \mathcal{F}_{t-1}) = \mathbb{V}ar(\{(\mu_\delta \mathbf{1}') \odot \beta_\lambda^{(y)'}\} y_t \big| \mathcal{F}_{t-1}) \\
&= \{(\mu_\delta \mathbf{1}') \odot \beta_\lambda^{(y)'}\} \mathbb{V}ar(y_t \big| \mathcal{F}_{t-1}) \{(\mu_\delta \mathbf{1}') \odot \beta_\lambda^{(y)'}\}' \\
&= \underbrace{\{(\mu_\delta \mathbf{1}') \odot \beta_\lambda^{(y)'}\}}_{=M_4} \text{diag}\left(M_2 + M_3 \begin{bmatrix} y_{t-1} \\ \delta_{t-1} \end{bmatrix} \right) \{(\mu_\delta \mathbf{1}') \odot \beta_\lambda^{(y)'}\}'.
\end{aligned}
$$

Because the $\delta_{i,t}$'s are independent conditionally to $\mathcal{F}_{t-1}$, $B = \mathbb{E}(\mathbb{V}ar(\delta_t | y_t, \mathcal{F}_{t-1}) | \mathcal{F}_{t-1})$ is a diagonal matrix whose diagonal entries are the components of

$$
\begin{aligned}
& \mathbb{E}\left(2\mu_\delta \odot \mu_\delta \odot \alpha_\lambda + 2\text{diag}(\mu_\delta \odot \mu_\delta) \left(\beta_\lambda^{(y)'} y_t + \mathbf{C}' \delta_{t-1} \right) \big| \mathcal{F}_{t-1} \right) \\
= \; & 2\mu_\delta \odot \mu_\delta \odot \alpha_\lambda + 2\text{diag}(\mu_\delta \odot \mu_\delta) \left(\beta_\lambda^{(y)'} \left\{ \mu_y \odot \left(\nu_y + \alpha_y + \beta_y^{(y)'} y_{t-1} + \mathbf{I}' \delta_{t-1} \right) \right\} + \mathbf{C}' \delta_{t-1} \right) \\
= \; & 2\mu_\delta \odot \mu_\delta \odot \alpha_\lambda + 2\text{diag}(\mu_\delta \odot \mu_\delta) \beta_\lambda^{(y)'} \{ \mu_y \odot (\nu_y + \alpha_y) \} \\
& + 2\text{diag}(\mu_\delta \odot \mu_\delta) \beta_\lambda^{(y)'} \left\{ (\mu_y \mathbf{1}') \odot \beta_y^{(y)'} \right\} y_{t-1} \\
& + 2\text{diag}(\mu_\delta \odot \mu_\delta) \left[\beta_\lambda^{(y)'} \left\{ (\mu_y \mathbf{1}') \odot \mathbf{I}' \right\} + \mathbf{C}' \right] \delta_{t-1} =: M_5 + M_6 \begin{bmatrix} y_{t-1} \\ \delta_{t-1} \end{bmatrix}.
\end{aligned}
$$

The last step consists in computing $\mathbb{C}ov(y_t, \delta_t | \mathcal{F}_{t-1})$:

$$
\begin{aligned}
\mathbb{C}ov(y_t, \delta_t \big| \mathcal{F}_{t-1}) &= \underbrace{\mathbb{E}(\mathbb{C}ov(y_t, \delta_t \big| y_t, \mathcal{F}_{t-1}) \big| \mathcal{F}_{t-1})}_{=0} + \mathbb{C}ov(y_t, \mathbb{E}(\delta_t \big| y_t, \mathcal{F}_{t-1}) \big| \mathcal{F}_{t-1}) \\
&= \mathbb{C}ov(y_t, \mathbb{E}(\delta_t \big| y_t, \mathcal{F}_{t-1}) \big| \mathcal{F}_{t-1}) \\
&= \mathbb{C}ov(y_t, (\mu_\delta \mathbf{1}') \beta_\lambda^{(y)'} y_t \big| \mathcal{F}_{t-1}) = \mathbb{V}ar(y_t \big| \mathcal{F}_{t-1}) \beta_\delta^{(y)} (\mathbf{1} \mu_\delta') \\
&= \text{diag}\left(M_2 + M_3 \begin{bmatrix} y_{t-1} \\ \delta_{t-1} \end{bmatrix} \right) \beta_\delta^{(y)} (\mathbf{1} \mu_\delta') \\
&= \text{diag}\left(M_2 + M_3 \begin{bmatrix} y_{t-1} \\ \delta_{t-1} \end{bmatrix} \right) M_4'.
\end{aligned}
$$

■

A.3 Empirical Investigation of Credit Risk Channels

To gain an intuition about the added flexibility provided by each credit risk channel in the model, we present here a calibrated example.

A.3.1 Summary of the Monte Carlo Experiment

This subsection synthetically presents the main results of a two-entity calibration and simulation exercise that has been conducted to better understand the different channels at play in our framework. Detailed results in the following sections.

We consider an economy with two defaultable entities whose credit-event intensities are driven by a single common latent factor y_t, independent from the riskless rate r_t. In the baseline case, the entities are identical and feature a null recovery rate in case of default and the SDF shows pricing associated with y_t and r_t only. We calibrate the baseline model such that the average 5y CDS is at 85bps, and the credit-risk premium goes from 0bp at the short-end to 20bps for long maturities.

We then construct three alternative parameterizations by respectively allowing for *direct contagion* from 1 to 2 ($\mathbf{C}_{2,1} > 0$), *indirect contagion* from 1 to y_t ($\mathbf{I}_1 > 0$) and *surprise pricing* of entity 2 ($\mathbf{S}_2 > 0$). Each parameter is uniquely pinned down such that the 5y CDS spread of entity 2 is at 100bps, keeping y_t at its baseline average. Comparative statics shows that the *surprise* parameter has an effect mainly at the short-end through increased risk premia, and confounds with the direct contagion effect for long enough maturities. In contrast switching on *indirect contagion* has no effect on the short-end, but increases the slope of the CDS curve with respect to the two other cases.

We simulate long time series of factors and asset prices for each parameterization. Although resulting from particular parameterizations of the model, this exercise provides some guidance on whether each case produces observational differences in observed moments. Some effects are mechanical. In particular, activating contagion effects augments default-clustering effects. The effect of *surprise* is only observed on asset prices (since it does not affect the physical dynamics of the state variable), pushing the means of short-maturity CDS spreads upwards. Second-order moments may also help distinguish between different mechanisms: the auto-correlations of simulated spreads are for instance lower for direct than for indirect *contagion*.[30]

Therefore, though it is difficult to draw general conclusions from specific calibrations, the simulation results point towards identification possibilities. To investigate this further, we conduct a Monte Carlo experiment simulating 500 trajectories of 240 months for each parameterization, discriminating whether defaults are observed or not. We estimate unrestricted versions of the model, authorizing (direct and indirect) contagion and surprise mechanisms at the same time whereas the true model only features one of these channels. Estimation is performed with approximate-filtering pseudo Maximum Likelihood (filter-based ML) and unconditional Generalized Method of Moments (GMM), to compare the precision of both methods. Our results can be summarized as follow. First, the filter-based ML method proves more efficient in recovering the correct channel in finite samples. (This justifies our utilization of filtering in our empirical exercise in the paper.) Second, including the credit-event variables δ_t as observables in the filter increases the quality of estimation, even when no defaults are observed. Last, observing in-sample defaults improves the ability of the filter to correctly identify the mechanisms at play.

We detail these results afterwards.

A.3.2 A Benchmark Economy

We consider an economy with two defaultable entities with credit event variables $\delta_t = (\delta_{1,t},\, \delta_{2,t})$ and whose probability of suffering a credit event is driven by a single common factor y_t. The historical default intensities are parameterized as:

$$\lambda_{1,t} = \beta_\lambda^{(y)} y_t, \quad \text{and} \quad \lambda_{2,t} = \beta_\lambda^{(y)} y_t + \mathbf{C} \cdot \delta_{1,t-1}\,, \tag{a.16}$$

[30] The interpretation is the following: when entity 1 defaults at $t-1$, the credit-event intensity $\lambda_{2,t}$ jumps at t with $\delta_{1,t}$ through *contagion*, whereas it jumps through the feedback loop on y_t in the *indirect contagion* case. Both parameterization make the CDSs of entity 2 jump upwards. Since y_t is persistent however, this increase in CDSs only persists in the *indirect contagion* case, not in the direct one. This decreases the autocorrelation of CDSs when direct *contagion* is authorized.

and the scale parameters $\mu_{\delta_1} = \mu_{\delta_2} = \mu_\delta$ are the same for both entities. Both components of the Poisson mixing variable are drawn independently. The common factor y_t and the risk-free rate r_t are independent and characterized by Gamma dynamics:

$$\begin{aligned} \mathsf{P}_{y,t} \,\big|\, \mathcal{F}_{t-1} &\sim \mathcal{P}\left(\beta_y^{(y)} y_{t-1} + \mathbf{I} \cdot \delta_{1,t-1}\right) \quad \text{and} \quad y_t \big| \mathsf{P}_{y,t} \sim \Gamma_{\nu_y + \mathsf{P}_{y,t}}\left(\mu_y\right), \\ \mathsf{P}_{r,t} \big| r_{t-1} &\sim \mathcal{P}\left(\alpha_r + \beta_r r_{t-1}\right) \quad \text{and} \quad r_t \big| \mathsf{P}_{r,t} \sim \gamma_{\mathsf{P}_{r,t}}\left(\mu_r\right) . \end{aligned} \tag{a.17}$$

Last, the one-period SDF is given by:

$$M_{t-1,t} = \exp\left(-r_{t-1} + \theta_r r_t + \theta_y y_t + \mathbf{S} \cdot \delta_{2,t} - \psi^{\mathbb{P}}_{w,t-1}(\theta_w)\right) . \tag{a.18}$$

In our baseline case, parameters ($\mathbf{C}$, $\mathbf{I}$, $\mathbf{S}$) are set to zero.

A.3.3 Calibration of the Illustrative Example

Our baseline model's calibration is at the monthly frequency and is presented in Table A.1. In order to avoid any discrepancies between recovery conventions, we impose that the recovery rate is zero ($\mu_\delta = 50$, the RR being defined by Equation 15). In this case, CDS spreads are virtually indistinguishable from credit spreads. We calibrate the short-rate parameters such that it has a persistence $\mu_r \cdot \beta_r$ of 0.97, a mean of 3% annualized and a standard deviation of 1% annualized. The common factor y_t is assumed to be quite persistent, with an autocorrelation of 0.95. The rest of the parameters are picked such that reasonable term structures and risk premiums are obtained.

Table A.1: Baseline Scenario Calibration

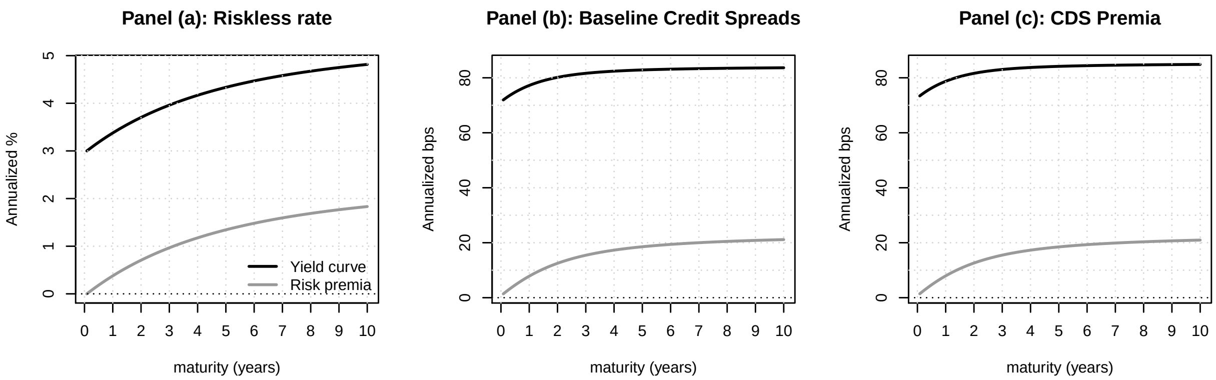
δ_t		y_t		r_t		$M_{t-1,t}$	
$\beta_\lambda^{(y)}$	$5 \cdot 10^{-4}$	$\beta_y^{(y)}$	0.95	β_r	118,172.6	θ_r	0.05
μ_δ	50	μ_y	1	μ_r	$8.21 \cdot 10^{-6}$	θ_y	0.01
$\mathbf{C}$	0	ν_y	0.06	α_r	9.1371	$\mathbf{S}$	0
		$\mathbf{I}$	0				

We present the term structures of the baseline scenario on Figure A.1. We assume that r_t and y_t are at their respective unconditional means of 3% (annualized) and 1.2 while the δ_t are fixed at 0. The riskless curve slopes up from 3% to 5% at the 10y maturity, and the term premium follows the same pattern from 0% to 2%. Panels (b) and (c) present the bond credit spreads and the CDS spreads, respectively. These two curves are virtually identical, small differences only resulting from the discrepancy between zero-coupon and par yields. In the following, we only focus on the term structure of CDSs to simplify exposition of results. The observed term structure of CDS spreads is upward sloping, from 70bps at the 1m maturity to nearly 85bps at the 10y maturity. Most of the upward sloping pattern is explained by increasing credit risk premiums, from 0bps to 20bps.

