Série des Documents de Travail

# n° 2020-01

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

**Alain MONFORT**[1]
**Fulvio PEGORARO**[2]
**Jean-Paul RENNE**[3]
**Guillaume ROUSSELLET**[4]



[1] Crest, alain.monfort@ensae.fr

[2] European Central Bank (DG Economics, Monetary Analysis Division) and ACPR (Direction d'Etude et Analyse des Risques), pegoraro@ensae.fr

[3] Corresponding author. University of Lausanne, HEC, jean-paul.renne@unil.ch.

[4] Desautels Faculty of Management, McGill University, guillaume.roussellet@mcgill.ca

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

Alain MONFORT*, Fulvio PEGORARO†, Jean-Paul RENNE‡ and Guillaume ROUSSELLET§

*First version: September 2017* *This version: December 2019*

**Abstract**

We propose a discrete-time affine pricing model that simultaneously allows for (i) the presence of systemic entities by departing from the no-jump condition on the factors' conditional distribution, (ii) contagion effects, (iii) and the pricing of credit events. Our affine framework delivers explicit pricing formulas for default-sensitive securities like bonds and credit default swaps (CDS). We estimate a multi-country version of the model and address economic questions related to the pricing of sovereign credit risk. Specifically, using euro-area data, we explore the influence of allowing for the pricing of credit events, we compare frailty and contagion channels, and we extract measures of depreciation-at-default from CDS denominated in different currencies.

**JEL Codes:** E43, G12.

**Key-words:** affine credit risk model, Gamma-zero distribution, no-jump condition, contagion, credit-event risk, sovereign credit risk and exchange rates.

*CREST, alain.monfort@ensae.fr

†European Central Bank (DG Economics, Monetary Analysis Division) and ACPR (Direction d'Etude et Analyse des Risques), pegoraro@ensae.fr

‡Corresponding author. University of Lausanne, HEC, jean-paul.renne@unil.ch

§Desautels Faculty of Management, McGill University, guillaume.roussellet@mcgill.ca

The authors are thankful to Torben ANDERSEN, Christian GOURIEROUX, Damir FILIPOVIC, Jing-Zhi HUANG, Loriano MANCINI, Julien HUGONNIER, Eric JONDEAU, Emanuel MOENCH, Nour MEDDAHI, Gerardo MANZO, Gustavo SCHWENKLER, Ricardo SCHECHTMAN, Patrick AUGUSTIN, Yaroslav MELNYK and Damien ACKERER for their useful comments. We are also grateful to seminar participants at EPFL Brownbag seminar 2016, Paris-Dauphine Finance seminar 2017, and CREST Financial Econometrics Seminar 2016, and to participants to TSE Financial Econometrics Conference 2016, ESEM (Geneva) 2016, NASM (Philadelphia) 2016, GSE Summer Forum 2016 (Barcelona), MFA 2017 (Chicago), EPFL School and Workshop on Dynamic Models in Finance 2017, $10^{th}$ annual SoFiE Conference (NYU Stern) 2017, IAAE Annual Conference 2017 (Sapporo), Advances in Fixed Income and Macro-Finance Research conference 2017 (Vancouver), Central Bank of Brazil XII Annual Seminar on Risk, Financial Stability and Banking (Sao Paulo) 2017 and C.R.E.D.I.T. 2017 (Venice).

The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Central Bank or the ACPR.

# 1 Introduction

The specification of no-arbitrage asset pricing models is concerned with the formulation of empirically realistic assumptions while maintaining a large degree of tractability. This trade-off is particularly problematic in credit-risk models, which require the modeling of the joint dynamics of risk factors ($y_t$) and of entity default indicators ($d_t$), along with their interplay reflecting financial and economic linkages between entities. In the tradition of Duffie and Singleton (1999), closed-form or semi closed-form pricing formulas for defaultable securities can be obtained in an affine intensity-based framework. In this class of models the vector $y_t$ is an affine process in the risk-neutral world and both the default intensities and the risk-free short rate are affine functions of these factors.[1] However, in order to get tractable pricing formula, existing studies usually resort to one or several of the following assumptions – each of them being debated in the theoretical or empirical literature.

First, the dynamics of $y_t$ does not depend on the vector of default indicators $d_t$ (i.e. $d_t$ does not cause $y_t$ in the Granger sense). This assumption – usually referred to as the no-jump condition – is made in particular in the doubly-stochastic Cox process framework used by e.g. Jarrow and Turnbull (1995), Lando (1998) or Duffie (2005).[2] If $y_t$ contains macroeconomic variables, this condition implies that the modeled entities are not "systemic". While this is reasonable when the entities are firms of small size, it is less realistic when large banks, insurance companies, or supranational and sovereign entities are considered. The no-jump condition is for instance relaxed by Bai et al. (2015), who consider systemic firms whose credit events have economy-wide effects. Benzoni et al. (2015) show how Bayesian updating of beliefs triggered by defaults also invalidates the no-jump condition.

Second, the default probabilities of different entities are independent given the path of $y_t$, hence there are no lagged or instantaneous contagion effects. In contrast, economic and financial linkages imply a significant amount of default clustering and dynamic contagion effects (see Jarrow and Yu,

[1]See e.g. Darolles et al. (2006) for a presentation of the properties of affine processes.

[2]A discussion of the no-jump assumption can be found in e.g. Duffie et al. (1996) and Duffie and Singleton (1999). Collin-Dufresne et al. (2004) propose formulas to value defaultable claims using expected risk-adjusted discounting provided that the expectation is taken under a new modified probability measure – which is different from the risk-neutral measure – that puts zero probability on paths where default occurs prior to the maturity, and is thus only absolutely continuous with respect to the risk-neutral probability measure.

2001; Ait-Sahalia et al., 2014; Bai et al., 2015; Benzoni et al., 2015; Duan and Miao, 2016; Azizpour et al., 2018).

Third, contrary to the factors driving the default intensity, the default event of any entity is usually not priced; that is, $d_t$ is absent from the stochastic discount factor (SDF). However, the pricing of credit surprises have been shown to be an important driver of corporate bond returns and to be a possible explanation of the so-called credit spread puzzle (see Driessen, 2005, Huang and Huang, 2012, and Gourieroux et al., 2014).[3]

Fourth, the recovery payment in case of default is typically defined as a constant or predetermined fraction (recovery rate) of an exposure-at-default given by the zero-coupon bond price that would have prevailed in case of no default; it is the recovery of market value (RMV) convention of Duffie and Singleton (1999).[4] However, several studies point to the existence of stochastic recovery rates (e.g. Altman et al., 2005; Das, 2009).

Hence, these four restrictive assumptions have been invalidated, one at a time, by theoretical and empirical studies. However, the literature still lacks a framework authorizing these channels being at play simultaneously. Empirically, this brings about the question of knowing which of these channels are the most relevant to explain the term structures of credit spreads and CDSs.

This paper introduces a general discrete-time affine positive credit-risk modeling framework able to simultaneously relax the above-listed assumptions of standard frameworks while maintaining tractable pricing formulas. The asset pricing model is based on the class of Vector Autoregressive Gamma processes introduced by Monfort et al. (2017), generalizing the ARG process introduced by Gourieroux and Jasiak (2006). These non-negative processes belong to the affine class and some of their components can stay at zero for prolonged periods of time. In our model, the default event of each entity $i$ is described by the latter type of components, called credit-event variable and

[3] The credit spread puzzle corresponds to the observation that corporate and sovereign bond spreads are seemingly higher than warranted by historical default rates (see e.g. D'Amato and Remolona, 2003; Almeida and Philippon, 2007; Gabaix, 2012; Giesecke et al., 2011).

[4] Alternative modeling conventions are: the recovery of face value (RFV) convention and the recovery of Treasury (RT) convention. While the exposure-at-default is the face value of the considered bond under RFV; it is the value of the otherwise equivalent default-free bond under RT. Whatever the convention used, most existing studies consider constant recovery rates.

denoted by $\delta_{i,t}$. The default date of any entity is defined as the first date at which $\delta_{i,t}$ becomes strictly positive. The other components of the multivariate process $(y_t)$ are pricing, or risk, factors. Some components of $y_t$ can be common factors, like a short rate, a non-negative transformation of macroeconomic variable, or a "frailty" factor – usually defined as an unobservable risk factor to which firms are jointly exposed (e.g. Das et al., 2007).

Our approach accommodates contagion. Feedbacks between credit-event variables $(\delta_t)$ capture direct contagion effects, or "mutual excitation" effects (Ait-Sahalia et al., 2014). Indirect contagion (or systemic risk) can also be obtained if the credit-event variables affect some components of $y_t$ which, in turn, influence other credit-event variables.

In the model, the SDF has a standard exponential-affine formulation. Importantly, though, the SDF depends not only on the factors $y_t$, but also on the credit-event variables $\delta_t$. This formulation allows to price credit events while preserving the affine structure of our multivariate process under the associated pricing – or risk-neutral – measure, which is instrumental to obtain pricing tractability.

We close the model specification of default-sensitive securities' payoffs by assuming, for any entity, a stochastic recovery rate given by an exponential-affine function of $(y_t, \delta_t)$. The recovery payoff is defined as the product of this recovery rate and of the exposure-at-default. The three usual types of exposures-at-default are considered: recovery of market value (RMV), recovery of face value (RFV) and recovery of Treasury (RT) (see Brennan and Schwartz, 1980; Duffie, 1998; Jarrow and Turnbull, 1995; Longstaff and Schwartz, 1995; Duffie and Singleton, 1999). When defaults are rare events, the identification of the driving factors of recovery rates is a challenging task (Pan and Singleton, 2008). Our empirical investigation selects the simplest specification of the recovery rate and leaves aside a thorough investigation of the recovery rate parameterization.

We provide closed-form recursive formulas to price defaultable zero-coupon bonds and Credit Default Swaps (CDS), for any maturity. The availability of closed-form formulas hinges on the affine property of the state vector $(y_t, \delta_t)$ under the risk-neutral measure. The fact that the physical dynamics of the state vector is also affine is particularly useful when it comes to estimate the model. Indeed, in this case, the dynamics can be written in a convenient vector autoregressive form and moments are easily computed, which opens the door to standard estimation techniques (e.g. Maximum Likelihood, ML, or Generalized Method of Moments, GMM). The Kalman filter can be used when some of the components of the state vector are unobserved.

To investigate the identifiability of the credit-risk channels, we perform a Monte Carlo experiment where we simulate several calibrated versions of our framework. We compare filtering-based maximum likelihood methods with GMM and provide evidence that only the former is able to detect direct and indirect contagion, and credit-event pricing when present in the data generating processes.

We use our framework to shed new light on the pricing of euro-area sovereign credit risk. More precisely, we jointly model the fluctuations of the sovereign CDS term structures of the four largest euro-area countries – France, Germany, Italy and Spain – and of Greece over the period January 2007 to July 2019. The five credit-event intensities are driven by a short rate, a frailty factor and country-specific factors, as in Ang and Longstaff (2013). However, contrary to this latter paper, we make the SDF explicit, thereby opening the door to the computation of credit-risk premia – defined as the differences between observed CDS spreads and those that would prevail if agents were not risk-averse. Having an explicit SDF is also an important distinction with respect to the study by Ait-Sahalia et al. (2014), who also estimate affine term-structure models of euro-area sovereign CDS.[5]

We show that one common factor and one country-specific factor for each country allow for a very good fit of CDS data. The estimation detects contagion effects, even when allowing for a frailty factor. Moreover, we find sizable credit-risk premia along the whole maturity spectrum. Typically, credit-risk premia account for more than half of CDS spreads at the five year maturity for France,

[5] Although Ait-Sahalia et al. (2014)'s dataset covers seven countries, they do not estimate a joint model and focus on small models involving two countries only. Besides, because they work only under the risk-neutral dynamics, they cannot examine sovereign risk premia.

Germany, Italy and Spain.

We also show how our framework offers the possibility to study quanto CDS spreads, which are the deviations between spreads of CDS on the same entity but denominated in different currencies. The quanto CDS written on a given defaultable entity contains information about the distribution of the exchange rate at the default time of this entity, i.e. about the expected depreciation-at-default (see e.g. Ehlers and Schonbucher, 2004; Augustin et al., 2020). In order to price CDSs whose payoffs are denominated in euros or in U.S. dollars, we simply augment the model with a EURUSD exchange rate equation and allow for depreciationary effects of sovereign defaults. Let us stress that this could not be captured in standard frameworks where feedbacks from defaults to common factors are ruled out.

According to the estimated specification of the exchange rate – obtained by optimizing the fit of observed quanto CDS spreads – sovereign defaults in France, Germany, Greece, Italy and Spain would be followed by average euro depreciations of, respectively, 15%, 20%, 0%, 6% and 8%.[6] Our results further suggest that it is the fact that the exchange rate jumps upon default – and not the correlation between the exchange rate and the conditional default probability – that is key to explain the fluctuation of quanto CDS spreads.[7] These findings are consistent with the so-called "Twin Ds" phenomenon, whereby sovereign defaults are accompanied by dramatic devaluations (see Reinhart, 2002; Na et al., 2018).

The remainder of the paper is organized as follows. Section 2 presents the general affine positive credit-risk modeling framework. Section 3 provides the associated explicit pricing formulas for defaultable bonds and CDSs. Section 4 develops the sovereign credit risk and Section 5 concludes. An online appendix provides proofs, technical results and details about the calibration of our sovereign credit-risk model.

[6] Our exchange-rate-augmented model maps quanto CDS spreads to the expected impact of sovereign defaults on exchange rate, making it possible to back out these (expected) impacts even if defaults are not observed in the sample.

[7] This is in line with the findings of Ehlers and Schonbucher (2004) and of Brigo et al. (2015). The former paper is based on CDS data for Japanese multinational corporations, the latter exploits Italian sovereign CDS data.

## 2 A General Affine Positive Credit Risk Modeling Framework

In this section, we build a no-arbitrage affine credit-risk model able to provide tractable pricing formulas when any combination of the following assumptions hold: (i) the exogeneity (or no-jump) condition is not satisfied, (ii) contagion is allowed, and (iii) credit-event risk is priced.

### 2.1 Notations and Statistical Assumptions

We consider an economy with $n$ defaultable entities indexed by $i$, firms or countries for instance. Each entity is associated with an indicator of default $d_{i,t}$, such that $d_{i,t} = 1$ if $i$ is affected by a credit event at time $t$, and $d_{i,t} = 0$ otherwise. We represent the arrival of a credit event for entity $i$ through a non-negative random process denoted by $\delta_{i,t}$ (see Section 2.2).

In addition to default and credit-event variables, our economy is governed by a set of $N_y$ common factors denoted by $y_t$. For ease of presentation we do not explicitly consider entity-specific factors $x_{i,t}$ (say) even if, under proper parameterization, some of the components of $y_t$ may play that role (as will be illustrated by Section 4).[8] We also use the notations $\delta_t = (\delta_{1,t}, \dots, \delta_{n,t})$ and $w_t = (y_t,\, \delta_t)$.

At this stage, it is important to represent the information set available to investors at date $t$. We denote by $\mathcal{F}_t^*$ the collection of present and past common factors $(y_t,\, y_{t-1},\, \dots)$, while $\mathcal{D}_{i,t}$ denotes the collection of all present and past entity-$i$ credit-event variables $(\delta_{i,t}, \delta_{i,t-1}, \dots)$ and $\mathcal{D}_t = \cup_{i=1}^{n} \mathcal{D}_{i,t}$. That is, $\mathcal{D}_t$ is the entire history of all present and past credit-event variables. The entire information set available to investors is thus given by $\mathcal{F}_t = \mathcal{D}_t \cup \mathcal{F}_t^*$.

### 2.2 Default Time Modeling

**Assumption 1** *The $k^{th}$ default date $\tau_i^{(k)}$ of entity $i$ is defined as:*

$$\tau_i^{(k)} = \inf \left\{ t > \tau_i^{(k-1)} : \{\delta_{i,t} > 0\} \cap \{\delta_{i,t-1} = 0\} \right\} ,$$

[8] see Monfort, Pegoraro, Renne, and Roussellet (2018) for a general specification where $x_{i,t}$ instantaneously depend on $y_t$.

*where* $\tau_i^{(0)} = 0$ *and* $\delta_{i,t} \geqslant 0$ *a.s. and is called credit-event variable.*

This definition accommodates non-absorbing default states. As highlighted by Guo et al. (2009), the credit event affecting a firm triggers in reality a period of resolution that can possibly lead this entity to insolvency after the default date or to remain solvent (see also Kraft and Steffensen, 2007). In the latter case, $\delta_{i,t}$ comes back to zero after it has jumped to positive values. Our assumed dynamics for the credit-event variables will feature such a mechanism.[9]

In the following, we introduce the main distributional assumptions for $\delta_t$ (Assumptions 2 and 3) and $y_t$ (Assumptions 4 and 5).

**Assumption 2** *Conditionally on* $(\mathcal{F}_t^*, \mathcal{D}_{t-1})$*, each credit-event variable* $\delta_{i,t}$*,* $i \in \{1, \dots, n\}$*, is independently drawn from a Gamma-zero distribution with intensity* $\lambda_{i,t}^{\mathbb{P}}$*. More precisely, there exists a Poisson distributed mixing variable* $\mathsf{P}_{i,t}$ *such that:*

$$\left(\mathsf{P}_{i,t} \,\middle|\, \mathcal{F}_t^*, \mathcal{D}_{t-1}\right) \overset{\mathbb{P}}{\sim} \mathcal{P}\left(\lambda_{i,t}^{\mathbb{P}}\right) \qquad \textit{and} \qquad \left(\delta_{i,t} \,\middle|\, \mathsf{P}_{i,t}\right) \overset{\mathbb{P}}{\sim} \Gamma_{\mathsf{P}_{i,t}}\left(\mu_{\delta_i}\right) , \tag{1}$$

*where* $\mu_{\delta_i} > 0$ *is the scaling parameter and* $\mathsf{P}_{i,t}$ *is the degree of freedom parameter, realized at date* $t$*, of the Gamma distribution* $\Gamma_{\mathsf{P}_{i,t}}(\mu_{\delta_i})$*. The associated Gamma-zero distribution is denoted* $\Gamma_0(\lambda_{i,t}^{\mathbb{P}}, \mu_{\delta_i})$*.*

According to Equation (1), the one-period-ahead survival probability of entity $i$, given $(\mathcal{F}_t^*, \mathcal{D}_{t-1})$, is the probability that the Poisson mixing variable is equal to 0, i.e. $e^{-\lambda_{i,t}^{\mathbb{P}}}$. (Indeed, when the mixing variable is equal to zero, the Gamma distribution collapses to a Dirac mass at zero.) Hence, the conditional probability of default is approximately equal to $\lambda_{i,t}^{\mathbb{P}}$ when this variable – the physical credit-event intensity – is small. We differ the specification of $\lambda_{i,t}^{\mathbb{P}}$ to the next section.

The following proposition summarizes the properties of an autonomous Autoregressive Gamma-

[9] In the case of sovereign debts, Asonuma and Trebesch (2016) find that 62% of debt restructuring episodes observed between 1978 and 2010 occurred post-default with an average duration of five years. The remaining 38% of restructuring episodes are preemptive, that is with the restructuring implemented prior to a credit event. The associated average duration is of one year.

zero process (featuring a Gamma-zero conditional distribution); more details can be found in Monfort, Pegoraro, Renne, and Roussellet (2017), who introduce this process.

**Proposition 2.1** *Let us assume that the random process* $(\ell_t)$ *is a* $ARG_0(\alpha_\ell, \beta_\ell, \mu_\ell)$ *process of order one, that is the conditional distribution of* $\ell_{t+1}$*, given* $\underline{\ell_t} = (\ell_t, \ell_{t-1}, \ldots)$*, is the Gamma-zero distribution:*

$$(\ell_{t+1}|\underline{\ell_t}) \sim \Gamma_0(\alpha_\ell + \beta_\ell\, \ell_t, \mu_\ell) \quad \textit{for} \quad \alpha_\ell \geq 0,\, \mu_\ell > 0,\, \beta_\ell > 0\,.$$

*Then, the conditional Laplace transform* $\varphi_{\ell,t}(u\,;\,\alpha_\ell, \beta_\ell, \mu_\ell)$ *of the* $ARG_0(\alpha_\ell, \beta_\ell, \mu_\ell)$ *process is given by:*

$$\varphi_{\ell,t}(u\,;\,\alpha_\ell, \beta_\ell, \mu_\ell) \quad := \quad \mathbb{E}\left[\exp\left(u\,\ell_{t+1}\right)\,|\,\underline{\ell_t}\right] = \exp\left[\frac{u\,\mu_\ell}{1 - u\,\mu_\ell}(\alpha_\ell + \beta_\ell\,\ell_t)\right]\,, \quad \textit{for} \quad u < \frac{1}{\mu_\ell}\,.$$

Proposition 2.1 shows that the conditional Laplace transform of $\ell_{t+1}$ is exponential-affine in $\ell_t$, thus formalizing the affine nature of the process. Affine processes have various key features making them particularly useful in asset-pricing models (see e.g. Duffie, 2001; Piazzesi, 2010).

In the context of the the $\text{ARG}_0$ process considered by Proposition 2.1, the affine property is obtained by specifying the intensity – that is also the parameter of a Poisson distribution (see Equation 1) – as an affine function of $\ell_t$. In the next subsection, we consider the multivariate extension of the $\text{ARG}_0$ process. As in the univariate case, the affine nature of the state vector will be obtained by making the credit-event intensity $\lambda^{\mathbb{P}}_{i,t}$ linearly depend on the state variables.

### 2.3 Credit-Event Intensity Specification

We hereby introduce the specification of default intensities as a function of the different variables in the economy.