A.3.4 Contagion, Systemic Credit Risk and Credit-Event Pricing

Our first experiment consists in relaxing successively the three channels provided by our credit risk model. We thus consider hereafter calibrations where either $\mathbf{C} > 0$, $\mathbf{I} > 0$ or $\mathbf{S} > 0$. In order to

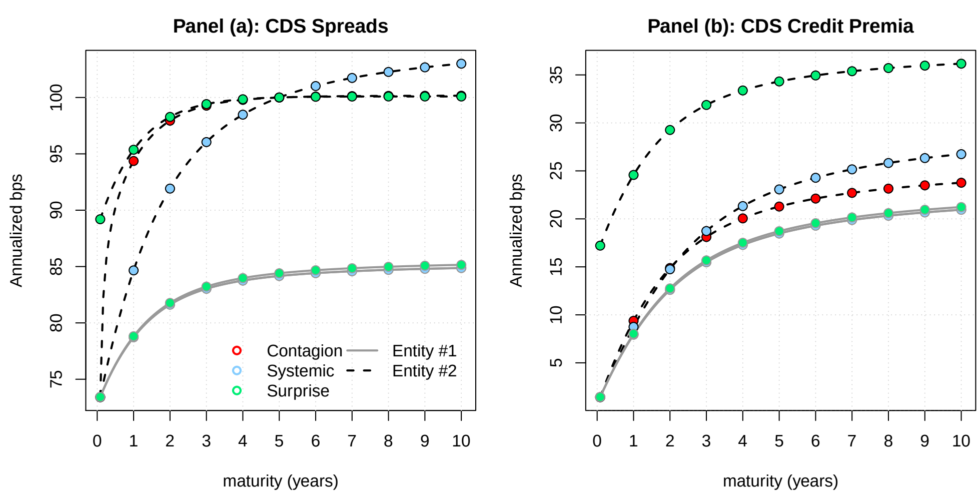
Figure A.1: Calibrated Yield Curves: Effect of Each Channel



Notes: This figure presents the term structures obtained for the baseline calibration presented in Table A.1. Panel (a) presents the yields and term premiums associated with riskless bonds, panel (b) presents the term structure of defaultable yields minus riskless yields, and panel (c) presents the CDS spreads and credit risk premiums. The term structures are obtained by assuming that r_t and y_t are at their means and that δ_t are null so no defaults have happened.

have comparable calibrations, we keep the values of r_t and y_t at the baseline unconditional mean. We pick the value of each parameter such that the 5y CDS of the entity 2 is equal to 100bps. We obtain that either $\mathbf{C} = 5.7561 \cdot 10^{-3}$, $\mathbf{I} = 0.6724$, or $\mathbf{S} = 3.5371 \cdot 10^{-3}$.

Figure A.2: Calibrated Yield Curves: Effect of Each Channel



Notes: This figure presents the term structures obtained for the alternative scenarios. Panel (a) presents the CDS spreads for each scenario, while panel (b) presents the associated credit risk premiums. The term structures are obtained by assuming that r_t and y_t are at the means implied by the baseline scenario and that δ_t are null so no defaults have happened. Term structures are presented for both entity 1 (solid grey lines) and 2 (black dashed lines). Contagion, systemic risk and credit event pricing scenarios are presented in red, blue and green, respectively.

The resulting yield curves are presented on Figure A.2. The solid grey lines present the results obtained for entity 1, which are virtually insignificant compared to the baseline. In contrast, the contagion, systemic and surprise scenarios propose three distinct term structures of CDS spreads on the second entity. First, switching on the contagion or systemic channels has a negligible effect on the very short end of the curve but creates a more upward sloping pattern than the baseline. The contagion scenario creates a curve that has more curvature and flattens out after the 5y maturity. In contrast, in the systemic risk scenario, the CDS curve has not yet plateaued at the 10y maturity. Second, both contagion and surprise scenarios have CDS term structures that are virtually the same after the 2y maturity. Third, the surprise scenario creates a large shift in the very short-end of the curve making it increase by more than 15bps at the 1m maturity. This effect is mainly operating through credit risk premiums, and the surprise scenario is the only one able to generate positive premiums at the very short-end (17bps, see Panel (b) of Figure A.2). This premiums is always at least 10bps above the credit risk premiums implied by the other scenarios.

A.3.5 Comparing Dynamics Implied by the Three Channels

We turn now to the study of the flexibility provided by each channels for the credit risk dynamics. We simulate four versions of our model, the baseline one and the three different scenarios. We simulate one trajectory of a million dates and compute associated statistics for each scenario.[31]

Table A.2 presents the obtained default probability of each entity, one-period ahead contagion probability $\mathbb{P}\left(\delta_{i,t} > 0 \middle| \delta_{j,t-1} > 0\right)$ and probability of simultaneous default, and conditional mean of the common factor y_t given that there was no default at $t-1$, that entity 1 defaulted at $t-1$, and that entity 2 has defaulted at $t-1$, respectively to measure Granger Causality. We also compute the same quantities for default events happening at t instead to measure instantaneous correlation.

Each scenario has typically the expected effect. Baseline default probabilities of each entity is 0.06%, and the contagion and simultaneous default probabilities are below 1%. With contagion, 23% of the defaults of entity 1 are followed by defaults of the second entity. The systemic risk channel increases the marginal probabilities to 0.09% because of the feedback loop, and it increases the contagion probabilities up to about 3%. The probability of simultaneous defaults also jumps up to 2.2%.[32]

As far as y_t is concerned, the contagion channel reduces slightly its average value necessary to observe a default of entity 2. The strongest effects can be observed when the systemic risk channel is switched on. Upon default of the first entity, the conditional mean of y_t jumps to more than 76 compared to 1.7 without default. Note that this also happens to a smaller extent upon default of the second entity, emphasizing that defaults tend to be more clustered in this scenario.

Up to now, our reasoning for identification of the different channels is based on the differences of dynamics before and after defaults occur. In practice, some entities will not experience any credit event in a given sample and the identification power resulting from observed asset prices can be questioned. Thus, we compare the dynamics of CDS obtained for each of these scenarios. Using the same simulated sample, we compare the conditional means and variances, autocorrelations and cross-correlations of the term structure of CDS spreads. All three scenarios unsurprisingly increase the mean and standard deviation of CDSs with respect to the baseline. The effects for the *contagion*

[31] We use the same shocks across scenarios for the simulation of $\mathsf{P}_{\delta,t}$, δ_t, $\mathsf{P}_{y,t}$, y_t, $\mathsf{P}_{r,t}$ and r_t. Since Gamma processes are conditionally heteroskedastic weak AR processes, we simulate uniform distributions and use inverse cumulative distribution functions to back out the simulations from the desired conditional distribution. Since the parameters differ across scenarios, the inverse CDFs will be different, thus creating the differences in the simulated data despite using the same uniform shocks as inputs. Any difference between scenarios are thus purely the result of difference in specifications.

[32] Note that since the only difference between the baseline and the surprise scenario lies in the SDF specification, both result in the same physical dynamics.

Table A.2: Moments of simulated factors

	Panel (a): Moments of credit event variables δ_t					
	Default Pr (%)		Contagion Pr (%)		Simultaneous Pr (%)	
	1	2	$1 \to 2$	$2 \to 1$	$1 \to 2$	$2 \to 1$
Baseline	0.06	0.06	0.82	0.64	0.99	0.96
Contagion	0.06	0.08	22.73	0.53	0.99	0.79
Systemic	0.09	0.09	2.87	1.11	2.18	2.10

	Panel (b): Moments of credit intensity factor y_t					
	No Dflt	Dflt #1	Dflt #2	No Dflt	Dflt #1	Dflt #2
	$\mathbb{E}[y_t \mid \delta_{t-1}]$			$\mathbb{E}[y_t \mid \delta_t]$		
Baseline	1.17	20.13	19.86	1.17	21.17	20.97
Contagion	1.17	20.13	19.54	1.17	21.17	20.60
Systemic	1.67	76.21	43.73	1.69	45.43	45.95

Notes: These Tables present the statistics obtained through simulations of length 1,000,000 of the baseline scenario of Table A.1 and the three scenarios. In panel (a), the first two columns present the average number of times δ_t is positive. Columns *Contagion Pr* counts the proportion of default of the one entity at t when the other has defaulted at $t-1$. Columns *Simultaneous Pr* counts the proportion of default of the one entity at t when the other has defaulted at the same time. The six columns of Panel (b) present the conditional mean of the default intensity y_t conditional on no default at $t-1$, default of entity 1 at $t-1$, default of entity 2 at $t-1$, and the same statistics for default at t, respectively.

and *surprise* scenarios are quite similar, and the average 1-year CDS spread jumps from 77bps to 92bps and 93bps respectively (see Table A.3, first four rows). In contrast, the *systemic* scenario makes the average 1-year CDS jump to 107bps. Second, the effects of the *contagion* and *surprise* scenarios are distinguishable through the auto- and cross-correlations of the CDS spreads. The baseline case produces first and twelfth order autocorrelation of 0.95 and 0.55, respectively, which drop down to 0.72 and 0.42 for the *contagion* case only. The effects are qualitatively similar across the term structure. We conclude that, in the context of this synthetic model, while the *systemic* channel conveys a bigger level impact on CDS spreads, the effects of *contagion* and *surprise* can be distinguished looking at the correlations of CDS spreads.

A.4 Monte Carlo Estimation Exercise

To get further insight on the identification power conveyed by each channel of the model, we conduct an estimation analysis on simulated trajectories. The framework is the synthetic one presented in Online Appendix A.3. Our objective is twofold. We estimate unrestricted versions of the model, authorizing contagion, systemic, and surprise channels at the same time whereas the true model

Table A.3: Moments of CDS spreads

		mean	sd	$\rho(1)$	$\rho(12)$	cor(1y)	cor(5y)	cor(10y)
1y	Baseline	76.62	233.53	0.9511	0.5471	1	0.9975	0.9915
	Contagion	91.75	280.70	0.7172	0.4157	1	0.9673	0.9553
	Systemic	106.80	417.72	0.9641	0.6475	1	0.9906	0.9654
	Surprise	92.93	283.81	0.9511	0.5472	1	0.9962	0.9877
5y	Baseline	84.21	120.16	0.9511	0.5480	0.9975	1	0.9983
	Contagion	100.32	145.04	0.8778	0.5077	0.9673	1	0.9974
	Systemic	115.57	269.14	0.9605	0.6291	0.9906	1	0.9919
	Surprise	100.41	145.32	0.9511	0.5479	0.9962	1	0.9976
10y	Baseline	85.38	76.19	0.9508	0.5473	0.9915	0.9983	1
	Contagion	101.03	92.93	0.8844	0.5106	0.9553	0.9974	1
	Systemic	114.81	197.69	0.9555	0.6044	0.9654	0.9919	1
	Surprise	101.00	93.08	0.9505	0.5465	0.9877	0.9976	1

Notes: These Tables present the statistics obtained through simulations of length 1,000,000 of the baseline scenario of Table A.1 and the three scenarios. The three blocks of rows compare the statistics for the CDS spreads of the 1y, 5y and 10y maturities. First two columns compare mean and standard deviations, the next two columns ($\rho(1)$ and $\rho(12)$) compare the first and twelfth order autocorrelation, and the remaining three columns compare the correlation with the other maturities.

only features one of the channels. This allows us to observe whether the channels are sufficiently different to be identified, even on finite samples. Estimation is performed by Maximum Likelihood (ML), where the likelihood function is computed by Kalman-filter techniques, and by unconditional GMM, allowing us to compare the precision of each method. We conduct the experiment over several samples such that some of them contain no observed defaults.

A.4.1 Framework

We assume, as is common in empirical works, that the common factor y_t is unobserved by the econometrician but δ_t is observable in real-time. She has also access to the term structures of bond credit spreads $\{\text{CS}_i(t,h)\}_{h\in H_i}$, where H_i is the discrete set of observable maturities and i refers to the defaultable entities.[33] More precisely, for any entity i and any maturity h, the bond

[33] Since the short-term rate is independent from the rest of the system, we can directly consider the bond credit spreads and forget about the riskless curve parameters. Since recovery rates are null, we focus on bond credit spreads only as information contained in the CDS curve is redundant. In a general case, despite the affine property of the model, CDS spreads are not affine in the factors (y_t, δ_t). This forces the econometrician to use a non-linear filter as

credit spreads are observed up to Gaussian white noise measurement errors independent across time, maturities and entities and with standard deviation $\sigma_\varepsilon = 1$bps. The set of parameters to be estimated is $\Theta = \{\rho_\delta, \beta_y, \nu_y, \theta_y, \mathbf{C}, \mathbf{I}, \mathbf{S}, \sigma_\varepsilon\}$, where $\rho_\delta := \mu_\delta \cdot \beta_\lambda^{(y)}$.[34] These measurement equations are accompanied with transition equations defining the VARG joint dynamics of y_t and δ_t as functions of Θ, which are detailed below in A.4.2. These equations together form the state-space model and allow us to proceed to approximate filtering maximum likelihood or moment-based estimation. These methods are described below.

A.4.2 Estimation Methods

Transition Equations: The conditional mean and the conditional variance-covariance of $w_t = (y_t, \delta_{1,t}, \delta_{2,t})'$ is given by:

$$\mathbb{E}\left(w_t|\mathcal{F}_{t-1}\right) = \begin{pmatrix} \nu_y \\ \rho_\delta \cdot \nu_y \\ \rho_\delta \cdot \nu_y \end{pmatrix} + \begin{pmatrix} \beta_y^{(y)} & \mathbf{I} & 0 \\ \rho_\delta \cdot \beta_y^{(y)} & \rho_\delta \cdot \mathbf{I} & 0 \\ \rho_\delta \cdot \beta_y^{(y)} & \rho_\delta \cdot \mathbf{I} + \mu_\delta \cdot \mathbf{C} & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \delta_{1,t-1} \\ \delta_{2,t-1} \end{pmatrix} \tag{a.19}$$

$$\text{Vech}\left[\mathbb{V}\left(w_t|\mathcal{F}_{t-1}\right)\right] = \begin{pmatrix} \nu_y \\ \rho_\delta \cdot \nu_y \\ \rho_\delta \cdot \nu_y \\ \rho_\delta \cdot \nu_y\left(2\mu_\delta + \rho_\delta\right) \\ \rho_\delta^2 \cdot \nu_y \\ \rho_\delta \cdot \nu_y\left(2\mu_\delta + \rho_\delta\right) \end{pmatrix} + \begin{pmatrix} 2\beta_y^{(y)} & 2\mathbf{I} & 0 \\ 2\rho_\delta \cdot \beta_y^{(y)} & 2\rho_\delta \cdot \mathbf{I} & 0 \\ 2\rho_\delta \cdot \beta_y^{(y)} & 2\rho_\delta \cdot \mathbf{I} & 0 \\ 2\rho_\delta \cdot \beta_y^{(y)}\left(\mu_\delta + \rho_\delta\right) & 2\rho_\delta\left(\mu_\delta + \rho_\delta\right) \cdot \mathbf{I} & 0 \\ 2\rho_\delta^2 \cdot \beta_y^{(y)} & 2\rho_\delta^2 \cdot \mathbf{I} & 0 \\ 2\rho_\delta \cdot \beta_y^{(y)}\left(\mu_\delta + \rho_\delta\right) & 2\rho_\delta\left(\mu_\delta + \rho_\delta\right) \cdot \mathbf{I} + 2\mu_\delta^2 \cdot \mathbf{C} & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \delta_{1,t-1} \\ \delta_{2,t-1} \end{pmatrix}$$

where $\rho_\delta = \mu_\delta \cdot \beta_\lambda^{(y)}$. It is easy to check that the system is second-order stationary iff $\beta_y^{(y)} < 1 - \rho_\delta \cdot \mathbf{I}$. From Equation (a.19) we obtain the semi-strong VAR representation of w_t:

$$w_t = \nu + \Phi w_{t-1} + \sqrt{\text{Vec}^{-1}\left[\Omega_0 + \Omega w_{t-1}\right]}\, \zeta_t\,, \tag{a.20}$$

where ζ_t is a standardized martingale difference, and Vec^{-1} is the operator transforming a vector into a matrix (column after column). We have:

$$\begin{aligned} &\mathbb{E}\left(w_t\right) = \left(I_3 - \Phi\right)^{-1}\nu\,, \quad \text{Vec}\left[\mathbb{V}\left(w_t\right)\right] = \left(I_9 - \Phi \otimes \Phi\right)^{-1}\left[\Omega_0 + \Omega\mathbb{E}\left(w_t\right)\right] \\ &\text{and} \quad \text{Cov}\left(w_t,\, w_{t-1}\right) = \Phi\mathbb{V}(w_t)\,. \end{aligned} \tag{a.21}$$

These formulas will be used to calculate moments of observable variables.

Filtering-based Estimation: The most standard approach for estimating term structure models with unobserved factors is based on approximate Kalman filtering (see e.g. Jong 2000). Compared with GMM-based methods, these methods allow to estimate the parameters and back out y_t at the

the extended Kalman filter. We adopt such a procedure in our real-data application in Section 4.

[34] We use $\mu_y = 1$ since this parameter is not identified.

same time. This comes at the cost of a higher computational complexity since the log-likelihood computation is performed iteratively and cannot be parallelized.

The main difficulty of the task lies in the non-linearities both in the transition and measurement equations: ζ_t is non-Gaussian and is characterized by a time-varying conditional covariance, and CDS spreads are non-linear functions of the state. A widely-employed method is the *extended Kalman filter* (EKF) which updates the filtered factors as if the data were Gaussian. This relies on two approximations, namely that *(i)* ζ_t is conditionally Gaussian, and *(ii)* CDS spreads can be dynamically approximated by a linear function of the states through a first-order Taylor expansion around their predicted values. Due to these approximations, the EKF does not provide a consistent estimator although Monte-Carlo studies show that the bias tend to be small in practice (see e.g. Duan and Simonato 1999; Monfort, Pegoraro, Renne, and Roussellet 2017).[35] Note that, in our context and for CDS spread formula, the derivatives computed with respect to y_t can be obtained analytically.

Another filtering-like approach is the so-called "inversion technique" based on Chen and Scott 1993. We describe this method more in detail below and leave it aside from our Monte-Carlo exercise for simplicity.

For all approximate filtering methods, consistency can be restored in principle by using indirect inference. However, such a refinement is likely to be heavy on the computational side, and it is unclear if restoring consistency matters from an empirical point of view. We thus also leave it aside in our Monte Carlo Experiment.

Inversion-based Estimation Inversion-based estimation methods are conceived around the idea that it is possible to recover the factors, date by date, by inverting the functions mapping the factors to the observables. Chen and Scott 1993 started with the idea that if some bonds are priced without errors, it is possible to exactly recover the values of the factors that generated them. While this is a very fast filtering method, it is subject to the arbitrary choice of which bonds to pick for exact pricing. This assumption can be relaxed by considering that certain portfolios of yields are priced without errors (see e.g. Joslin et al. 2011). In our context, a consistent approach would also require to enforce that $y_t > 0$ at all dates, which cannot be guaranteed for any model parameterization and dataset. In the general case, solving for latent factors requires numerical optimization through e.g. gradient-based methods (see also Andreasen and Christensen 2015). On key advantage with respect to filtering-based methods is that the set of optimization problems can be performed in parallel, speeding up the estimation process.

Once the time-series of y_t is obtained, the estimation consists in expressing the log-likelihood of the observables through Bayes rule. We denote by $\text{Obs}_t = \{\text{CS}_t, \text{CDS}_t, \delta_t\}$ the set of all CDS spreads, of all credit spreads, and of the credit event variables δ_t that are observable to the econometrician. We are looking for the one-period conditional log-likelihood function $\mathcal{L}\left(\text{Obs}_t|\underline{\text{Obs}_{t-1}}\right)$. We also denote by Obs_t^* the set of observables deprived of one credit spread. When the model is well-specified, there exists an invertible and deterministic function $g_t(\bullet)$ such that $\text{Obs}_t = g_t\left(\text{Obs}_t^*, y_t\right)$. The conditional quasi log-likelihood can be written in terms of both Obs_t^* and y_t:

$$\mathcal{L}\left(\text{Obs}_t|\underline{\text{Obs}_{t-1}}\right) = \mathcal{L}\left(g_t\left(\text{Obs}_t^*, y_t\right)|\underline{\text{Obs}_{t-1}}\right) = \mathcal{L}\left(\text{Obs}_t^*, y_t|\underline{\text{Obs}_{t-1}}\right) + \log\left|\frac{\partial g_t^{-1}\left(\text{Obs}_t\right)}{\partial \text{Obs}_t}\right|.$$

[35] More accurate approximations can be obtained for approximate filters. The second-order extended Kalman filter uses second-order Taylor approximation to perform the filtering recursions. The UKF uses a set of so-called "sigma-points" that are propagated through the non-linear state-space in the filtering recursions. The reader may refer to Christoffersen, Dorion, Jacobs, and Karoui (2014) for the latter.

For all dimensions but one, the function g_t is equal to identity since it transforms elements of Obs_t into itself. The last dimension is trivial when only bond credit spreads are used (because they are affine in y_t), and more complicated when adding CDSs (that are not affine in y_t). This leads the Jacobian matrix to be triangular with only one element on the diagonal different from one, and its determinant is exactly equal to that entry, denoted by $\ell_{y,t}$. Next, we can use Bayes rule to expand the conditional log-likelihood as:

$$\mathcal{L}\left(\text{Obs}_t \middle| \underline{\text{Obs}_{t-1}}\right) = \mathcal{L}\left(\text{CS}_t^*,\, \text{CDS}_t \middle| y_t,\, \delta_t\right) + \mathcal{L}\left(y_t,\, \delta_t \middle| \underline{\text{Obs}_{t-1}}\right) + \log|\ell_{y,t}|\;. \tag{a.22}$$

The first term of the log-likelihood represents the joint Gaussian distribution of the measurement errors ε_t and η_t. The second term represents the dynamics of the risk factors and can be approximated by a conditionally Gaussian log-likelihood using the transition Equations (a.19).[36]

Moments-based Estimation: One of the key advantages of writing an affine model is that both conditional and marginal moments of all factors are available analytically. This naturally opens the way for method of moments estimation. Although it would be possible to use instruments to attain the efficiency bound of the GMM estimator, we abstract from efficiency issues and directly consider marginal moments here.[37] This also has the natural advantage to avoid having to filter y_t values.