**Assumption 3** *For any entity* $i$*, the physical credit-event intensity is given by the following affine function of* $y_t$ *and* $\delta_{t-1}$*:*

$$\lambda^{\mathbb{P}}_{i,t} = \alpha_{\lambda_i} + \beta^{(y)'}_{\lambda_i} y_t + \mathbf{C}'_i \delta_{t-1}\,, \tag{2}$$

*where* $\alpha_{\lambda_i}$ *is a scalar,* $\beta^{(y)}_{\lambda_i}$ *is a size-*$N_y$ *vector, and* $\mathbf{C}_i$ *has* $n$ *non-negative entries, such that* $\lambda^{\mathbb{P}}_{i,t} \geqslant 0$ *a.s.*

Depending on whether $\mathbf{C}_i = \mathbf{0}$, the intensity processes are either $\mathcal{F}^*_t$- or $\mathcal{F}^*_t \cup \mathcal{D}_{t-1}$-adapted. The case where at least one component $\mathbf{C}_{i,e}$ $(i \neq e)$ is different from zero is important since it allows our framework to feature *direct contagion* between at least two entities, or mutually exciting processes. To see this, remember that a credit event happens for entity $e$ at date $t-1$ if its credit-event variable $\delta_{e,t-1}$ jumps from zero to a strictly positive value. If $\mathbf{C}_{i,e}$ is positive, the intensity $\lambda^{\mathbb{P}}_{i,t}$ increases at date $t$, thus generating a higher default probability for entity $i$. Our model can then reproduce cross-excitation, a crucial feature in the credit-risk modeling literature (see e.g. Giesecke and Zhu, 2013; Ait-Sahalia et al., 2014, 2015).[10,11]

Because of the potential persistence of $y_t$, the credit-event variable of entity $i$ can remain positive for several periods after the first jump of before going back to zero. Through contagion effects, this may increase other credit-event intensities persistently, having a long-lasting impact on the price of their defaultable securities (see following sections).

To close the model, we have to specify the dynamics of $y_t$ or, equivalently, to characterize its conditional distribution. Before proposing a specific conditional distribution (in Subsection 2.4), let us define the general context under which we get an affine state vector $(y_t, \delta_t)$ – and therefore closed-form pricing formulas:

**Assumption 4** *Given* $\mathcal{F}_{t-1}$*, the stochastic process* $\{y_t\}$ *has an exponential-affine Laplace transform:*

$$\varphi^{\mathbb{P}}_{y_{t-1}}(u_y) \quad := \quad \mathbb{E}\left[\exp(u_y' y_t) \,\middle|\, \mathcal{F}_{t-1}\right] = \exp\left[A^{(y)}_y(u_y)' y_{t-1} + A^{(\delta)}_y(u_y)' \delta_{t-1} + B_y(u_y)\right] ,$$

[10] Compared with the model of Ait-Sahalia et al. (2014), our assumption is that any jump of $\delta_{i,t}$ automatically triggers a credit event while they assume that a jump has a certain probability to generate a default (see their Equation 5). While such a mechanism could be introduced in our formulation, we leave it aside for simplicity.

[11] Our model also allows for self-excitation – in the Hawkes sens – if $\mathbf{C}_{i,i} > 0$. Note however that a situation where $\mathbf{C}_{i,i} > 0$ for some $i$'s but $\mathbf{C}_{i,e} = 0$ for $i \neq e$ is not consistent with contagion phenomena (unless the model features indirect contagion, see the discussion below Assumption 5).

*where $u_y$ is an argument of size $N_y$.*

Assumption 4 effectively breaks down the no-jump condition whenever the image of the loading function $A_y^{(\delta)}(u_y)$ contains values different from zero. In other words, the credit events of any entity $\delta_{i,t}$ can have an impact on the common factors dynamics through Granger-causality. In this case, process $y_t$ is not autonomous. We call this mechanism indirect contagion (or systemic risk) since the default of a single entity can for instance have an impact on one of the components of $y_t$ representing the state of the economy, which will in turn feedback onto higher credit-event intensities.

Combining Assumptions 2, 3 and 4, we show that $w_t$ is an affine process:

**Proposition 2.2** *Under Assumptions 2, 3 and 4, the stochastic process $\{w_t\}$ is affine under the historical probability measure $\mathbb{P}$. That is, the conditional Laplace transform of $w_t$ given $\mathcal{F}_{t-1}$ is an exponential-affine function of $w_{t-1}$. Formally:*

$$\varphi^{\mathbb{P}}_{w_{t-1}}(u_w) \quad := \quad \mathbb{E}\left[\exp(u_y' y_t + u_\delta' \delta_t) \,\middle|\, \mathcal{F}_{t-1}\right] = \exp\left[A_w^{(y)}(u_w)' y_{t-1} + A_w^{(\delta)}(u_w)' \delta_{t-1} + B_w(u_w)\right] ,$$

*where $u_w = (u_y, u_\delta)$, and the functions $A_w^{(y)}$, $A_w^{(\delta)}$ and $B_w$ are given by:*

$$\begin{aligned}
A_w^{(y)}(u_w) &= A_y^{(y)}\left(u_y + \beta_\lambda^{(y)} \frac{u_\delta \odot \mu_\delta}{\mathbf{1} - u_\delta \odot \mu_\delta}\right) , \; A_w^{(\delta)}(u_w) = A_y^{(\delta)}\left(u_y + \beta_\lambda^{(y)} \frac{u_\delta \odot \mu_\delta}{\mathbf{1} - u_\delta \odot \mu_\delta}\right) + \mathbf{C}\,\frac{u_\delta \odot \mu_\delta}{\mathbf{1} - u_\delta \odot \mu_\delta} , \\
B_w(u_w) &= B_y\left(u_y + \beta_\lambda^{(y)} \frac{u_\delta \odot \mu_\delta}{\mathbf{1} - u_\delta \odot \mu_\delta}\right) + \alpha_\lambda' \frac{u_\delta \odot \mu_\delta}{\mathbf{1} - u_\delta \odot \mu_\delta} ,
\end{aligned}$$

*where $\beta_\lambda^{(y)}$ is the $(N_y \times n)$ matrix whose columns are $\beta_{\lambda_i}^{(y)}$, $\mathbf{C}$ is the $(n \times n)$ matrix whose columns are $\mathbf{C}_i$, and $\alpha_\lambda$ is the vector of all $\alpha_{\lambda_i}$. The operator $\odot$ is the (Hadamard) element-by-element product and the $\div$ operator is taken element-by-element.*

**Proof** See Online Appendix A.1.1.

### 2.4 Proposed Affine Vector Autoregressive Gamma Dynamics

Assumption 4 leaves some freedom about the specification of the dynamics of $y_t$. Assumption 5 describes a particularly convenient choice that we develop further in the empirical application (see Chen and Filipovic, 2007, for a continuous-time approach).[12]

**Assumption 5** *The common factors $y_{j,t}$, $j \in \{1, \ldots, N_y\}$, feature the following non-central Gamma dynamics :*

$$\left(\mathsf{P}_{y_j,t} \,\middle|\, \mathcal{F}_{t-1}\right) \overset{\mathbb{P}}{\sim} \mathcal{P}\left(\beta_{y_j}^{(y)\prime} y_{t-1} + \mathbf{I}_j' \delta_{t-1}\right) \qquad \textit{and} \qquad \left(y_{j,t} \,\middle|\, \mathsf{P}_{y_j,t}\right) \overset{\mathbb{P}}{\sim} \Gamma_{\nu_{y_j} + \mathsf{P}_{y_j,t}}\left(\mu_{y_j}\right) . \tag{3}$$

*All parameters of Equation (3) are non-negative and of adapted dimension. Moreover, conditionally on $\mathcal{F}_{t-1}$, the scalar components of $y_t$ are assumed to be independent.*

Contrary to the credit-event variables $\delta_t$, the risk factors $y_{j,t}$ cannot stay at zero when parameters $\nu_{y_j}$ are strictly positive. Two relevant characteristics stand out from this specification. First, the vectors of parameters $\mathbf{I}_j$ represent the transmission channel of what we have called systemic risk in the previous subsection. If entity $e$ defaults at $t-1$, its credit-event variable $\delta_{e,t-1}$ jumps to a positive value, which increases the conditional mean of the common factor $y_{j,t}$ as long as Granger-causality is allowed by $\mathbf{I}_{j,e} > 0$. Parameters $\mathbf{I}_{j,e} > 0$ *and* $\beta_{\lambda_i}^{(y)}$ then open the way to an *indirect contagion* channel from the systemic entity $e$ (featuring $\delta_{e,t-1} > 0$) to another entity $i$ via the common factor $y_{j,t}$. (The chosen notation $\mathbf{I}_j$ makes reference to the indirect nature of the contagion channel; in the same way, $\mathbf{C}_i$ was referring to direct contagion.) This feature results from the recursive specification of the two sets of variables: while Equation (2) implies that $y_t$ instantaneously causes $\delta_t$, Equation (3) shows that $y_t$ depends on $\delta_{t-1}$.

Under Assumption 5, risk factors $y_t$ are non-negative. This may represent a problem if one wants to include observable risk factors with negative values among $y_t$'s components. In some cases,

[12] Alternative admissible dynamics have to be such that the credit-event intensities stay positive. One such example can be found in Roussellet (2019), who uses quadratic combinations of Gaussian processes.

this problem may be circumvented by employing non-negative transformations of the considered variables.

The state process $\{w_t\} = \{(y_t,\, \delta_t)\}$ resulting from Assumptions 2, 3 and 5 is called recursive Vector Autoregressive Gamma (VARG) process.[13] The following proposition complements Proposition 2.2 in the VARG context. Specifically, it gives the forms of functions functions $A_y^{(y)}$, $A_y^{(\delta)}$ and $B_y$ resulting from Assumption 5.

**Proposition 2.3** *Under Assumptions 2, 3 and 5, the Laplace transform of $w_t$, conditional on $\mathcal{F}_{t-1}$, is exponential-affine in $w_{t-1}$ and given by Proposition 2.2, with functions $A_y^{(y)}$, $A_y^{(\delta)}$ and $B_y$ given by:*

$$A_y^{(y)}(u_y) \;\; = \;\; \beta_y^{(y)} \frac{u_y \odot \mu_y}{\mathbf{1} - u_y \odot \mu_y} \, , \;\; A_y^{(\delta)}(u_y) = \mathbf{I} \, \frac{u_y \odot \mu_y}{\mathbf{1} - u_y \odot \mu_y} \, , \;\; B_y(u_y) = -\nu_y' \log \left[\mathbf{1} - u_y \odot \mu_y\right] \, ,$$

*where $\beta_y^{(y)}$ is the $(N_y \times N_y)$ matrix whose columns are $\beta_{y_j}^{(y)}$, $\mathbf{I}$ is the $(n \times N_y)$ matrix whose columns are $\mathbf{I}_j$, and $\mu_y$ and $\nu_y$ are the vectors stacking together the individual elements with the same notations. The operator $\odot$ is the Hadamard element-by-element product, and the $\log(\cdot)$ and $\div$ operators are taken element-by-element.*

**Proof** See Online Appendix A.1.2.

As for standard affine processes, our assumed dynamics for $w_t$ has convenient properties in terms of conditional cumulants, stationarity conditions and predictions. These properties directly derive from the semi-strong VAR representation of the state-vector dynamics (see Online Appendix A.2).[14]

[13]The Vector Autoregressive Gamma (VARG) process is a multivariate generalization of the ARG process. The notation is introduced in Monfort et al. (2017) and considers conditionally independent components, as is the case here (see also Dai et al., 2010, and Creal and Wu, 2015, for alternative approaches). The generalization to the case of conditional dependence is proposed by Monfort et al. (2018).

[14]In particular, once the semi-strong VAR representation is known, the state vector is strictly stationary iff the eigenvalues of the auto-regressive matrix (denoted by $M_1$ in Online Appendix A.2) are strictly lower than one in modulus.