Several types of moments can be used for estimation. In particular, the conditional and marginal default probabilities are closed-form functions of the parameters in Θ:

$$\begin{aligned}
\mathbb{P}_{t-1}\left(\delta_{2,t} > 0\right) &= 1 - e^{-\frac{\beta_\lambda^{(y)}}{1+\beta_\lambda^{(y)}}\beta_y^{(y)} y_{t-1} - \left(\frac{\beta_\lambda^{(y)}}{1+\beta_\lambda^{(y)}}\mathbf{I}+\mathbf{C}\right)\delta_{1,t-1} - \nu_y \log\left(1+\beta_\lambda^{(y)}\right)} && \text{(a.23)}\\
\mathbb{P}\left(\delta_{2,t} > 0\right) &= 1 - \left[\left(1+\beta_\lambda^{(y)}\right)\prod_{i=1}^{+\infty}\left(1 + p_i + \frac{\mu_\delta q_i}{1+\mu_\delta q_i}\right)\right]^{-\nu_y}, && \text{(a.24)}
\end{aligned}$$

where the recursions for p_i and q_i are provided below. Second, the moments of bond credit spreads are those of an affine transformation of the factors and are thus attainable in closed-form, including mean, variance, and autocovariance for instance. Last, including moments of the CDS data is more challenging because of the nonlinearity of the pricing formula. One can circumvent this problem by either using a simulated method of moments (SMM) or by performing a first-order Taylor expansion of the exponential functions in the CDS pricing formula.

Recursions for Default Probabilities The recursions for the default probabilities are given by:

$$\begin{aligned}
p_n &= \frac{p_{n-1} + \mu_\delta q_{n-1}\left(\beta_\lambda^{(y)} + p_{n-1}\right)}{1 + p_{n-1} + \mu_\delta q_{n-1}\left(1 + \beta_\lambda^{(y)} + p_{n-1}\right)} \cdot \beta_y^{(y)}\\
q_n &= \frac{p_{n-1} + \mu_\delta q_{n-1}\left(\beta_\lambda^{(y)} + p_{n-1}\right)}{1 + p_{n-1} + \mu_\delta q_{n-1}\left(1 + \beta_\lambda^{(y)} + p_{n-1}\right)} \cdot \mathbf{I}\,,
\end{aligned}$$

[36]Note that it would be technically possible to use the exact likelihood for the autoregressive gamma processes, but it can only be expressed with Bessel functions whose computation involve numerically intensive methods.

[37]Optimal instrumentation can be performed by using a continuum of moments as in Carrasco, Chernov, Florens, and Ghysels (2007).

where the initial values are given by $p_1 = \frac{\beta_\lambda^{(y)}}{1+\beta_\lambda^{(y)}}\beta_y^{(y)}$ and $q_1 = \frac{\beta_\lambda^{(y)}}{1+\beta_\lambda^{(y)}}\mathbf{I} + \mathbf{C}\cdot\mathbb{1}_{\{i=2\}}$. Let us show this result by computing the default probability of the second entity.

$$\begin{aligned}
\mathbb{P}_{t-1}\left(\delta_{2,t}=0\right) &= \mathbb{P}_{t-1}\left(\mathsf{P}_{2,t}=0\right)\\
&= \mathbb{E}_{t-1}\left[\mathbb{P}_{t-1}\left(\mathsf{P}_{2,t}=0\right)\big|y_t\right]\\
&= \mathbb{E}_{t-1}\left[\exp\left(-\beta_\lambda^{(y)}y_t - \mathbf{C}\cdot\delta_{1,t-1}\right)\right]\\
&= \exp\left(-\mathbf{C}\cdot\delta_{1,t-1} - \frac{\beta_\lambda^{(y)}}{1+\beta_\lambda^{(y)}}\left(\beta_y^{(y)}y_{t-1} + \mathbf{I}\cdot\delta_{1,t-1}\right) - \nu_y\log\left(1+\beta_\lambda^{(y)}\right)\right) .
\end{aligned}$$

We can thus write:

$$\mathbb{P}_{t-1}\left(\delta_{2,t}=0\right) = \exp\left(-q_1\delta_{1,t-1} - p_1 y_{t-1} - a_1\right) ,$$

where p_1 and q_1 are given by the expressions above and $a_1 = \nu_y\log(1+\beta_\lambda^{(y)})$. Using the law of iterated expectations, we can write:

$$\mathbb{P}_{t-n}\left(\delta_{2,t}=0\right) = \exp\left(-a_1\right)\times\mathbb{E}_{t-n}\left[\exp\left(-q_1\delta_{1,t-1} - p_1 y_{t-1}\right)\right] .$$

Since the joint process w_t is affine, this expression can be transformed as:

$$\mathbb{P}_{t-n}\left(\delta_{2,t}=0\right) = \exp\left(-q_n\delta_{1,t-n} - p_n y_{t-n} - a_n\right) .$$

The recursions can be obtained by going one step further in the law of iterated expectations:

$$\begin{aligned}
\mathbb{P}_{t-n}\left(\delta_{2,t}=0\right) &= \mathbb{E}_{t-n}\left[\exp\left(-q_{n-1}\delta_{1,t+1-n} - p_{n-1}y_{t+1-n} - a_{n-1}\right)\right]\\
&= e^{-a_{n-1}}\times\mathbb{E}_{t-n}\left[\exp\left(-\frac{\mu_\delta q_{n-1}}{1+\mu_\delta q_{n-1}}\beta_\lambda^{(y)}y_{t+1-n} - p_{n-1}y_{t+1-n}\right)\right]\\
&= e^{-a_{n-1}}\times\mathbb{E}_{t-n}\left[\exp\left(-\left[p_{n-1} + \frac{\mu_\delta q_{n-1}}{1+\mu_\delta q_{n-1}}\beta_\lambda^{(y)}\right]y_{t+1-n}\right)\right]\\
&= \exp\Bigg[-a_{n-1} - \frac{p_{n-1} + \frac{\mu_\delta q_{n-1}}{1+\mu_\delta q_{n-1}}\beta_\lambda^{(y)}}{1+p_{n-1} + \frac{\mu_\delta q_{n-1}}{1+\mu_\delta q_{n-1}}\beta_\lambda^{(y)}}\left(\beta_y^{(y)}y_{t-n} + \mathbf{I}\cdot\delta_{1,t-n}\right)\\
&\quad - \nu_y\log\left(1+p_{n-1} + \frac{\mu_\delta q_{n-1}}{1+\mu_\delta q_{n-1}}\beta_\lambda^{(y)}\right)\Bigg] .
\end{aligned}$$

We simplify:

$$\frac{p_{n-1} + \frac{\mu_\delta q_{n-1}}{1+\mu_\delta q_{n-1}}\beta_\lambda^{(y)}}{1+p_{n-1} + \frac{\mu_\delta q_{n-1}}{1+\mu_\delta q_{n-1}}\beta_\lambda^{(y)}} = \frac{p_{n-1}\left(1+\mu_\delta q_{n-1}\right) + \mu_\delta q_{n-1}\beta_\lambda^{(y)}}{\left(1+p_{n-1}\right)\left(1+\mu_\delta q_{n-1}\right) + \mu_\delta q_{n-1}\beta_\lambda^{(y)}} = \frac{p_{n-1} + \mu_\delta q_{n-1}\left(p_{n-1} + \beta_\lambda^{(y)}\right)}{1+p_{n-1} + \mu_\delta q_{n-1}\left(1+p_{n-1} + \beta_\lambda^{(y)}\right)} .$$

By identification we obtain:

$$\begin{aligned}
p_n &= \frac{p_{n-1} + \mu_\delta q_{n-1}\left(\beta_\lambda^{(y)} + p_{n-1}\right)}{1+p_{n-1} + \mu_\delta q_{n-1}\left(1+\beta_\lambda^{(y)} + p_{n-1}\right)}\cdot\beta_y^{(y)}\\
q_n &= \frac{p_{n-1} + \mu_\delta q_{n-1}\left(\beta_\lambda^{(y)} + p_{n-1}\right)}{1+p_{n-1} + \mu_\delta q_{n-1}\left(1+\beta_\lambda^{(y)} + p_{n-1}\right)}\cdot\mathbf{I}\\
a_n &= a_{n-1} + \nu_y\log\left(1+p_{n-1} + \frac{\mu_\delta q_{n-1}}{1+\mu_\delta q_{n-1}}\beta_\lambda^{(y)}\right) .
\end{aligned}$$

Developing the recursion on a_n, we have:

$$a_n = \nu_y \log\left(1 + \beta_\lambda^{(y)}\right) + \nu_y \sum_{i=1}^{n-1} \log\left(1 + p_i + \frac{\mu_\delta q_i}{1 + \mu_\delta q_i} \beta_\lambda^{(y)}\right) .$$

A.4.3 Estimation Details

We simulate trajectories of length 20 years (240 periods) and obtain 500 trajectories where no default is observed and 500 where at least one default is observed. We do this for each of the four scenarios used in the comparative statics, i.e. baseline, contagion, systemic, and surprise (see Subsection A.3). For each trajectory, we estimate the set of parameters in Θ. We impose no restrictions beside all parameters being positive and the stationarity condition (see online Appendix A.4.2). For each method, we perform the same exercise using either bond credit spreads only, or bond credit spreads and credit event variables. To ensure comparability across methods, we initialize the parameter values as if all channels were switched on at the same time.

For the approximate filtering, we initialize our filter at the marginal mean and variance of the process. When δ_t is included in the set of observables, we impose no measurement errors and initialize its value at zero, with zero variance and covariance with y_t. For moment-based estimation, we operate an optimal two-step estimation where the second-step weighting matrix is adjusting for the autocorrelation of moments using Newey-West formula with 5 lags. We include the mean, variance-covariance and first order autocorrelation of the 10 credit spreads, resulting in 165 moments. When δ_t is included in the observables, we add the mean, variance-covariance and default frequency of the two credit-event processes, resulting in 7 additional moments.

A.4.4 Results

We present the estimation results for the approximate filtering in Tables A.4 and A.5, excluding and including δ_t, respectively. The GMM estimation results are provided in Tables A.6 and A.7, with a similar structure as the filtering results. The main result of the Monte-Carlo exercise is that the approximate filter is relatively more efficient in estimating the parameters and detecting which channel is switched on than our GMM-based method. We detail this result below.

Looking at the filtering results, we observe that the average and median estimated parameters are nearly always close to the true value, irrespective of the inclusion of δ_t. When there are observed defaults in the sample, the confidence bands tend to shrink, consistently with the intuition that more information provides more discriminatory power. Additionally, the filtering method is very efficient in separating the effects of each different channel. For the baseline, systemic and surprise scenarios, Table A.4 shows that the median of estimated parameters is close to zero when the channel is switched off, and close to the parameter value otherwise. The only exception is for the contagion scenario when δ_t is not included. When defaults are observed, the average parameter value of $\widehat{\mathbf{C}} \cdot 10^{-3}$ is 3.04 below the true value of 5.756 but the average value of $\widehat{\mathbf{I}}$ is slightly positive at 0.025 and the average of $\widehat{\mathbf{S}} \cdot 10^{-3}$ is 2.198 (see Table A.4). When no defaults are observed, the problem is amplified and the contagion parameter gets to virtually zero while the other two are inflated. This problem is nearly entirely corrected by adding the δ_t in the observables, which disciplines the estimation method (Table A.5). the contagion parameter $\widehat{\mathbf{C}} \cdot 10^{-3}$ now jumps to 5.2 when defaults are observed, and 3.15 when they are not. The filter still attributes somewhat of an effect to the surprise parameter (0.288 and 1.371, respectively), but the effect is largely dampened.

Including the credit-event processes in the set of observables may however have drawbacks. First, it can create numerical instability for several trajectories. For both the baseline and the surprise cases when defaults are observed, the average of the parameter ρ_δ goes to more than 18, compared to the true value of 0.025. However, this problem is likely due to only a few trajectories since the medians are exactly equal to the true value and the confidence intervals are contained, with a 95% quantile equal to 0.092 and 0.027, respectively. Second, adding defaults in the observables increases substantially the computation time needed for convergence of estimation (see Table A.8). Last, including δ_t in the observables suppresses the filtering errors on the credit event series but automatically increases the errors on the common factor y_t.

Our GMM estimation shows at least two major issues with respect to the approximate filtering method. First, irrespective of whether δ_t is included for estimation or not, the averages of parameters and confidence bands are much larger than for the approximate filter, up to very unreasonable values. When we include moments about δ_t in the estimation, the results usually get worse and some parameters explode to accommodate for the jumps on the time series. Second, For all cases, the GMM estimators are almost incapable of retrieving which channel was switched on. We conclude from this exercise that a GMM method based on marginal moments alone cannot precisely pinpoint the credit risk channels in finite samples.

Table A.4: Parameter Estimates: Approximate Filter without δ_t in the set of observable

			ν_y	β_y	ρ_δ	θ_y	$\mathbf{I}$	$\mathbf{C} \cdot 10^3$	$\mathbf{S} \cdot 10^3$	σ_ϵ
		True	0.06	0.95	0.025	0.01	0 (0.6724)	0 (5.7561)	0 (3.5371)	1
Baseline	$\delta_t = 0$	Mean (median)	0.067 (0.062)	0.95 (0.958)	0.023 (0.024)	0.009 (0.005)	0 (0)	0.002 (0)	0.002 (0)	1 (1.001)
		90% CI	[0.054 - 0.093]	[0.914 - 0.969]	[0.017 - 0.027]	[0 - 0.029]	[0 - 0.002]	[0 - 0.015]	[0 - 0.014]	[0.975 - 1.027]
	$\delta_t > 0$	Mean (median)	0.061 (0.06)	0.957 (0.963)	0.025 (0.025)	0.006 (0.003)	0 (0)	0.001 (0)	0.001 (0)	1 (1)
		90% CI	[0.057 - 0.066]	[0.93 - 0.969]	[0.022 - 0.027]	[0 - 0.02]	[0 - 0.001]	[0 - 0.007]	[0 - 0.005]	[0.973 - 1.025]
Contagion	$\delta_t = 0$	Mean (median)	0.059 (0.056)	0.944 (0.953)	0.026 (0.027)	0.013 (0.009)	0.042 (0.042)	0.009 (0.001)	3.28 (3.292)	1.003 (1.002)
		90% CI	[0.051 - 0.073]	[0.905 - 0.963]	[0.021 - 0.029]	[0.003 - 0.034]	[0.026 - 0.055]	[0 - 0.047]	[3.186 - 3.335]	[0.977 - 1.03]
	$\delta_t > 0$	Mean (median)	0.056 (0.057)	0.952 (0.955)	0.027 (0.026)	0.01 (0.007)	0.025 (0.028)	3.04 (0.561)	2.198 (2.979)	1.108 (1.006)
		90% CI	[0.051 - 0.062]	[0.937 - 0.967]	[0.024 - 0.029]	[0.002 - 0.017]	[0 - 0.047]	[0 - 5.975]	[0 - 3.32]	[0.974 - 1.068]
Systemic	$\delta_t = 0$	Mean (median)	0.064 (0.062)	0.93 (0.951)	0.023 (0.024)	0.021 (0.009)	0.675 (0.675)	0.017 (0.02)	0.085 (0.036)	1 (1)
		90% CI	[0.06 - 0.075]	[0.733 - 0.964]	[0.019 - 0.025]	[0.002 - 0.13]	[0.611 - 0.718]	[0 - 0.041]	[0 - 0.358]	[0.973 - 1.024]
	$\delta_t > 0$	Mean (median)	0.06 (0.06)	0.949 (0.956)	0.025 (0.025)	0.01 (0.007)	0.671 (0.674)	0.024 (0)	0.015 (0)	1.152 (0.999)
		90% CI	[0.06 - 0.063]	[0.928 - 0.967]	[0.024 - 0.025]	[0.001 - 0.021]	[0.658 - 0.685]	[0 - 0.035]	[0 - 0.065]	[0.973 - 1.026]
Surprise	$\delta_t = 0$	Mean (median)	0.063 (0.061)	0.946 (0.952)	0.024 (0.025)	0.012 (0.008)	0.002 (0)	0.008 (0)	3.507 (3.52)	0.999 (1)
		90% CI	[0.055 - 0.074]	[0.901 - 0.963]	[0.02 - 0.027]	[0.003 - 0.036]	[0 - 0.01]	[0 - 0.031]	[3.425 - 3.545]	[0.973 - 1.024]
	$\delta_t > 0$	Mean (median)	0.06 (0.06)	0.952 (0.955)	0.025 (0.025)	0.009 (0.007)	0.001 (0)	0.005 (0)	3.526 (3.532)	0.999 (0.999)
		90% CI	[0.057 - 0.063]	[0.934 - 0.964]	[0.024 - 0.026]	[0.003 - 0.018]	[0 - 0.006]	[0 - 0.018]	[3.49 - 3.542]	[0.972 - 1.023]

Notes: For each scenario, we simulate 1,000 trajectories of length 240, including 500 where no defaults are observed ($\delta_t = 0$) and 500 where at least one default is observed ($\delta_t > 0$). The table presents the mean and median of estimated parameters across the 500 simulations on the first row, and their 5% and 95% quantiles on the second row.

Table A.5: Parameter Estimates: Approximate Filter with δ_t in the set of observable

			ν_y	β_y	ρ_δ	θ_y	**I**	$\mathbf{C} \cdot 10^3$	$\mathbf{S} \cdot 10^3$	σ_ϵ
		True	0.06	0.95	0.025	0.01	0 (0.6724)	0 (5.7561)	0 (3.5371)	1
Baseline	$\delta_t = 0$	Mean (median)	0.064 (0.061)	0.902 (0.959)	0.024 (0.024)	0.042 (0.005)	1551.668 (0)	0.072 (0)	0.017 (0)	1.353 (1.001)
		90% CI	[0.052 - 0.086]	[0.628 - 0.969]	[0.016 - 0.028]	[0 - 0.2]	[0 - 0.002]	[0 - 0.004]	[0 - 0.02]	[0.976 - 1.028]
	$\delta_t > 0$	Mean (median)	0.059 (0.06)	0.921 (0.967)	18.472 (0.025)	0.033 (0.002)	0.025 (0)	0.301 (0)	0.189 (0)	5.052 (1.005)
		90% CI	[0.005 - 0.075]	[0.622 - 0.989]	[0.016 - 0.092]	[0 - 0.184]	[0 - 0.019]	[0 - 2.726]	[0 - 1.908]	[0.975 - 34.024]
Contagion	$\delta_t = 0$	Mean (median)	0.059 (0.059)	0.819 (0.947)	0.024 (0.025)	0.087 (0.011)	0.018 (0.017)	3.15 (3.221)	1.371 (1.272)	1 (1.001)
		90% CI	[0.049 - 0.069]	[0.523 - 0.968]	[0.018 - 0.028]	[0.001 - 0.266]	[0 - 0.045]	[0.014 - 5.668]	[0.014 - 3.229]	[0.974 - 1.027]
	$\delta_t > 0$	Mean (median)	0.062 (0.06)	0.938 (0.964)	0.269 (0.025)	0.014 (0.002)	0.02 (0.001)	5.195 (5.746)	0.288 (0.004)	1.589 (1.001)
		90% CI	[0.055 - 0.064]	[0.758 - 0.969]	[0.022 - 0.027]	[0 - 0.109]	[0 - 0.026]	[2.604 - 5.76]	[0 - 1.613]	[0.974 - 1.039]
Systemic	$\delta_t = 0$	Mean (median)	0.062 (0.061)	0.904 (0.953)	0.024 (0.025)	0.035 (0.008)	0.663 (0.671)	0.02 (0)	0.014 (0.001)	1 (1.001)
		90% CI	[0.057 - 0.069]	[0.687 - 0.969]	[0.02 - 0.026]	[0 - 0.158]	[0.573 - 0.734]	[0 - 0.116]	[0 - 0.068]	[0.974 - 1.025]
	$\delta_t > 0$	Mean (median)	0.065 (0.06)	0.947 (0.964)	8.361 (0.025)	0.013 (0.003)	0.647 (0.674)	0.388 (0)	0.264 (0)	5.223 (1.005)
		90% CI	[0.035 - 0.066]	[0.84 - 0.972]	[0.022 - 0.051]	[0 - 0.062]	[0.085 - 0.705]	[0 - 3.86]	[0 - 2.47]	[0.976 - 31.95]
Surprise	$\delta_t = 0$	Mean (median)	0.061 (0.06)	0.833 (0.948)	0.023 (0.024)	0.078 (0.011)	0.002 (0)	0.02 (0)	3.509 (3.525)	1 (1)
		90% CI	[0.053 - 0.071]	[0.573 - 0.968]	[0.018 - 0.027]	[0.001 - 0.231]	[0 - 0.009]	[0 - 0.108]	[3.404 - 3.56]	[0.974 - 1.025]
	$\delta_t > 0$	Mean (median)	0.061 (0.06)	0.937 (0.965)	18.236 (0.025)	0.015 (0.002)	0.016 (0)	0.669 (0)	3.421 (3.532)	2.139 (1.002)
		90% CI	[0.054 - 0.067]	[0.765 - 0.969]	[0.021 - 0.027]	[0 - 0.105]	[0 - 0.008]	[0 - 0.278]	[3.073 - 3.545]	[0.974 - 1.47]

Notes: For each scenario, we simulate 1,000 trajectories of length 240, including 500 where no defaults are observed ($\delta_t = 0$) and 500 where at least one default is observed ($\delta_t > 0$). The table presents the mean and median of estimated parameters across the 500 simulations on the first row, and their 5% and 95% quantiles on the second row.

Table A.6: Parameter Estimates: two-step GMM without δ_t in the set of observable

			ν_y	β_y	ρ_δ	θ_y	$\mathbf{I}$	$\mathbf{C} \cdot 10^3$	$\mathbf{S} \cdot 10^3$	σ_ϵ
		True	0.06	0.95	0.025	0.01	0 (0.6724)	0 (5.7561)	0 (3.5371)	1
Baseline	$\delta_t = 0$	Mean (median)	5.464 (0.478)	0.982 (0.99)	0.006 (0.001)	0.006 (0.002)	22.036 (0.602)	0.855 (0)	0.526 (0)	8.663 (1)
		90% CI	[0.03 - 17.851]	[0.95 - 1]	[0 - 0.025]	[0 - 0.032]	[0 - 80.312]	[0 - 5.756]	[0 - 3.537]	[0 - 25.506]
	$\delta_t > 0$	Mean (median)	0.943 (0.157)	0.98 (0.976)	0.011 (0.006)	0.001 (0)	3.236 (0.11)	0.104 (0)	0.064 (0)	3.574 (0.129)
		90% CI	[0.074 - 5.124]	[0.973 - 1]	[0 - 0.033]	[0 - 0.004]	[0 - 6.444]	[0 - 0.006]	[0 - 0.003]	[0 - 23.377]
Contagion	$\delta_t = 0$	Mean (median)	8.993 (0.298)	0.977 (0.979)	0.006 (0.002)	0.009 (0.002)	68.598 (0.672)	2.506 (2.896)	1.713 (1.65)	6.516 (0.383)
		90% CI	[0.029 - 39.255]	[0.949 - 1]	[0 - 0.025]	[0 - 0.037]	[0.001 - 546.549]	[0 - 5.756]	[0 - 3.537]	[0 - 23.244]
	$\delta_t > 0$	Mean (median)	2.991 (0.174)	0.976 (0.973)	0.01 (0.006)	0.002 (0)	34.472 (0.451)	6.304 (3.589)	0.355 (0)	5.549 (0.12)
		90% CI	[0.06 - 6.985]	[0.952 - 1]	[0 - 0.028]	[0 - 0.01]	[0 - 71.553]	[2.11 - 19.822]	[0 - 3.131]	[0 - 24.749]
Systemic	$\delta_t = 0$	Mean (median)	18.859 (0.06)	0.966 (0.955)	0.009 (0.003)	0.012 (0.01)	188.442 (0.842)	2.218 (0.744)	2.216 (2.31)	5.072 (0.934)
		90% CI	[0.029 - 139.454]	[0.945 - 1]	[0 - 0.025]	[0 - 0.04]	[0.672 - 1353.627]	[0 - 5.756]	[0 - 3.538]	[0 - 25.76]
	$\delta_t > 0$	Mean (median)	12.482 (0.232)	0.97 (0.969)	0.011 (0.009)	0.002 (0)	141.167 (1.176)	1.265 (0.63)	0.635 (0)	2.94 (0.21)
		90% CI	[0.06 - 86.309]	[0.95 - 0.998]	[0 - 0.026]	[0 - 0.01]	[0.557 - 1094.95]	[0.008 - 5.756]	[0 - 3.537]	[0 - 23.474]
Surprise	$\delta_t = 0$	Mean (median)	9.327 (0.309)	0.977 (0.983)	0.005 (0.002)	0.009 (0.002)	77.915 (0.633)	2.451 (2.964)	1.668 (1.552)	6.621 (0.355)
		90% CI	[0.027 - 37.5]	[0.947 - 1]	[0 - 0.025]	[0 - 0.039]	[0 - 516.982]	[0 - 5.756]	[0 - 3.537]	[0 - 23.064]
	$\delta_t > 0$	Mean (median)	3.649 (0.185)	0.975 (0.973)	0.017 (0.005)	0.002 (0)	39.598 (0.6)	3.085 (3.194)	0.486 (0)	2.551 (0.017)
		90% CI	[0.06 - 10.453]	[0.954 - 1]	[0 - 0.028]	[0 - 0.01]	[0.001 - 123.776]	[0.953 - 3.822]	[0 - 3.225]	[0 - 21.23]

Notes: For each scenario, we simulate 1,000 trajectories of length 240, including 500 where no defaults are observed ($\delta_t = 0$) and 500 where at least one default is observed ($\delta_t > 0$). The table presents the mean and median of estimated parameters across the 500 simulations on the first row, and their 5% and 95% quantiles on the second row. Estimation is performed using two-step GMM, where the weighting matrix in the first step is fixed at the inverse mean of each moment, and the second step weighting matrix is computed using Newey-West standard deviations with 5 lags. Since we have 165 individual moments, numerical instability can arise in which case the second-step weighting matrix is not adjusted from autocovariance.