### 2.5 The Stochastic Discount Factor and Credit-Risk Pricing

We assume the existence of a representative investor who prices all assets in the economy such that no-arbitrage holds.[15] We focus on the following financial instruments: risk-free debt, debt issued by all defaultable entities $i$, and credit-risk derivatives on all entities.

The following assumption pertains to the short-term risk-free rate and the SDF.

**Assumption 6** *The risk-free one-period yield between $t-1$ and $t$, denoted by $r_{t-1}$, is given by an affine function of the factors:*

$$r_{t-1} = \xi_0 + \xi_y' y_{t-1} + \xi_\delta' \delta_{t-1}\,, \tag{4}$$

*where $\xi_0$, $\xi_y$, and $\xi_\delta$ are respectively a scalar, a $N_y$-dimensional vector and a $n$-dimensional vector. The one-period SDF is denoted $M_{t-1,t}$ and given by:*

$$M_{t-1,t} = \exp\left(-r_{t-1} + \theta_y' y_t + \mathbf{S}'\delta_t - \log\left[\varphi^{\mathbb{P}}_{w_{t-1}}(\theta_y, \mathbf{S})\right]\right)\,, \tag{5}$$

*where $(\theta_y, \mathbf{S})$ are the risk-correction parameters, or "prices of risk."*

Although our SDF formulation is primarily motivated by computational reasons, its exponential-affine formulation can stem from structural models where agents feature CRRA or Epstein-Zin preferences and where consumption growth is affine in $w_t$.[16] In particular, these structural approaches imply that $\mathbf{S}_i > 0$ if the credit event associated with entity $i$ coincide with a drop in consumption, which is for instance consistent with empirical evidence on the effect of sovereign defaults (see e.g. Reinhart and Rogoff, 2011; Mendoza and Yue, 2012; Trebesch and Zabel, 2017).[17]

[15] In the discrete-time context, it can be shown that under the assumptions of (a) existence of uniqueness of a price, (b) price linearity and continuity and (c) absence of arbitrage opportunity, there exists a unique positive SDF. This derives from a conditional version of the Riesz representation theorem (see e.g. Hansen and Richard, 1987).

[16] If the representative agent features time-separable CRRA preferences and if consumption growth between dates $t-1$ and $t$ is denoted by $\Delta c_t$, then it is easily shown that the SDF $M_{t-1,t}$ is proportional to $\exp(-\gamma\Delta c_t)$, where $\gamma$ is the relative risk aversion parameter. Eraker (2008) proposes an approach to solve for an approximated exponential affine SDF when the representative agent features Epstein and Zin (1989)'s preferences and when $\Delta c_t$ linearly depends on an affine process. Bai et al. (2015) also obtain a pricing kernel that depends, in an exponential affine way, on credit-event variables in the context of a production economy populated by firms (see their Equations 25).

[17] The fact that consumption drops coincide with sovereign defaults is also consistent with the disaster-risk literature,

By construction, the SDF specification of Equation (5) is consistent with the basic no-arbitrage relationship that $\mathbb{E}\left[M_{t-1,t}\middle|\mathcal{F}_{t-1}\right] = e^{-r_{t-1}}$. The price of risk $\theta_y$, which is standard in the context of credit-risk modeling, drives a wedge between the physical and risk-neutral moments of the factors driving the credit-event intensities. The vector of parameters $\mathbf{S}$ allows us to relax the widely-used assumption according to which credit events do not enter the representative investor's SDF – and are therefore not a priced source of risk. In line with Gourieroux et al. (2014), we call this mechanism credit-event pricing (or surprise pricing, hence the notation $\mathbf{S}$).

Now that the physical dynamics and the SDF are known, we can price assets whose payoffs depend on future values of the state vector. Specifically, the date-$t$ price of an asset providing the payoff $P_{t+1}$ on date $t+1$ is given by $\mathbb{E}(M_{t,t+1}P_{t+1}|\mathcal{F}_t)$. This price can also be written $\mathbb{E}^{\mathbb{Q}}(e^{-r_t}\,P_{t+1}|\mathcal{F}_t)$, where the one-period-ahead change of measure from $\mathbb{P}$ (the physical measure) to $\mathbb{Q}$ (the risk-neutral measure) is given by $M_{t,t+1}/\mathbb{E}(M_{t,t+1}|\mathcal{F}_t)$. Determining the risk-neutral measure often facilitates pricing. In the remaining of this subsection, we explore the risk-neutral dynamics resulting from our assumptions.

To start with, the following proposition states that the risk-neutral conditional Laplace transform of $w_t$ – which characterizes its risk-neutral dynamics – is exponential affine and is readily available from its physical conditional Laplace transforms (given in Proposition 2.2) and from the prices of risk (defined in Assumption 6).

**Proposition 2.4** *Under Assumptions 2, 3, 4 (or 5) and 6, the stochastic process $\{w_t\}$ has an exponential-affine conditional Laplace transform given $\mathcal{F}_{t-1}$ under the risk-neutral measure:*

$$\varphi^{\mathbb{Q}}_{w_{t-1}}(u_w) \quad := \quad \mathbb{E}^{\mathbb{Q}}\left[\exp(u_y'y_t + u_\delta'\delta_t)\,\middle|\,\mathcal{F}_{t-1}\right] = \exp\left[A_w^{\mathbb{Q}(y)}(u_w)'y_{t-1} + A_w^{\mathbb{Q}(\delta)}(u_w)'\delta_{t-1} + B_w^{\mathbb{Q}}(u_w)\right],$$

where sovereign default can be triggered by exogenous disasters having dramatic recessionary effects (see e.g. Barro, 2006; Gabaix, 2012).

*where the loadings functions are given by:*

$$\begin{aligned} A_w^{\mathbb{Q}^{(\ell)}}(u_w) &= A_w^{(\ell)}\big(u_y + \theta_y,\, u_\delta + \mathbf{S}\big) - A_w^{(\ell)}\big(\theta_y,\, \mathbf{S}\big)\,,\ \ell \in \{y, \delta\}\,, \\ B_w^{\mathbb{Q}}(u_w) &= B_w\big(u_y + \theta_y,\, u_\delta + \mathbf{S}\big) - B_w\big(\theta_y,\, \mathbf{S}\big)\,, \end{aligned}$$

*where $A_w^{(\ell)}(u_w)$, $\ell = \{y, \delta\}$, and $B_w(u_w)$ are defined in Proposition 2.2.*

**Proof** Straightforward application of the Esscher transform. ■

Proposition 2.4 underlines the wide use of affine processes in asset-pricing models: the exponential-affine specification of the SDF preserves the affine property when we move from the historical to the risk-neutral measure. This leads to closed-form pricing formulas not only for securities paying off an exponential-affine function of the risk-factors $w_t$, but also featuring payoffs of the form $\mathbb{1}_{\{\delta_{i,t}>0\}}$ (see Section 3).

While Proposition 2.4 perfectly characterizes the risk-neutral dynamics of $w_t$, it does not make the risk-neutral conditional distributions of $\delta_t$ and of $y_t$ explicit. Remarkably, the risk-neutral conditional distributions of $\delta_t$ and of $y_t$ remain of the same type as their physical counterpart. This is formalized in the next two propositions, which echo the assumptions made on the (physical) conditional distributions of $\delta_t$ and of $y_t$, respectively by Assumptions 2 and 5.

**Proposition 2.5** *Under Assumptions 2, 3, 4 (or 5) and 6, and conditionally on $\mathcal{F}_t^*$, $\mathcal{D}_{t-1}$, the credit-event variables $\delta_{i,t}$ are Gamma-zero distributed under the risk-neutral probability measure $\mathbb{Q}$. In particular, there exists a risk-neutral credit-event intensity process $\lambda_{i,t}^{\mathbb{Q}}$ adapted to $\mathcal{F}_t^*$, $\mathcal{D}_{t-1}$, such that:*

$$\big(\mathsf{P}_{i,t}\,\big|\,\mathcal{F}_t^*,\, \mathcal{D}_{t-1}\big) \overset{\mathbb{Q}}{\sim} \mathcal{P}\left(\lambda_{i,t}^{\mathbb{Q}}\right) \qquad \text{and} \qquad \big(\delta_{i,t}\,\big|\,\mathsf{P}_{i,t}\big) \overset{\mathbb{Q}}{\sim} \Gamma_{\mathsf{P}_{i,t}}\left(\mu_{\delta_i}^{\mathbb{Q}}\right)\,, \tag{6}$$

*where $\mu_{\delta_i}^{\mathbb{Q}} = \dfrac{\mu_{\delta_i}}{1 - \mathbf{S}_i \mu_{\delta_i}}$ and the risk-neutral credit-event intensity is given by:*

$$\lambda_{i,t}^{\mathbb{Q}} = \alpha_{\lambda_i}^{\mathbb{Q}} + \beta_{\lambda_i}^{\mathbb{Q}^{(y)'}} y_t + \mathbf{C}_i^{\mathbb{Q}'} \delta_{t-1} = \frac{\lambda_{i,t}^{\mathbb{P}}}{1 - \mathbf{S}_i \mu_{\delta_i}}\,, \tag{7}$$

*with* $\alpha^{\mathbb{Q}}_{\lambda_i} = \frac{1}{1-\mathbf{S}_i\mu_{\delta_i}}\alpha_{\lambda_i}$, $\beta^{\mathbb{Q}(y)}_{\lambda_i} = \frac{1}{1-\mathbf{S}_i\mu_{\delta_i}}\beta^{(y)}_{\lambda_i}$, *and* $\mathbf{C}^{\mathbb{Q}}_i = \frac{1}{1-\mathbf{S}_i\mu_{\delta_i}}\mathbf{C}_i$.

**Proof** Straightforward application of the Esscher transform to the distribution associated with the conditional Laplace transform $\varphi_{\delta_{i,t}}(u;\, \lambda^{\mathbb{P}}_{i,t},\, \mu_{\delta_i}) = \mathbb{E}\left[\exp(u\delta_{i,t})\middle|\mathcal{F}^*_t,\, \mathcal{D}_{t-1}\right]$. ∎

Proposition 2.5 emphasizes an important property for the risk-neutral credit-event intensities. For standard credit-risk models, credit-event prices of risk are null ($\mathbf{S} = 0$), implying that physical and risk-neutral intensities are identical functions of all risk factors (see Equation 7). In other words, $\lambda^{\mathbb{Q}}_{i,t} = \lambda^{\mathbb{P}}_{i,t}$ if credit events are not priced by the representative investor. In that case, credit-risk premia arise only because the conditional moments of the risk factors $w_t$ are different when taken under the risk-neutral measure or under the historical one. Instead, when the credit-event risk is priced ($\mathbf{S}_i > 0$), the risk-neutral intensities become proportional to the physical ones, i.e. $\lambda^{\mathbb{Q}}_{i,t} = \frac{\lambda^{\mathbb{P}}_{i,t}}{1-\mathbf{S}_i\mu_{\delta_i}}$, as in Jarrow et al. (2005) and Driessen (2005) (see also Duffie, 2005, and references therein). This creates an additional wedge between the physical and risk-neutral moments of the credit-event variables $\delta_{i,t}$.

After $\delta_t$, let us now consider the risk-neutral conditional distribution of $y_t$.