Table A.7: Parameter Estimates: two-step GMM with δ_t in the set of observable

			ν_y	β_y	ρ_δ	θ_y	$\mathbf{I}$	$\mathbf{C} \cdot 10^3$	$\mathbf{S} \cdot 10^3$	σ_ϵ
		True	0.06	0.95	0.025	0.01	0 (0.6724)	0 (5.7561)	0 (3.5371)	1
Baseline	$\delta_t = 0$	Mean (median)	2.439 (0.604)	0.981 (0.994)	0.009 (0)	0.01 (0.003)	7.9 (0.007)	0.498 (0)	0.391 (0.027)	10.518 (9.746)
		90% CI	[0.027 - 8.573]	[0.946 - 1]	[0 - 0.026]	[0 - 0.036]	[0 - 28.731]	[0 - 5.728]	[0 - 3.547]	[0 - 24.942]
	$\delta_t > 0$	Mean (median)	232.836 (0.148)	0.937 (0.974)	2.38 (0.008)	0.013 (0.001)	6.978 (0.003)	5.065 (0)	0.539 (0.012)	8.364 (2.016)
		90% CI	[0.007 - 5.27]	[0.763 - 1]	[0 - 2.548]	[0 - 0.086]	[0 - 9.811]	[0 - 27.642]	[0 - 3.559]	[0 - 31.74]
Contagion	$\delta_t = 0$	Mean (median)	2.031 (0.305)	0.976 (0.986)	0.008 (0.001)	0.015 (0.003)	18.614 (0.259)	0.886 (0.002)	2.842 (3.142)	8.722 (1.577)
		90% CI	[0.025 - 6.355]	[0.942 - 1]	[0 - 0.036]	[0 - 0.042]	[0 - 53.117]	[0 - 3.463]	[0.611 - 4.238]	[0 - 23.472]
	$\delta_t > 0$	Mean (median)	1310.754 (0.085)	0.905 (0.963)	11.127 (0.015)	0.015 (0.002)	28.31 (0.004)	7.529 (3.541)	2.208 (2.338)	15.615 (3.78)
		90% CI	[0.003 - 6.243]	[0.478 - 1]	[0 - 35.717]	[0 - 0.077]	[0 - 38.253]	[0 - 33.785]	[0 - 4.835]	[0 - 62.978]
Systemic	$\delta_t = 0$	Mean (median)	6.158 (0.06)	0.964 (0.955)	0.013 (0.004)	0.017 (0.01)	113.15 (0.788)	1.523 (0.022)	2.648 (2.499)	8.602 (2.869)
		90% CI	[0.027 - 35.126]	[0.941 - 0.999]	[0 - 0.025]	[0 - 0.043]	[0.111 - 787.997]	[0 - 5.76]	[1.716 - 3.575]	[0 - 27.46]
	$\delta_t > 0$	Mean (median)	166.152 (0.071)	0.915 (0.958)	7.063 (0.022)	0.016 (0.004)	57.041 (0.202)	5.619 (0.056)	1.497 (1.439)	20.619 (0.859)
		90% CI	[0.009 - 15.369]	[0.717 - 0.998]	[0 - 2.009]	[0 - 0.084]	[0 - 135.381]	[0 - 29.606]	[0 - 3.572]	[0 - 87.21]
Surprise	$\delta_t = 0$	Mean (median)	1.719 (0.326)	0.976 (0.987)	0.008 (0.001)	0.014 (0.003)	14.027 (0.066)	0.921 (0)	2.84 (3.208)	8.63 (1.563)
		90% CI	[0.025 - 6.279]	[0.942 - 1]	[0 - 0.037]	[0 - 0.042]	[0 - 46.632]	[0 - 3.618]	[0.564 - 4.383]	[0 - 23.357]
	$\delta_t > 0$	Mean (median)	156.27 (0.078)	0.929 (0.961)	2.682 (0.015)	0.015 (0.003)	24.432 (0.008)	6.365 (2.108)	2.188 (3.117)	8.28 (0.909)
		90% CI	[0.007 - 8.172]	[0.648 - 1]	[0 - 7.818]	[0 - 0.079]	[0 - 50.089]	[0 - 31.839]	[0 - 4.53]	[0 - 37.27]

Notes: For each scenario, we simulate 1,000 trajectories of length 240, including 500 where no defaults are observed ($\delta_t = 0$) and 500 where at least one default is observed ($\delta_t > 0$). The table presents the mean and median of estimated parameters across the 500 simulations on the first row, and their 5% and 95% quantiles on the second row. Estimation is performed using two-step GMM, where the weighting matrix in the first step is fixed at the inverse mean of each moment and to one for the moments of δ_t, and the second step weighting matrix is computed using Newey-West standard deviations with 5 lags. Since we have more than 165 individual moments, numerical instability can arise in which case the second-step weighting matrix is not adjusted from autocovariance.

Table A.8: Computational Time and Filtering Errors: Approximate Filter

				Mean	Stdev	5% quantile	Median	95% quantile
Baseline	$\delta_t = 0$	Time (sec)	(wo/ δ_t)	143.63	120.68	53.96	110.7	326.35
			(w/ δ_t)	535.29	630.11	80.79	246.28	1985.27
		$\widehat{y}_t - y_t$	(wo/ δ_t)	0.018	0.118	-0.071	0.002	0.169
			(w/ δ_t)	-0.047	0.271	-0.499	0.001	0.189
		$\widehat{\delta}_{1,t} - \delta_{1,t}$	(wo/ δ_t)	0.027	0.084	0	0	0.159
		$\widehat{\delta}_{2,t} - \delta_{2,t}$		0.021	0.054	0	0	0.13
	$\delta_t > 0$	Time (sec)	(wo/ δ_t)	103.84	94.95	49.09	79.53	237.41
			(w/ δ_t)	334.72	475.33	54.49	164.83	1516.16
		$\widehat{y}_t - y_t$	(wo/ δ_t)	-0.01	0.113	-0.177	0.001	0.12
			(w/ δ_t)	-0.262	1.604	-2.948	0.025	1.459
		$\widehat{\delta}_{1,t} - \delta_{1,t}$	(wo/ δ_t)	-0.065	2.216	0	0.004	0.461
		$\widehat{\delta}_{2,t} - \delta_{2,t}$		-0.085	2.352	0	0.004	0.425
Contagion	$\delta_t = 0$	Time (sec)	(wo/ δ_t)	175.49	199.65	65.72	131.08	349.53
			(w/ δ_t)	843.21	768.66	78.25	435.3	2217.19
		$\widehat{y}_t - y_t$	(wo/ δ_t)	-0.057	0.186	-0.402	-0.001	0.042
			(w/ δ_t)	0.01	0.143	-0.124	0.001	0.183
		$\widehat{\delta}_{1,t} - \delta_{1,t}$	(wo/ δ_t)	0.052	0.175	0	0	0.29
		$\widehat{\delta}_{2,t} - \delta_{2,t}$		0.02	0.054	0	0	0.129
	$\delta_t > 0$	Time (sec)	(wo/ δ_t)	204.87	217.6	67.33	140.11	634.86
			(w/ δ_t)	345.95	415.41	58.52	169.2	1284.22
		$\widehat{y}_t - y_t$	(wo/ δ_t)	-0.163	0.384	-0.983	-0.004	0.029
			(w/ δ_t)	0.131	0.346	-0.201	0.095	0.715
		$\widehat{\delta}_{1,t} - \delta_{1,t}$	(wo/ δ_t)	0.134	1.589	0	0.001	0.365
		$\widehat{\delta}_{2,t} - \delta_{2,t}$		-0.089	2.987	0	0.006	0.428
Systemic	$\delta_t = 0$	Time (sec)	(wo/ δ_t)	303.18	339.04	75.56	175.94	1106.98
			(w/ δ_t)	597.7	641.31	69.36	296.12	1963.39
		$\widehat{y}_t - y_t$	(wo/ δ_t)	0.064	0.447	-0.457	0.05	0.612
			(w/ δ_t)	0.016	0.088	-0.05	0.001	0.142
		$\widehat{\delta}_{1,t} - \delta_{1,t}$	(wo/ δ_t)	0.119	0.435	0	0	0.752
		$\widehat{\delta}_{2,t} - \delta_{2,t}$		0.022	0.053	0.001	0.002	0.131
	$\delta_t > 0$	Time (sec)	(wo/ δ_t)	224.16	331.7	58.32	114.32	824.34
			(w/ δ_t)	281.62	394.55	48.72	145.86	1163.66
		$\widehat{y}_t - y_t$	(wo/ δ_t)	0.035	1.901	-2.13	0.06	1.738
			(w/ δ_t)	-0.298	3.569	-5.238	-0.103	4.618
		$\widehat{\delta}_{1,t} - \delta_{1,t}$	(wo/ δ_t)	0.237	2.206	0	0.001	2.327
		$\widehat{\delta}_{2,t} - \delta_{2,t}$		-0.083	3.132	0.001	0.018	0.797
Surprise	$\delta_t = 0$	Time (sec)	(wo/ δ_t)	177.85	184.27	67.54	129.83	387.26
			(w/ δ_t)	1016.89	757.9	212.24	602.95	2281.57
		$\widehat{y}_t - y_t$	(wo/ δ_t)	0.005	0.068	-0.063	0.001	0.085
			(w/ δ_t)	0.026	0.139	-0.068	0.002	0.226
		$\widehat{\delta}_{1,t} - \delta_{1,t}$	(wo/ δ_t)	0.051	0.182	0	0.002	0.27
		$\widehat{\delta}_{2,t} - \delta_{2,t}$		0.021	0.054	0	0	0.13
	$\delta_t > 0$	Time (sec)	(wo/ δ_t)	167.14	153.83	62.26	117.63	412.91
			(w/ δ_t)	428.1	473.33	74.17	277.57	1594.72
		$\widehat{y}_t - y_t$	(wo/ δ_t)	-0.012	0.071	-0.132	0.001	0.06
			(w/ δ_t)	0.058	0.463	-0.539	0.045	0.712
		$\widehat{\delta}_{1,t} - \delta_{1,t}$	(wo/ δ_t)	-0.021	2.311	0	0.004	0.717
		$\widehat{\delta}_{2,t} - \delta_{2,t}$		-0.085	2.352	0	0.004	0.426

Notes: In the case where δ_t is included in the measurement equations and filtered, the filtering errors are null by construction and unreported. Computations where performed in parallel on the ComputeCanada cluster where all CPUs are *Intel Platinum 8160F Skylake 2.1Ghz*.