**Proposition 2.6** *If we assume that the historical dynamics of $w_t$ is described by the recursive VARG process of Assumptions 2, 3 and 5, and under the SDF specification of Assumption 6, then, conditionally on $\mathcal{F}_{t-1}$ and under the risk-neutral measure, the components of $y_t$ are independent and follow a non-central Gamma distribution (as under the physical measure). More precisely, for $y_{j,t}$, $j \in \{1, \ldots, N_y\}$:*

$$\left(\mathsf{P}_{y_j,t}\,\middle|\,\mathcal{F}_{t-1}\right) \overset{\mathbb{Q}}{\sim} \mathcal{P}\left(\beta^{\mathbb{Q}(y)'}_{y_j} y_{t-1} + \mathbf{I}^{\mathbb{Q}'}_j \delta_{t-1}\right) \qquad \text{and} \qquad \left(y_{j,t}\,\middle|\,\mathsf{P}_{y_j,t}\right) \overset{\mathbb{Q}}{\sim} \Gamma_{\nu_{y_j}+\mathsf{P}_{y_j,t}}\left(\mu^{\mathbb{Q}}_{y_j}\right), \tag{8}$$

*where the risk-neutral parameters are given by:*

$$\beta^{\mathbb{Q}(y)}_{y_j} = \beta^{(y)}_{y_j}\frac{1}{1-\mu_{y_j}\widetilde{\theta}_{y_j}},\; \mu^{\mathbb{Q}}_{y_j} = \frac{\mu_{y_j}}{1-\mu_{y_j}\widetilde{\theta}_{y_j}}, \quad \text{and} \quad \mathbf{I}^{\mathbb{Q}}_j = \mathbf{I}_j\frac{1}{1-\mu_{y_j}\widetilde{\theta}_{y_j}},$$

*with*

$$\widetilde{\theta}_y = \theta_y + \beta_\lambda^{(y)} \frac{\mathbf{S} \odot \mu_\delta}{\mathbf{1} - \mathbf{S} \odot \mu_\delta}. \tag{9}$$

**Proof** Straightforward application of Propositions 2.5 and 2.4. See Monfort, Pegoraro, Renne, and Roussellet (2018) for a proof in a general framework. ■

Propositions 2.5 and 2.6 show that the risk-neutral distribution of the state vector $w_t$ thus remains recursive VARG. The recursive structure also underlines another important effect on the prices of risk applied to the common factors. Because $y_t$ instantaneously feeds back on the credit-event variables $\delta_t$, it is associated with two sources of priced risk in the modified risk-adjustment parameter $\widetilde{\theta}_y$. The representative investor can dislike upwards movements of $y_t$ itself, for instance because it represents worsened economic conditions; this is represented by $\theta_y$. Second, the investor prices movements in $y_t$ because of the frailty-like impact on the default probabilities of some entities (second term on the right-hand side of Equation 9). It is therefore natural to see $\mathbf{S}$, the prices of risk associated with credit-event variables, appearing in the pricing of the common factors $y_t$.

### 2.6 The Reverse-Order Multi-Horizon Laplace Transform

Before exploring the pricing properties of the framework introduced in the previous sections, we explain hereafter why and how multi-horizon Laplace transforms can be computed in a fast way when their arguments feature a so-called "reverse-order structure". This approach allows for an efficient computation of all default-sensitive asset prices we consider, i.e. any security whose cashflows can be expressed as exponential-affine combinations of $w_t$, or featuring terms of the form $\mathbb{1}_{\{\delta_{i,t}>0\}}$. We first introduce the general definition of the multi-horizon Laplace transform and we then discuss the particular reverse-order case.

**Proposition 2.7** *Let us consider a horizon $h \in \mathbb{N}$ and a set of arguments $(u_1, \ldots, u_h)$, where each vector $u_j$ is of dimension $N_y + n$. The h-period-ahead risk-neutral multi-horizon Laplace transform* *of the affine process $\{w_t\}$, for arguments $(u_1, \ldots, u_h)$, is given by:*

$$\begin{aligned}\varphi^{\mathbb{Q}}_{w_t}(u_1, \ldots, u_h) &:= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(\sum_{j=1}^{h} u_j' w_{t+j}\right) |\mathcal{F}_t\right] \\ &= \exp\Big(\mathcal{A}_h(u_1, \ldots, u_h)' w_t + \mathcal{B}_h(u_1, \ldots, u_h)\Big),\end{aligned} \tag{10}$$

*where the loadings functions $\mathcal{A}_h$ and $\mathcal{B}_h$ are defined for any arguments $(v_1, \ldots, v_h)$ through the following recursive system (for $h > 0$):*

$$\begin{cases} \mathcal{A}_h(v_1, \ldots, v_h) = A^{\mathbb{Q}}_w(v_1 + \mathcal{A}_{h-1}(v_2, \ldots, v_h)) \\ \mathcal{B}_h(v_1, \ldots, v_h) = B^{\mathbb{Q}}_w(v_1 + \mathcal{A}_{h-1}(v_2, \ldots, v_h)) + \mathcal{B}_{h-1}(v_2, \ldots, v_h), \end{cases} \tag{11}$$

*initialized with $\mathcal{A}_0 = \mathbf{0}$ and $\mathcal{B}_0 = 0$.*

***Proof*** *See Online Appendix A.1.3.* ■

For two sets of arguments $U_{h_1} = (u_1, \ldots, u_{h_1})$ and $U_{h_2} = (u_1, \ldots, u_{h_2})$, with $h_1 < h_2$, the respective $h_1$-step and $h_2$-step recursions (11) have to be run separately. More precisely, the multiple calls to functions $A^{\mathbb{Q}}_w$ and $B^{\mathbb{Q}}_w$ (Equation 11), that are necessary to evaluate the multi-horizon Laplace transform $\varphi^{\mathbb{Q}}_{w_t}(u_1, \ldots, u_{h_1})$, are of no use to compute $\varphi^{\mathbb{Q}}_{w_t}(u_1, \ldots, u_{h_2})$ even if the first $h_1$ vectors of $U_{h_1}$ and $U_{h_2}$ are the same, because the arguments to be called in $A^{\mathbb{Q}}_w$ and $B^{\mathbb{Q}}_w$ are never the same. For example, while the computation of $\varphi^{\mathbb{Q}}_{w_t}(u_1)$ involves $A^{\mathbb{Q}}_w(u_1)$ and $B^{\mathbb{Q}}_w(u_1)$, the computation of $\varphi^{\mathbb{Q}}_{w_t}(u_1, u_2)$ involves $A^{\mathbb{Q}}_w(u_2)$, $B^{\mathbb{Q}}_w(u_2)$, $A^{\mathbb{Q}}_w(u_1 + A^{\mathbb{Q}}_w(u_2))$ and $B^{\mathbb{Q}}_w(u_1 + A^{\mathbb{Q}}_w(u_2))$. If $u_1 \neq u_2$, there is no overlap in the computations.

From a computational point of view, the situation is more favorable when the different sets of considered arguments are (also) nested, but organized in "reverse order"; that is when we want to compute $\varphi^{\mathbb{Q}}_{w_t}(u_h, \ldots, u_1)$, with $h$ growing. Indeed, simply changing the order of the arguments of

$\mathcal{A}_h$ and $\mathcal{B}_h$ in Equation (11), we see that:

$$\begin{cases} \mathcal{A}_h(u_h,\dots,u_1) &= A_w^{\mathbb{Q}}\left(u_h + \mathcal{A}_{h-1}(u_{h-1},\dots,u_1)\right) \\ \\ \mathcal{B}_h(u_h,\dots,u_1) &= B_w^{\mathbb{Q}}\left(u_h + \mathcal{A}_{h-1}(u_{h-1},\dots,u_1)\right) + \mathcal{B}_{h-1}(u_{h-1},\dots,u_1)\,, \end{cases} \tag{12}$$

which shows that $\mathcal{A}_h(u_h,\dots,u_1)$ and $\mathcal{B}_h(u_h,\dots,u_1)$ – which determine $\varphi_{w_t}^{\mathbb{Q}}\,(u_h,\,\dots,\,u_1)$ – are directly deduced from $\mathcal{A}_{h-1}(u_{h-1},\dots,u_1)$ and $\mathcal{B}_{h-1}(u_{h-1},\dots,u_1)$ – which themselves determine $\varphi_{w_t}^{\mathbb{Q}}\,(u_{h-1},\,\dots,\,u_1)$. Looking back at the example of the previous paragraph, the computation of $\varphi_{w_t}^{\mathbb{Q}}(u_1)$ and $\varphi_{w_t}^{\mathbb{Q}}(u_2,u_1)$ both involve $A_w^{\mathbb{Q}}(u_1)$ and $B_w^{\mathbb{Q}}(u_1)$ as a first step.

In other words, the computations of $\varphi_{w_t}^{\mathbb{Q}}\,(u_1)$, $\varphi_{w_t}^{\mathbb{Q}}\,(u_2,u_1)$, $\dots$, $\varphi_{w_t}^{\mathbb{Q}}\,(u_h,\,\dots,\,u_1)$ involve $h$ calls of functions $A_w^{\mathbb{Q}}$ and $B_w^{\mathbb{Q}}$, against $h(h+1)/2$ calls when the arguments do not feature the reverse-order structure, that is for instance when we have to compute $\varphi_{w_t}^{\mathbb{Q}}\,(u_1)$, $\varphi_{w_t}^{\mathbb{Q}}\,(u_1,u_2)$, $\dots$, $\varphi_{w_t}^{\mathbb{Q}}\,(u_1,\,\dots,\,u_h)$

The following subsections make an intensive use of conditional multi-horizon Laplace transforms applied on arguments satisfying the reverse-order property. Moreover, the involved reverse-order structure is systematically characterized by only two vectors $u$ and $v$ (say), i.e. $\varphi_{w_t}^{\mathbb{Q}}\,(u,\,\dots,\,u,\,v)$. For ease of presentation, we will adopt the following notation:

$$\varphi_{w_{t(h)}}^{\mathbb{Q}}\,(u,\,v) = \varphi_{w_t}^{\mathbb{Q}}\,(u,\,\dots,\,u,\,v) \tag{13}$$

and, by abuse of notation, we will replace $\mathcal{A}_h(u,\dots,u,v)$ and $\mathcal{B}_h(u,\dots,u,v)$ by $\mathcal{A}_h(u,v)$ and $\mathcal{B}_h(u,v)$, respectively.

## 3 Defaultable Asset Pricing

In this section, we consider the problem of computing the no-arbitrage price of default-free and defaultable securities under Assumptions 2, 3, 4 (or 5) and 6. We focus on two types of assets, namely bonds and credit default swaps (CDSs). We start by providing notations and assumptions regarding the assets' payoffs in case of default.

### 3.1 Recovery Rate and Recovery Conventions

We denote by $B^*(t,h)$ the price of a default-free bond of residual maturity $h$ at time $t$, and by $B_i(t,h)$ the price of a defaultable bond issued by entity $i$. By no-arbitrage, we have:

$$B^*(t,h) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r_t - \dots - r_{t+h-1}} \,\middle|\, \mathcal{F}_t\right] .$$

As is standard in affine models, the price of this bond is obtained as a closed-form exponential-affine function of $w_t$.