Table A.9: Computational Time: two-step GMM

			Mean	Stdev	5% quantile	Median	95% quantile
Baseline	$\delta_t = 0$	(wo/ δ_t)	295.96	230.18	80.94	222.65	778.9
		(w/ δ_t)	204.56	151.86	73.98	140.32	518.84
	$\delta_t > 0$	(wo/ δ_t)	244.57	181.95	100.49	181.28	556.46
		(w/ δ_t)	138.16	79.24	73.42	121.21	250.22
Contagion	$\delta_t = 0$	(wo/ δ_t)	259.1	210.83	63.38	209.06	614.02
		(w/ δ_t)	191.59	142.99	78.45	143.34	532.8
	$\delta_t > 0$	(wo/ δ_t)	271.07	204.78	79.78	197.42	676.03
		(w/ δ_t)	142.26	72.48	75.78	127.19	272.14
Systemic	$\delta_t = 0$	(wo/ δ_t)	220.68	157.16	56.6	189.96	559.98
		(w/ δ_t)	170.17	76.61	78.67	163.92	330.99
	$\delta_t > 0$	(wo/ δ_t)	209.96	169.93	56.57	161.95	508.72
		(w/ δ_t)	141.67	59.51	75.76	137.55	239.16
Surprise	$\delta_t = 0$	(wo/ δ_t)	252.56	208.05	60.16	208.3	599.67
		(w/ δ_t)	192.9	148.03	79.33	142.25	534.37
	$\delta_t > 0$	(wo/ δ_t)	208.51	148.09	70.6	159.27	482.75
		(w/ δ_t)	143.26	69.59	78.12	130.85	257.65

Notes: In the case where δ_t is included in the measurement equations and filtered, the filtering errors are null by construction and unreported. Computations where performed in parallel on the ComputeCanada cluster where all CPUs are *Intel Platinum 8160F Skylake 2.1Ghz.*

A.5 Sovereign Credit Risk Application (Section 4)

A.5.1 Calibration of μ_{δ_i} in the Sovereign Credit Risk Application

Data on 1983-2015 sovereign defaults – and more specifically to the associated recovery rates – are used to calibrate μ_{δ_i}, which defines the scale of credit events $\delta_{i,t}$. Default data are from Moody's (2016). They cover 22 sovereign defaults: Russia (1998), Pakistan (1999), Ecuador (1999), Ukraine (2000), Ivory Coast (2000), Argentina (2001), Moldova (2002), Uruguay (2003), Nicaragua (2003), Grenada (2004), Dominican Republic (2005), Belize (2006), Seychelles (2008), Ecuador (2008), Jamaica (2010), Greece (2012), Greece (2012), Belize (2012), Cyprus (2013), Jamaica (2013), Argentina (2013), Ukraine (2015).

Two kinds of recovery rate estimates are considered by Moody's (2016, Exhibit 11). The first one is based on the 30-day post-default price or distressed exchange trading price. The second is the ratio of the present value of cash flows received as a result of the distressed exchange versus those initially promised, discounted using yield to maturity immediately prior to default. For each default, we compute the average of the two ratios when both are available and we take the only one that is available otherwise. Let's denote by $\overline{\varrho}_i$, $i \in 1,\dots,22$, the resulting recovery rates. Panel (a) of Figure A.3 shows an histogram of $-\log(\overline{\varrho}_i)$.

Conditional on a default at date t (i.e. $\delta_{i,t} > 0$), the distribution of $\delta_{i,t}$ is approximately a gamma distribution with a unit shape parameter and a scale parameter of μ_{δ_i}. (The approximation is accurate if the date-t probabilities of default, conditional on $(\underline{w}_{t-1}, y_t)$ are small.) Note further that, under the RMV specification used in our application, we have $\delta_{i,t} \equiv -\log(\varrho_{i,t})$. Therefore, the sample average of the $-\log(\overline{\varrho}_i)$, that is 0.6, is used as an estimate of μ_{δ_i}. The black solid line appearing on Figure A.3 shows the resulting approximate distribution of $-\log(\varrho_{i,t})$.

A.5.2 Maximum Sharpe Ratio between Dates t and $t+h$

The maximum Sharpe ratio of an investment realized between dates t to $t+h$ is given by (see Hansen and Jagannathan 1991):

$$\mathcal{M}_{t,t+h} = \frac{\sqrt{\mathbb{V}ar_t(M_{t,t+h})}}{\mathbb{E}_t(M_{t,t+h})}.$$

Using the notation $M_{t,t+1} = \exp(\mu_{0,m} + \mu_{1,m}' w_{t+1} + \mu_{2,m}' w_t)$ (where the $\mu_{i,m}$'s are for instance easily deduced from Equation 26), we have:

$$\begin{aligned} M_{t,t+h} &= \exp\left(h\mu_{0,m} + \mu_{2,m}' w_t\right) \times \\ &\quad \exp\left([\mu_{1,m} + \mu_{2,m}]' w_{t+1} + \dots + [\mu_{1,m} + \mu_{2,m}]' w_{t+h-1} + \mu_{1,m}' w_{t+h}\right) \end{aligned}$$

Therefore, using the notation $\varphi^{\mathbb{P}}_{w_t(h)}(u, v) \equiv \mathbb{E}_t(\exp(u' w_{t+1} + \dots + u' w_{t+h-1} + v' w_{t+h}))$, we get:

$$\begin{aligned} \mathcal{M}_{t,t+h} &= \frac{\sqrt{\varphi^{\mathbb{P}}_{w_t(h)}(2[\mu_{1,m} + \mu_{2,m}], 2\mu_{1,m}) - \varphi^{\mathbb{P}}_{w_t(h)}(\mu_{1,m} + \mu_{2,m}, \mu_{1,m})^2\}}}{\varphi^{\mathbb{P}}_{w_t(h)}(\mu_{1,m} + \mu_{2,m}, \mu_{1,m})} \\ &= \sqrt{\exp\{\log \varphi^{\mathbb{P}}_{w_t(h)}(2[\mu_{1,m} + \mu_{2,m}], 2\mu_{1,m}) - 2\log \varphi^{\mathbb{P}}_{w_t(h)}(\mu_{1,m} + \mu_{2,m}, \mu_{1,m})\} - 1}. \end{aligned}$$

When w_t is an affine process, $\varphi^{\mathbb{P}}_{w_t(h)}(u, v)$ is available in closed-form using recursive formulas replacing $\mathbb{Q}$ by $\mathbb{P}$ parameters).

In practice, we impose the constraint that the maximum Sharpe ratio is below 1 by infinitely penalizing the value of the filter-based likelihood for sets of parameters that does not meet this requirement. This is the approach proposed by Duffee 2010 (see his Subsection 6.5). We acknowledge that this penalization procedure can create numerical instability during the optimization procedure and renders the inference procedure non-standard.

References

Andreasen, M., and B. J. Christensen. 2015. "The SR approach: A new estimation procedure for non-linear and non-Gaussian dynamic term structure models". *Journal of Econometrics* 184 (2): 420–451.

Carrasco, M., M. Chernov, J.-P. Florens, and E. Ghysels. 2007. "Efficient estimation of general dynamic models with a continuum of moment conditions". *Journal of Econometrics* 140 (2): 529–573.

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

Christoffersen, P., C. Dorion, K. Jacobs, and L. Karoui. 2014. "Nonlinear Kalman Filtering in Affine Term Structure Models". *Management Science* 60 (9): 2248?2268.

Duan, J.-C., and J. G. Simonato. 1999. "Estimating exponential-affine term Structure models by Kalman filter". *Review of Quantitative Finance and Accounting* 13 (2): 111–135.

Duffee, G. 2010. *Sharpe Ratios in Term Structure Models.* Economics Working Paper Archive 575. The Johns Hopkins University,Department of Economics.

Hansen, L. P., and R. Jagannathan. 1991. "Implications of Security Market Data for Models of Dynamic Economies". *Journal of Political Economy* 99 (2): 225–262.

Jarrow, R. A., and S. Turnbull. 1995. "Pricing Derivatives on Financial Securities Subject to Credit Risk". *Journal of Finance* 50 (1): 53–86.

Jong, F. de. 2000. "Time Series and Cross-Section Information in Affine Term-Structure Models". *Journal of Business & Economic Statistics* 18, no. 3 (): 300–314.

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

Longstaff, F., and E. Schwartz. 1995. "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt". *Journal of Finance* 50 (3): 789–821.

Monfort, A., F. Pegoraro, J.-P. Renne, and G. Roussellet. 2017. "Staying at Zero with Affine Processes: An Application to Term Structure Modelling". *Journal of Econometrics* 201, no. 2 (): 348–366.

Moody's. 2016. *Sovereign Default and Recovery Rates, 1983-2015.* Moody's Investor Service Data Report.