**Proposition 3.1** *The price of the risk-free bond of any residual maturity $h$ can be computed explicitly through the multi-horizon Laplace transform recursions as:*

$$\begin{aligned} B^*(t,h) &= \exp\left[-r_t - (h-1)\xi_0\right] \times \varphi^{\mathbb{Q}}_{w_t(h-1)}(-\xi,-\xi) \\ &= \exp\left\{-h\xi_0 + \left[\mathcal{A}_{h-1}(-\xi,\,-\xi) - \xi\right]' w_t + \mathcal{B}_{h-1}(-\xi,\,-\xi)\right\} , \end{aligned}$$

*where $\xi = (\xi_y,\, \xi_\delta)$ is defined in Equation (4) and where functions $\mathcal{A}_{h-1}$ and $\mathcal{B}_{h-1}$ can be evaluated using system (12).*

We now consider the case of defaultable bonds and, for ease of notation, $\tau_i^{(k)}$ defined in Assumption 1 will be denoted $\tau_i$. We assume that when entity $i$ defaults, all its outstanding bonds are terminated and provide a (possibly stochastic) recovery payment for each unit of face value. The following definition summarizes the potential recovery assumptions commonly adopted in the literature (see Brennan and Schwartz, 1980; Duffie, 1998; Duffie and Singleton, 1999).

**Definition 3.1** *The recovery payment for bond $B_i(\tau_i,h)$ in case of default of the issuer entity $i \in \{1,\,\dots,\,n\}$ at time $\tau_i$ is given by the product of a recovery rate and a recovery payment, denoted by $\varrho_{i,\tau_i}$ and $\Pi_{i,\tau_i}(h)$, respectively. The following assumptions on the recovery payment are commonly* *made:*

- *Recovery of Market Value (RMV):* $\Pi_{i,\tau_i}(h) = \widetilde{B}_i(\tau_i, h)$
- *Recovery of Face Value (RFV):* $\Pi_{i,\tau_i}(h) = 1$
- *Recovery of Treasury (RT):* $\Pi_{i,\tau_i}(h) = B^*(\tau_i, h)$,

*where* $\widetilde{B}_i(\tau_i, h)$ *would be the price of the defaultable bond of entity* $i$ *at time* $\tau_i$ *if there had been no credit event.*

**Assumption 7** *The recovery rate is given by an exponential-affine function of the state process* $w_t$*, written as:*

$$\varrho_{i,\tau_i} = \exp\left(-\omega_{i,0} - \omega_i^{(y)'}\, y_{\tau_i} - \omega_i^{(\delta)}\, \delta_{i,\tau_i}\right) , \tag{14}$$

*or, more compactly,* $\varrho_{i,\tau_i} = \exp\left(-\omega_{i,0} - \omega_i^{(w)'}\, w_{\tau_i}\right)$*, where all parameters are of adapted size.*

Whenever we assume that the entire dynamics of $w_t$ are defined by Assumptions 2, 3 and 5, it is easy to impose that the recovery rate is bounded between 0 and 1 by forcing all parameters in Equation (14) to be non-negative. This equation specifies a stochastic recovery rate whose time-varying magnitude may depend on common and entity-specific factors. An interesting particular case arises when we simply assume $\omega_{i,0} = 0$, $\omega_i^{(y)} = \mathbf{0}$ and $\omega_i^{(\delta)} = 1$:

$$\varrho_{i,\tau_i} = \exp\left(-\ \delta_{i,\tau_i}\right) . \tag{15}$$

This specification delivers a clean interpretation of the process $\delta_{i,t}$. Relation (15) indeed formalizes the idea of a stochastic recovery rate equal to one as long as $\delta_{i,t} = 0$, and leaving the unitary upper bound at the default time $\tau_i$ with a reduction whose magnitude depends on the size of the credit-event variable on the default date ($\delta_{i,\tau_i}$). The size of the jump in $\delta_{i,t}$ represents of the "severity" of the credit event. It jointly determines the effective recovery rate on bonds, the systemic impact on common factors (through $\mathbf{I}$ parameters) and the increase in other entities credit-event intensities $\lambda_t$

through direct contagion effects (parameters $\mathbf{C}$).

We focus hereafter on the RMV and RFV conventions. The RT case is presented in the online Appendix A.1.9. We consider the recovery rate specification (15) for the RMV convention. The RFV case is treated under the more general specification (14).

### 3.2 Defaultable Bond Pricing

We first establish a general expression for the price of a defaultable bond – without specifying the recovery convention. Assume that entity $i$ has not defaulted at time $t$. It is useful to rewrite the default indicator formalizing a default event happening at time $t+k$ as:

$$\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}} \times \mathbb{1}_{\{\delta_{i,t+k}>0\}} = \mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}} - \mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}. \tag{16}$$

The bond trading at price $B_i(t,h)$ on date $t$ provides its holder with a single payoff between dates $t+1$ and $t+h$: this payoff is either $\varrho_{i,t+k}\Pi_{i,t+k}(h-k)$ (settled on date $t+k$) if the credit event happens at time $t+k$ ($k \leqslant h$) or 1 (settled on date $t+h$) if the default does not happen during the life of the bond ($t+h<\tau_i$). Accordingly, the price of the bond has to satisfy:

$$\begin{aligned} B_i(t,h) &= \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{\ell=0}^{k-1} r_{t+\ell}\right)\left[\mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}} - \mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}}\right]\varrho_{i,t+k}\Pi_{i,t+k}(h-k) \,\middle|\, \mathcal{F}_t\right] \\ &+ \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{\ell=0}^{h-1} r_{t+\ell}\right)\mathbb{1}_{\{\delta_{i,t:t+h}=\mathbf{0}\}} \,\middle|\, \mathcal{F}_t\right]. \end{aligned} \tag{17}$$

Looking at (17), it is not obvious that the conditional expectations will be obtained in closed-form, nor is it clear that multi-horizon Laplace transforms will be useful to compute such quantities. The following Lemma will prove to be crucial for the derivation of closed-form pricing formula[18]

**Lemma 3.1** *Let $Z_1$ be a random variable valued in $\mathbb{R}^d$ ($d \geq 1$) and $Z_2$ be a random variable valued in $\mathbb{R}^+ = [0,+\infty)$. Suppose that $\mathbb{E}\left[\exp\left(u_1' Z_1 - u_2 Z_2\right)\right]$ exists for a given $u_1$ and $u_2 \geq 0$. Then, we*

[18]This lemma can be seen as a generalization of Lemma 2.1 in Monfort et al. (2017) (see also Chen and Filipovic, 2007).

*have:*

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

**Proof** See Online Appendix A.1.4. ■

Replacing $Z_1$ by sets of future $w_t$ and $Z_2$ by sums of $\delta_{t+\ell+1}$, we obtain some of the expectations appearing in Equation (17) as limits of the multi-horizon Laplace transform (whose computation benefits from a reverse order structure, see Subsection 2.6). Depending on the recovery assumption, the pricing formulas can simplify further and are summarized by the following propositions.

**Proposition 3.2** *Under the RMV assumption (see Definition 3.1), assuming that the recovery rate is given by Equation (15), then the no-arbitrage price of the defaultable bond satisfies:*

$$B_i(t,h) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{\ell=0}^{h-1}(r_{t+\ell}+\delta_{i,t+\ell+1})\right)\Big|\mathcal{F}_t\right], \tag{18}$$

*and can be computed as follows:*

$$\begin{aligned} B_i(t,h) &= \exp\left[-r_t-(h-1)\xi_0\right]\times\varphi^{\mathbb{Q}}_{w_t(h)}\left(-\xi-\mathrm{e}_{\delta_i},\,-\mathrm{e}_{\delta_i}\right) \\ &= \exp\left\{-h\xi_0+\left[\mathcal{A}_h\left(-\xi-\mathrm{e}_{\delta_i},\,-\mathrm{e}_{\delta_i}\right)-\xi\right]' w_t+\mathcal{B}_h\left(-\xi-\mathrm{e}_{\delta_i},\,-\mathrm{e}_{\delta_i}\right)\right\} , \end{aligned} \tag{19}$$

*where the vector $\mathrm{e}_{\delta_i}$ is a selection vector such that $\mathrm{e}_{\delta_i}' w_t = \delta_{i,t}$ and where functions $\mathcal{A}_h$ and $\mathcal{B}_h$ can be evaluated using system (11).*

**Proof** See online Appendix A.1.5. ■

Equation (18) is a key result of the paper. It shows that our framework still leads to the familiar no-arbitrage bond pricing formula based on the default-adjusted short rate $(r_{t+\ell}+\delta_{i,t+\ell+1})$, in spite of the fact that the credit events are sources of risk that are priced. This result can be seen as a discrete-time generalization of the RMV setting proposed by Duffie and Singleton (1999).

In the case where no default is observed throughout a sample, $\delta_{i,t}$ is uniformly equal to zero for all observed dates. However, the price of defaultable bonds do not collapse to that of risk-free bonds as long as the distribution of future $\delta_{i,t}$'s, conditional on $\mathcal{F}_t$, is not concentrated at zero.

**Proposition 3.3** *Under the RFV assumption, assuming that the recovery rate is given by Equation (14), the no-arbitrage price of the defaultable bond is given by the following sum of exponentially-affine functions:*

$$\begin{aligned} B_i(t,h) &= \lim_{u\to+\infty} e^{-r_t}\left[\sum_{k=1}^{h} e^{-\omega_{i,0}-ue'_{\delta_i}w_t-(k-1)\xi_0}\left(\varphi^{\mathbb{Q}}_{w_t(k)}\left[-\xi-ue_{\delta_i},\,-\omega_i^{(w)}\right]-\varphi^{\mathbb{Q}}_{w_t(k)}\left[-\xi-ue_{\delta_i},\,-ue_{\delta_i}-\omega_i^{(w)}\right]\right)\right. \\ &\qquad \left. +\; e^{-(h-1)\xi_0-ue'_{\delta_i}w_t}\varphi^{\mathbb{Q}}_{w_t(h)}\left[-\xi-ue_{\delta_i},\,-ue_{\delta_i}\right]\right], \end{aligned} \tag{20}$$

*where $\varphi^{\mathbb{Q}}_{w_t(k)}(u,v)=\exp\{\mathcal{A}_k(u,v)'w_t+\mathcal{B}_k(u,v)\}$ is the conditional multi-Laplace transform that can be evaluated using system (12).*

**Proof** See online Appendix A.1.6. ■

In the following section, we will use the pricing formula (20) to obtain the closed-form CDS pricing result. For the sake of precision, we will use the notation $B_i^{\mathrm{RFV}}\left(t,h;\,\omega_{i,0},\,\omega_i^{(w)}\right)$ to refer to this defaultable bond price under RFV with the recovery rate defined as in Equation (14).

### 3.3 Credit Default Swap (CDS) Valuation

Let us now consider the problem of CDS pricing. A CDS is a derivative contract where a protection buyer accepts to regularly pay a fixed rate called CDS premium (or spread) to a protection seller as long as the underlying entity does not suffer a credit event. In case of default, the contract terminates and the protection seller provides the loss-given-default on a reference bond to the protection buyer.[19] We denote by $\mathcal{S}_i(t,h)$ this CDS spread, set at date $t$ with maturity $t+h$. We

[19]This description of the CDS contract is stylized and neglects particular institutional features such as the auction process when a credit event is triggered, the cheapest-to-deliver premium, potential counterparty and liquidity risk embedded in these contracts, or the fact that CDSs can be quoted in a different currency than the underlying bond. Although our assumptions may appear simplistic, they are in line with most of the reduced-form CDS term structure literature.

assume in the following that the notional is equal to one. The CDS spread is such that the present value of the payments made by the protection buyer (the fixed leg) is equal to 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 $t+k$ ($k \leq h$), the cash flow on this date is $\mathcal{S}_i(t,h)$, and is independent of $k$. The present value of the fixed-leg payments is denoted by $\mathrm{PB}_i(t,h)$ and is given by:

$$\mathrm{PB}_i(t,h) = \mathcal{S}_i(t,h) \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}} \left[ \exp\left( -\sum_{\ell=0}^{k-1} r_{t+\ell} \right) \mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}} \,\Big|\, \mathcal{F}_t \right]. \tag{21}$$

Under the RFV convention, the protection seller will make a payment of $(1-\varrho_{i,t+k})$ (the Loss-Given-Default) at date $t+k$ in case of default over the time interval $]t+k-1, t+k]$. The present value of this promised payment is given by:

$$\mathrm{PS}_i(t,h) = \sum_{k=1}^{h} \mathbb{E}^{\mathbb{Q}} \left[ \exp\left( -\sum_{\ell=0}^{k-1} r_{t+\ell} \right) (1-\varrho_{i,t+k}) \left[ \mathbb{1}_{\{\delta_{i,t:t+k-1}=\mathbf{0}\}} - \mathbb{1}_{\{\delta_{i,t:t+k}=\mathbf{0}\}} \right] \,\Big|\, \mathcal{F}_t \right]. \tag{22}$$

Both expressions can be easily obtained using the same methods as for defaultable bonds. The pricing result is expressed in the following proposition.

**Proposition 3.4** *The no-arbitrage CDS spread $\mathcal{S}_i(t,h)$, negotiated at date $t$ and associated with a credit default swap (CDS) maturing in $h$ periods, is such that the present values of the protection buyer and seller are equal, thus given by:*

$$\mathcal{S}_i(t,h) = \frac{B_i^{\mathrm{RFV}}(t,h;\, 0,\, \mathbf{0}) - B_i^{\mathrm{RFV}}\left(t,h;\, \omega_{i,0},\, \omega_i^{(w)}\right)}{\lim\limits_{u\to+\infty} \sum_{k=1}^{h} e^{-k\xi_0 - (\xi + u e_{\delta_i})' w_t} \times \varphi^{\mathbb{Q}}_{w_t(k)}\left(-\xi - u e_{\delta_i}, -u e_{\delta_i}\right)}, \tag{23}$$

*where $B_i^{\mathrm{RFV}}$ is given by Equation (20) in Proposition 3.3.*

**Proof** See online Appendix A.1.7. ∎

As for bond prices under the RFV convention, CDS spreads are explicit but not given by an exponential-affine function of $w_t$. (Equation (23) expresses them as ratios of sums of exponentially-affine functions.) It should be noted that computing bond prices and CDS spreads partly involves the same multi-horizon Laplace transforms, thus reducing the overall number of recursions that need to be performed to price all assets in the economy. The applications presented in Section 4 consider CDS spreads denominated both in domestic and foreign currencies. The latter case is presented in Appendix A.1.8.

# 4 Applications

This section gathers the results of several empirical illustrations of our framework. We first conduct a Monte Carlo experiment to observe the influence of the different credit-risk channels authorized by our framework, and to compare the respective performances of maximum likelihood and moment-based estimation methods to identify these channels. The results are presented in Subsection 4.1. In the remaining subsections, we exploit our framework to study the pricing of sovereign credit risk. Using euro-area data, we notably explore the influence of allowing for the pricing of credit events, we compare frailty and contagion channels, and we extract measures of depreciation-at-default from CDS denominated in different currencies.

## 4.1 A Stylized Calibration and Simulation Exercise

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 can be found in online Appendices A.3 and A.4.

We consider a stylized economy with two defaultable entities whose credit-event intensities are driven by a single common latent factor $y_t$, independent from the autonomous short-term 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 significantly increases the slope of the CDS curve with respect to the two other cases.

Next, 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, authorizing 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. This results, which comes from the fact that $\lambda^{\mathbb{Q}}_{2,t} > \lambda^{\mathbb{P}}_{2,t}$ when $\mathbf{S}_2 > 0$, will also be discussed in Subsection 4.6, on real data. These simulations also illustrate how second-order moments may help distinguish between different mechanisms: the auto-correlations of simulated spreads are for instance lower for direct than for indirect *contagion*.[20]

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,

[20] 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.

discriminating whether defaults are observed or not (see online Appendix A.4). 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 below.) Second, including the credit-event variables $\delta_t$ as observables in the filter increases the quality of estimation, even when no defaults are observed (i.e. when $\delta_t = 0$ for all dates $t$). Last, observing in-sample defaults improves the ability of the filter to correctly identify the mechanisms at play.

### 4.2 A Model for European Sovereign Credit Risk

In this section, we exploit the framework presented above to study the pricing of sovereign credit risk. We focus on five euro-area countries: the four largest ones, that are France, Germany, Italy and Spain – accounting for 75% of the 19-country euro area GDP – and Greece, which defaulted on March 9, 2012.[21] In spite of the high credit quality of the first four countries, the associated sovereign CDS spreads have reached relatively high levels over the last twelve years, especially during the so-called euro-area sovereign debt crisis initiated in late 2009.

The framework we propose can be seen as an extension of Ait-Sahalia et al. (2014) along the following dimensions: (a) whereas the models estimated in the latter study involve pairs of countries, ours jointly accounts for five economies, (b) by specifying the SDF, we explicitly model investors' risk preferences, opening the door to risk-premium analysis and the extraction of physical probabilities of default and (c) our model allows for both a common factor and country-specific ones, while

[21] On March 9, 2012, the International Swaps and Derivatives Association (ISDA) decided the payment on Greek CDSs. The ISDA indeed considered that the Greek legislation that forced losses on all private creditors constituted a credit event.

Ait-Sahalia et al. (2014)'s models entail only country-specific factors.

We estimate different versions of the model, which allows us to revisit three credit-risk issues in the sovereign context. First, we examine the influence of allowing for credit-event pricing, that is when sovereign defaults directly affect the SDF (Subsection 4.6). Second, we discuss the differences resulting from allowing for frailty and/or contagion in the model (Subsection 4.7). Third, we extend the model in order to investigate quanto CDS spreads, that are spread differentials between euro- and dollar-denominated CDSs (Subsection 4.8).

We consider $n = 5$ economies. There are three types of components in vector $y_t$: the first is the short-term rate $r_t$; the second, denoted by $z_t$, is a frailty factor influencing all countries; the last $n$ components of $y_t$, gathered in $x_t = (x_{1,t}, \dots, x_{n,t})$, are country-specific factors, in the sense that $x_{i,t}$ intervenes only in the default intensity of country $i$. The historical default intensities are given by:

$$\lambda^{\mathbb{P}}_{i,t} = \beta^{(x)}_{\lambda,i} x_{i,t} + \mathbf{C}'_i \delta_{t-1}, \tag{24}$$

where $\beta^{(x)}_{\lambda,i}$ is a scalar and where vector $\mathbf{C}_i$ is of the form $c_i \boldsymbol{\kappa}_c$, $i = 1, \dots, n$, $c_i$ being a non-negative scalar and $\boldsymbol{\kappa}_c$ being an $n$-dimensional vector of country weights (summing to one). To simplify the analysis, we do not allow for potential indirect contagion here. More precisely, we set the $\mathbf{I}_j$ parameters appearing in Equation (3) to zero.[22]

Equation (24) is consistent with the general formulation (2), with $\alpha_{\lambda_i} = 0$ and $\beta^{(y)}_{\lambda_i} = (0, 0, \beta^{(x)}_{\lambda,i} e_i)$, where $e_i$ denotes the $i^{th}$ column of the $n \times n$ identity matrix. Conditionally on $\mathcal{F}_{t-1}$, the components of $y_t$ are independent and we have:

$$\begin{aligned}
& r_t \big| \mathsf{P}_{r,t}, \mathcal{F}_{t-1} \sim \gamma_{\mathsf{P}_{r,t}}(\mu_r) && \text{where} \quad \mathsf{P}_{r,t} \big| \mathcal{F}_{t-1} \sim \mathcal{P}\left(\alpha_r + \beta_r r_{t-1}\right), \\
& z_t \big| \mathsf{P}_{z,t}, \mathcal{F}_{t-1} \sim \Gamma_{\nu_z + \mathsf{P}_{z,t}}(1) && \text{where} \quad \mathsf{P}_{z,t} \big| \mathcal{F}_{t-1} \sim \mathcal{P}\left(\beta^{(z)}_z z_{t-1}\right), \\
& x_{i,t} \big| \mathsf{P}_{x_i,t}, \mathcal{F}_{t-1} \sim \Gamma_{\mathsf{P}_{\nu_x + x_i,t}}(1) && \text{where} \quad \mathsf{P}_{x_i,t} \big| \mathcal{F}_{t-1} \sim \mathcal{P}\left(\beta^{(z)}_x z_{t-1} + \beta^{(x)}_x x_{i,t-1}\right).
\end{aligned} \tag{25}$$

[22] Preliminary estimations pointed towards the non-significance of such parameters.

Last, the one-period SDF is given by:

$$M_{t-1,t} = \exp\left(-r_{t-1} + \theta_r r_t + \theta_z z_t + \boldsymbol{\kappa}'_M(\theta_x x_t + \mathbf{S}\delta_t) - \psi^{\mathbb{P}}_{w,t-1}(\theta_w)\right), \tag{26}$$

where $\boldsymbol{\kappa}_M$ is a vector of country weights summing to one and where $\theta_w = (\theta_r, \theta_z, \theta_x \boldsymbol{\kappa}_M, \mathbf{S}\boldsymbol{\kappa}_M)$. The previous formulation is consistent with Equation (5).

## 4.3 Data

The data are monthly and cover the period from January 2007 to July 2019 (ends of month). CDS spreads and proxies of risk-free zero-coupon yields are extracted from Thomson Reuters Datastream.[23] We remove CDS spreads that do not fluctuate for three consecutive months, for this indicates low trading volumes. (In Datastream, in the absence of quotes, the last-observed ones are repeated.) We also remove (Greek) CDS spreads that are higher than 20,000 basis points.[24]

For CDS spreads and risk-free yields, the following maturities are considered: 1, 2, 3, 5 and 10 years. We therefore have 35 measurement equations: 25 (=$5 \times n$) correspond to CDS spreads, 5 correspond to risk-free zero-coupon yields and 5 ($= n$) correspond to the $\delta_{i,t}$. The latter are all null except for one instance: for March 2012, when the Greek sovereign default took place, the Greek credit-event variable $\delta_{i,t}$ is set to $-\log(0.22)$, consistently with an observed recovery rate of 22% (see Coudert and Gex, 2013).

## 4.4 Estimation Strategy

Most of the parameters are estimated by maximizing the (approximate) likelihood function obtained as a by-product of the extended Kalman filter (see online Appendix A.4.2 for details and

[23]The zero-coupon yields are bootstrapped from the euro swap yield curve by Thomson Reuters Datastream; mnemonics are BDWX0073R (1-year maturity) to BDWX0082R (10-year maturity).