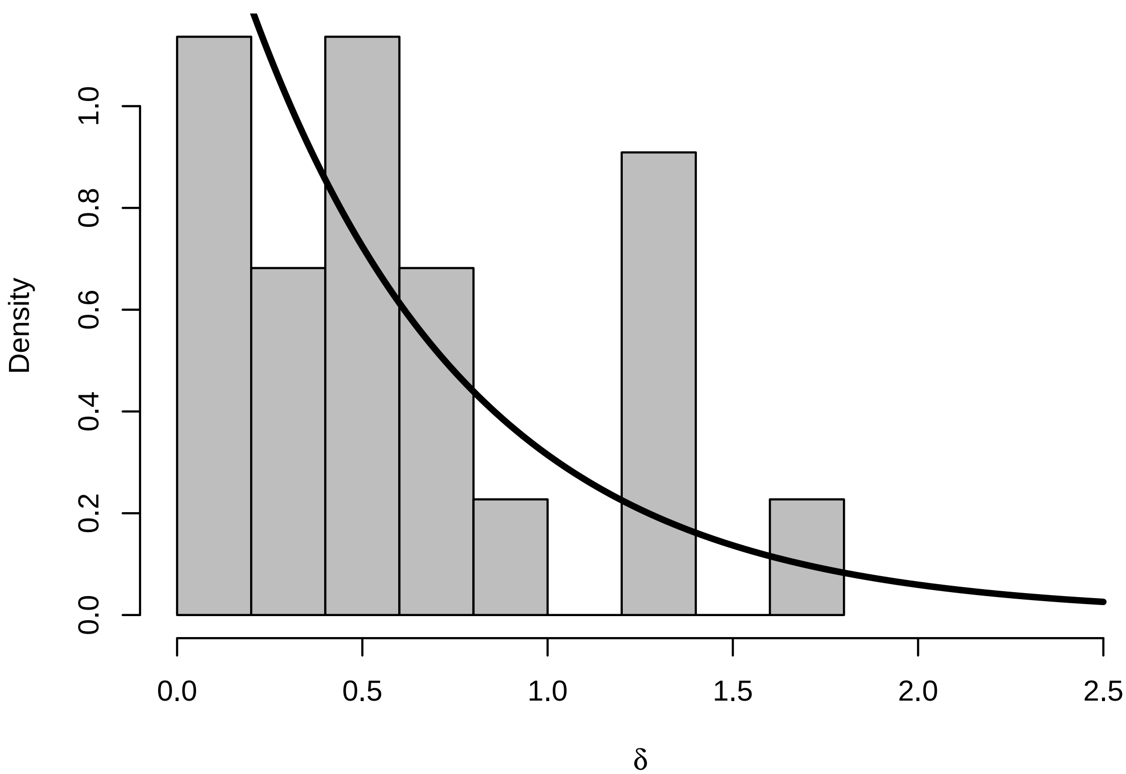
A.5.3 Tables and Figures

Table A.10: Parameter estimates

Model		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\beta^{(x)}_{\lambda,DE}$	$\times 10^5$	3.388	3.269	6.630	4.097	16.706	1.281	21.403	2.666
$\beta^{(x)}_{\lambda,FR}$	$\times 10^5$	8.281	8.038	14.579	10.299	10.758	2.746	15.374	5.300
$\beta^{(x)}_{\lambda,IT}$	$\times 10^4$	5.043	4.934	9.260	6.614	0.659	1.320	0.589	2.599
$\beta^{(x)}_{\lambda,SP}$	$\times 10^4$	2.809	2.862	4.499	3.775	1.633	0.813	1.673	1.456
$\beta^{(x)}_{\lambda,GR}$	$\times 10^2$	1.469	1.206	2.929	2.201	1.076	0.322	1.468	0.818
c_{DE}		0.000	–	0.000	–	0.040	–	0.045	–
c_{FR}		0.017	–	0.020	–	0.220	–	0.261	–
c_{IT}		0.078	–	0.133	–	2.597	–	3.485	–
c_{SP}		0.280	–	0.564	–	2.460	–	1.831	–
c_{GR}		0.016	–	0.495	–	11.838	–	47.849	–
$\boldsymbol{\kappa}_{c,DE}$		0.023	–	0.040	–	0.051	–	0.059	–
$\boldsymbol{\kappa}_{c,FR}$		0.448	–	0.002	–	0.490	–	0.514	–
$\boldsymbol{\kappa}_{c,IT}$		0.000	–	0.000	–	0.173	–	0.212	–
$\boldsymbol{\kappa}_{c,SP}$		0.522	–	0.944	–	0.279	–	0.210	–
$\boldsymbol{\kappa}_{c,GR}$		0.007	–	0.014	–	0.007	–	0.005	–
ν_z	$\times 10^2$	44.014	64.854	60.875	50.276	–	–	–	–
$1-\beta^{(z)}_z$	$\times 10^3$	26.434	25.576	26.125	38.538	–	–	–	–
ν_x	$\times 10^1$	0.060	0.005	0.018	0.000	0.061	1.381	0.067	1.544
$\beta^{(z)}_x$	$\times 10^2$	0.196	0.171	0.135	0.270	–	–	–	–
$1-\beta^{(x)}_x$	$\times 10^2$	2.511	1.822	2.163	3.556	1.603	0.712	0.394	2.906
α_r	$\times 10^2$	0.463	0.457	0.446	0.466	0.437	3.939	0.478	3.461
μ_r	$\times 10^5$	0.548	0.556	0.561	0.574	0.539	0.570	0.665	0.919
β_r	$\times 10^{-5}$	1.826	1.797	1.784	1.741	1.856	1.752	1.503	1.088
θ_z	$\times 10^3$	6.572	7.348	6.648	14.511	–	–	–	–
θ_x	$\times 10^2$	8.639	7.594	8.848	11.999	9.698	4.849	6.629	10.995
θ_r	$\times 10^{-2}$	2.154	2.181	2.234	2.136	2.281	0.248	2.083	0.279
$\mathbf{S}$		2.332	1.933	–	–	0.448	3.084	–	–
ℓ		0.109	0.000	0.055	0.000	0.353	0.106	0.588	0.000
σ_{RF}		0.291	0.291	0.291	0.291	0.287	0.295	0.291	0.292
η_{CDS}		0.152	0.155	0.150	0.156	0.171	0.259	0.166	0.257
Sharpe		1.00	1.00	0.44	1.00	0.55	1.00	0.35	0.98
log-lik.		–13768	–13847	–13819	–13917	–14270	–15136	–14246	–15213

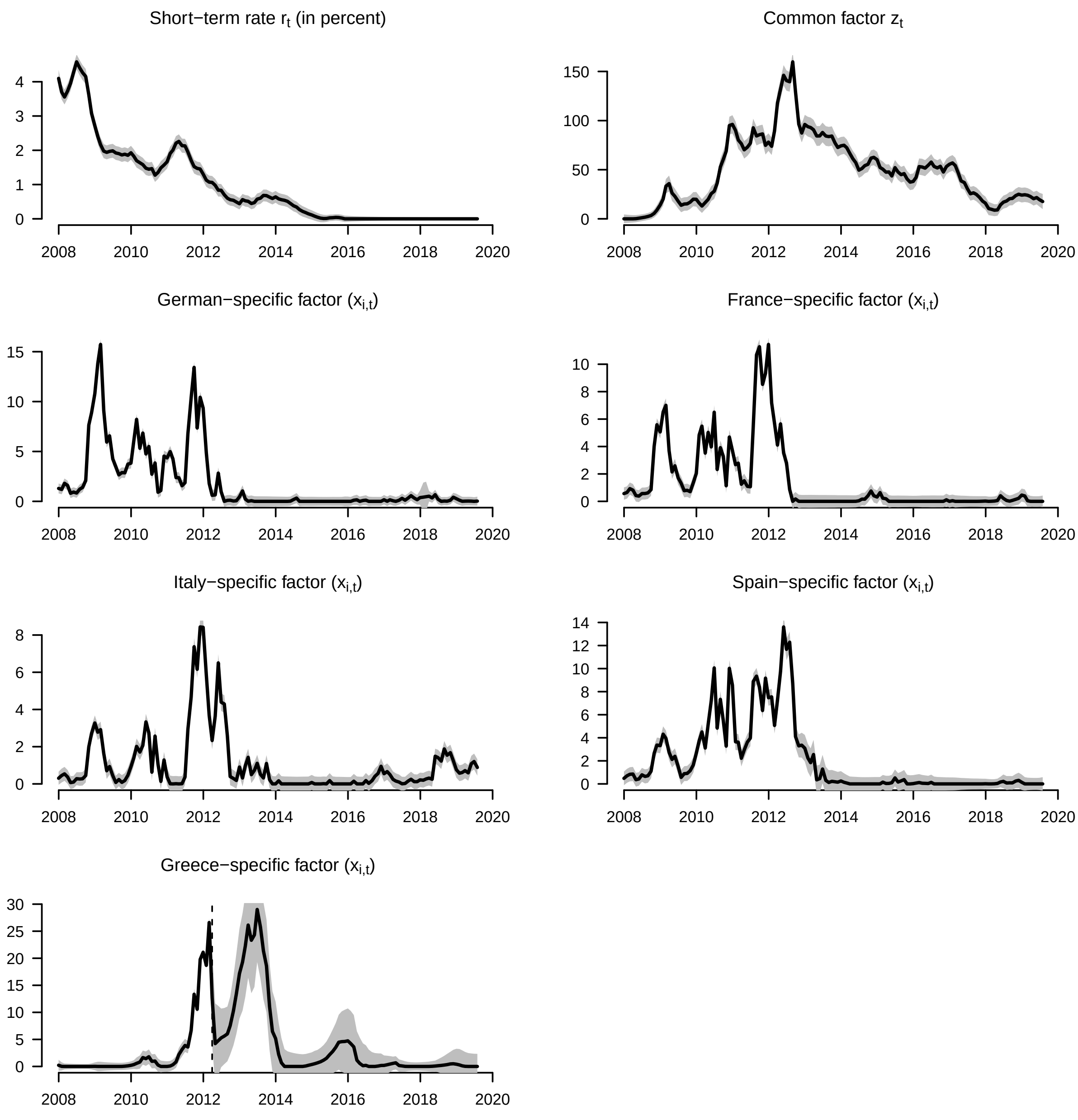
Note: Models (Eqs. 24, 25 and 26) are estimated by MLE. "$-$" indicates parameters that are constrained to be equal to zero. Model (1) is the baseline model; Models (5) to (8) feature no frailty factor (z_t) and Models (2), (4), (6) and (8) feature no contagion. σ_{RF} is the standard deviation of the measurement errors associated with risk-free zero-coupon yields, expressed in percent. The standard deviation of the measurement errors associated with a given CDS spread is equal to η_{CDS} multiplied by the sample standard deviation of the considered CDS spread. In Equation (24), $\mathbf{C}_i$ is given by $c_i\boldsymbol{\kappa}_c$, the $n = 5$ components of $\boldsymbol{\kappa}_c$ being given in the table. The components of the vector of country weights $\boldsymbol{\kappa}_M$, appearing in the SDF (Equation 26), sum to one and are proportional to countries' 2018 GDPs raised to the power of ℓ. "Sharpe" reports the sample average of the one-year maximum Sharpe ratio. The set of admissible parameters is restricted to the area resulting in an average maximum Sharpe ratios that is lower than 1.

Figure A.3: Sovereign recovery rates



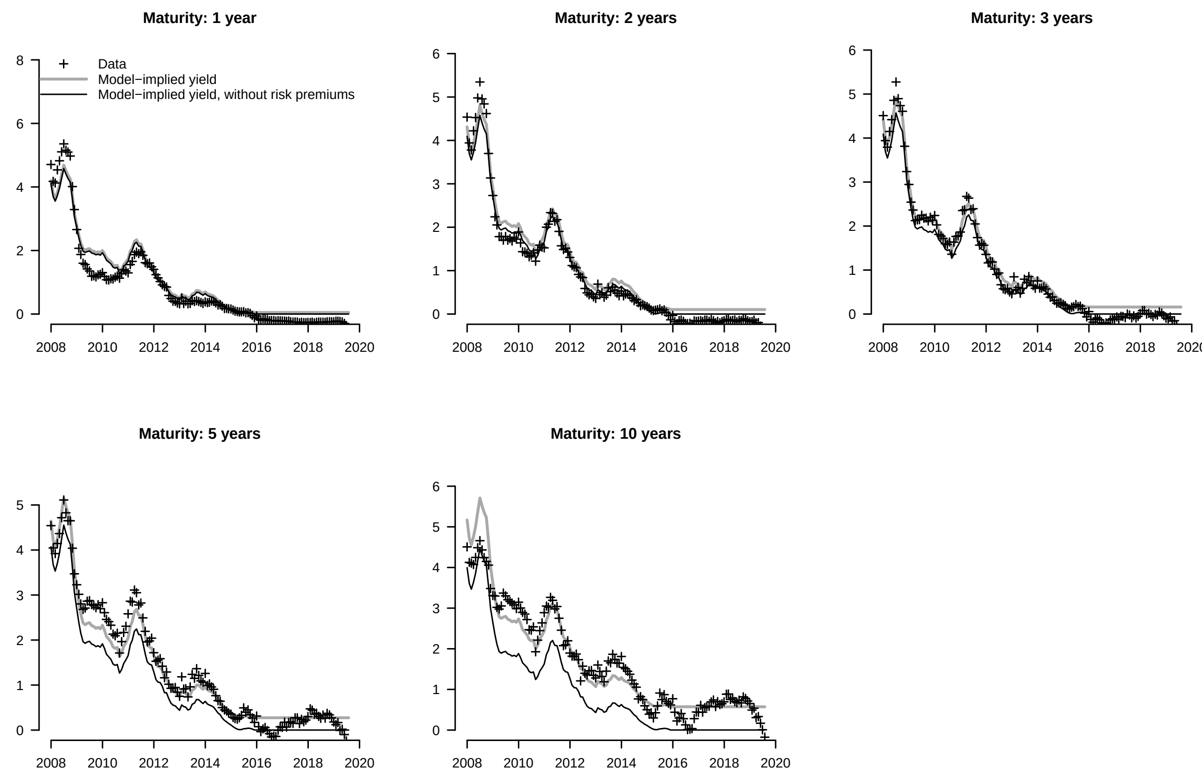
Note: This figure displays an histogram of $-\log(\overline{\varrho}_i)$, where $\overline{\varrho}_i$, $i \in 1, \ldots, 22$, are estimates of the recovery rates associated with sovereign defaults that took place over the last thirty years (Moody's 2016). In the RMV specification, $-\log(\overline{RR})$ is identical to the credit-event variable δ. The red line shows the density function of a gamma distribution with a shape parameter of 1 and a scale parameter of 0.6, which is the sample mean of $-\log(\overline{\varrho})$. In the model, this gamma distribution approximately corresponds to the distribution of $\delta_{i,t}$ conditional on default (i.e. on $\delta_{i,t} > 0$).

Figure A.4: Estimated factors



Note: This figure displays the estimated (smoothed) components of $y_t = [r_t, z_t, x_t']'$ (see Subsection 4.1 for a description of these factors). The grey areas are symmetric two-standard-deviation bands, reflecting Kalman-smoothing uncertainty. Since the Kalman smoother is approximate for non-Gaussian factors, these confidence bands are approximated as well (and can be negative); this is merely an indication of smoothing uncertainty. As regards factors z_t and x_t, the wideness of the grey bands for the first few periods results from the absence of CDS data before December 2007. For Greece: the vertical dashed bar indicates the default period (March 2012).

Figure A.5: Observed vs model-implied risk-free yields



Note: The gray lines correspond to the model-implied risk-free yields, expressed in percent. The data span the period from January 2007 to July 2019 at the monthly frequency. The thin black line corresponds to (model-implied) $\mathbb{P}$ risk-free yields, defined as the credit-risk-free yields that would be observed if agents were not risk averse (obtained by setting the prices of risk, i.e. θ_x, θ_y, θ_r and $\mathbf{S}$, to zero).