[24]The reason why CDS spreads can potentially be above 10,000 basis points (100%) is that the payments of the premium leg are usually made on a quarterly basis (the payment being equal to the annualized spread divided by 4). The CDS spread can therefore be equal to up to 40,000 basis points if the default is almost certain in the coming month and if the recovery rate is expected to be close to zero.

references regarding this type of estimation technique). To facilitate or discipline the estimation, some parameters are calibrated or constrained.

First, in order to diminish the number of parameters to be estimated, we assume that the $n$ components of $\kappa_M$ (Equation 26) are functions of countries' Gross Domestic Product (GDP). Specifically, denoting by $GDP_i$ the GDP of country $i$, we assume that $\kappa_M$ is proportional to $[GDP_1^\ell, \dots, GDP_n^\ell]$, where $\ell$ is a parameter to be estimated.[25] Second, we set $\mu_{\delta_i} = 0.6$, which makes our model consistent with the 1983-2015 average of sovereign-default recovery rates (see online Appendix A.5.1, making use of the Moody's, 2016, dataset). Third, in the spirit of Cochrane and Saa-Requejo (2000), we impose an upper bound for the sample average of the maximum one-year Sharpe ratio. As advocated by Cochrane and Saa-Requejo (2000), this bound is set to 1. It is worth noting that the fact that maximum Sharpe ratios are available in close form in an affine model is instrumental to make this approach feasible (see online Appendix A.5.2 for details regarding this computation).[26]

Table A.10 reports parameter estimates for eight versions of the model, whose parameterizations are summarized in Table 1. Models (1) is the "complete" version of the model, allowing for pricing of credit events (i.e. $\mathbf{S} \neq 0$), contagion (i.e. $\mathbf{C}_i \neq 0$) and a frailty component $z_t$. The latter component is absent from Models (5) to (8). Models (2), (4), (6) and (8) do not allow for contagion. Finally, $\mathbf{S}$ is set to zero in Models (3), (4), (7) and (8).

The penultimate row of Table A.10 reports the sample average of the maximum Sharpe ratios; the bound appears to be binding for half of the estimated models, including the complete one (Model (1)). The maximum values of the log-likelihood functions are reported in the last row of Table A.10; the highest value is naturally obtained for Model (1), the most complete model.

[25] We use 2018 GDPs, as measured by Eurostat.

[26] Preliminary estimations of the model – without this third restriction – yielded to extreme and unreasonable maximum Sharpe ratios, which was reflected in extremely large credit-risk premia. The last phenomenon echoes findings by Duffee (2010), who documents that maximum Sharpe ratios are often far too large when one estimates unconstrained no-arbitrage yield curve models.

## 4.5 Estimation Results (Complete Model)

Before turning to the comparison of the different models (Subsections 4.6 and 4.7), let us focus on the model featuring credit-event pricing, contagion and a frailty component (Model (1)).

Figure 1 illustrates the model fit by comparing observed CDS spreads (black crosses) to the model-implied ones (grey solid line). It appears that the model is able to capture a large share of credit spreads' fluctuations, across time and maturities. On the same figure, black solid lines represent those counterfactual CDS spreads that would be observed if the prices of risk $\theta_r$, $\theta_z$, $\theta_x$ and $\mathbf{S}$ (see Equation (26)) were equal to zero. This characterizes a counterfactual world where investors are not risk averse. By definition, credit-risk premia are the differentials between the latter spreads – dubbed "$\mathbb{P}$ CDS spreads" – and model-implied ones – the "$\mathbb{Q}$ CDS spreads". Figure 1 therefore confirms that credit-risk premia are substantial for all maturities, including short ones (12 months). As discussed in the next subsection, allowing for credit-event pricing (i.e. $\mathbf{S} > 0$) is instrumental to obtain sizable short term credit-risk premia.

The existence of credit-risk premia translates into differences between physical and risk-neutral probabilities of default (Figure 2). Let us consider the average ratios between risk-neutral and physical default probabilities. Figure 2 shows that, at the five-year horizon, these $\mathbb{Q}/\mathbb{P}$ ratios go from 1.3 (Greece) to 3.5 (France and Germany). These ratios are intermediary for Italy and Spain, with respective values of 2.6 and 2.9. The previous observations are suggestive of a negative relationship between the credit-riskiness of a country and the $\mathbb{Q}/\mathbb{P}$ ratios. This finding echoes results from the corporate credit-risk literature, according to which the part of spreads accounted for by credit loss expectations reflects a smaller fraction of yield spreads for investment-grade bonds than for lower credit-quality bonds (see e.g. Table 1 of D'Amato and Remolona, 2003, or Huang and Huang, 2012).

Estimated factors are shown in the online Appendix (Figure A.4). Unsurprisingly given its construction, the frailty factor $z_t$ is strongly correlated to the CDS spreads, with an average correlation of 57% across countries and maturities. Most of CDSs' variability is explained by country-specific factors $x_{i,t}$ up to the end of 2012. From 2013 onward, the frailty factor $z_t$ takes over explaining the term structures of all countries' CDSs whereas all factors $x_{i,t}$ converge to virtually zero and barely fluctuate. Unreported results – available upon request – also show that this factor relates to the euro-area unemployment rate and to the European economic policy uncertainty indicator computed by Baker et al. (2016).

### 4.6 The Pricing of Sovereign Credit Events

A limited number of credit-risk studies explicitly distinguish the risk of credit spread changes – if no default occurs – from the risk of the default event itself. As shown by Jarrow et al. (2005), the latter form of risk, called credit-event risk, cannot be priced if default jumps are conditionally independent across an infinite number of entities. As a consequence, after having found evidence of credit-event pricing in the context of large U.S. private firms, Driessen (2005) concludes that default jumps are not conditionally independent across the considered firms, or that not enough corporate bonds are traded to fully diversify the default jump risk. By contrast, Bai et al. (2015) find that when contagion is introduced within a general equilibrium framework for an economy comprising a large number of firms, credit-event risk premia have an upper bound of a few basis points.[27]

Because the number of sovereign entities is far smaller than the number of private borrowers, Jarrow et al. (2005)'s conditions for the absence of default event pricing are a priori not satisfied in the sovereign credit-risk context. And, as a matter of fact, our econometric results point to the existence of sovereign default event pricing. Indeed, the differences between the maximum log-likelihoods obtained for the versions of the model where $\mathbf{S}$ is restricted to be null and the respective versions where $\mathbf{S}$ is free – e.g. Model (3) vs Model (1), or Model (4) vs Model (2) – are well above critical values based on $\chi^2$ distributions.

What are the economic implications of sovereign credit-event pricing? Recall that if $\mathbf{S} = 0$, then

[27] Bai et al. (2015) focus on the returns of the asset value of firms and do not explicitly consider the prices of medium-term to long-term financial instruments. In their model, each firm is associated with a jump process. The asset value of firm $i$ falls when its own jump process is activated or, to a lesser extent, when it is the case for the jump process of another firm (which is how contagion is modeled). Because the jump intensities are constant, this model does not feature self- nor cross-excitation.

the physical and risk-neutral default intensities are the same (see Subsection 2.5). Since default intensities are closely related to the one-period credit spreads, having $\mathbf{S} = 0$ tends to contain credit-risk premia for short maturities. This is illustrated by Figure 3, which displays, for Model (1) and for two different dates, the model-implied term structures of CDS spreads (black solid lines), together with counterfactual term structures that are obtained after setting $\mathbf{S}$ to zero, all else being equal (dotted line). We also show the spreads that would prevail if all prices of risk were equal to zero (grey solid line). It appears that, for short-maturities, credit-risk premia are essentially accounted for by the credit-event price of risk $\mathbf{S}$.

The influence of credit-event pricing is further illustrated by Table 2, which reports the average shares of CDS spreads accounted for by credit-risk premia across the different estimated models and for two horizons: one short horizon of 6 months and a longer horizon of 5 years. In this table, figures in bold font indicate the models where $\mathbf{S}$ is allowed to be strictly positive. As expected, even for short horizons, credit-risk premia are substantial for those models allowing for credit-event pricing. For instance, the 6-month credit-risk premia amount to about 50% of CDS spreads in Model (1), against 10% in Model (3), the only difference between these two models being that $\mathbf{S} = 0$ (before estimation of the remaining parameters) in the latter. Differences are lower for larger maturities.

These results suggest that one may substantially underestimate short-term credit-risk premia when using a credit-risk model that does not allow for credit-event pricing. This, in turn, may lead to overestimation of short-term probabilities of default.

### 4.7 Frailty and Contagion

This application is also connected to the strand of the credit-risk literature exploring the influence of frailty and/or contagion mechanisms in credit-risk models.

Exploring the default history of U.S. industrial firms between 1979 and 2004, Das et al. (2007) show that one cannot reconcile observed default clustering to standard credit-risk models where the default intensities depend solely on observable macro-finance variables. Default clustering can however be accounted for by "frailty," by which many firms can be jointly exposed to one or more unobservable risk factors (see Duffie et al., 2009). Another potential source of clustering is contagion, through which the default by one entity has a direct impact on the health of other firms. Azizpour et al. (2018) reject the hypothesis that U.S. corporates' default times are correlated only because of frailty-like mechanisms.

The previous studies build on the availability of rich datasets of corporate defaults. Sovereign defaults are far rarer, making it particularly challenging to disentangle frailty from contagion in the sovereign context.[28] An advantage of the sovereign case is however the availability of various financial instruments conveying information regarding the market perception of this risk along both time and maturity dimensions. In the aftermath of the euro-area sovereign debt crisis, several studies have documented the correlation of sovereign credit spreads (see Beirne and Fratzscher, 2013; Ludwig, 2014; Lucas et al., 2014; Caporin et al., 2018). Nevertheless, very few studies have explicitly considered frailty or contagion mechanisms to account for the dynamics of the whole term structure of sovereign spreads. And when they do, these studies consider only one or the other phenomenon, but not both at the same time. Typically, while Ait-Sahalia et al. (2014) allow for contagion but not frailty in their no-arbitrage model; the inverse holds true for Ang and Longstaff (2013).

Mechanically, the values of the maximum log-likelihoods obtained with models featuring a frailty factor – Models (1), (2), (3) and (4) – are substantially larger than those associated with models with no frailty – respectively Models (5), (6), (7) and (8). Unfortunately, because the parameters of the dynamics of $z_t$ are not identifiable under the null hypothesis of no frailty, the distribution of the likelihood-ratio test statistic is non-standard in this situation (see e.g. Hansen, 1992). By contrast, we can test whether the contagion parameters, that are the $c_i$'s and the $\boldsymbol{\kappa}_i$'s are jointly

[28] Both types of channel have been mentioned as potential sources of sovereign default clustering. Longstaff et al. (2011) suggest that co-movements in sovereign credit risk may reflect a strong influence of global macroeconomic factors, which is rather suggestive of the frailty mechanism. Bai et al. (2012) put the emphasis on feedback mechanisms between credit and liquidity risks that may give rise to contagious spillover effects among sovereign entities. Benzoni et al. (2015) explore an alternative contagion-like mechanism, whereby agents updating fragile beliefs they have about the state of the economy.

statistically significant by employing likelihood-based tests. The test statistic being well beyond the critical values at any confidence level, we are led to reject the hypothesis of no contagion pricing.

Notwithstanding this econometric evidence, what are the economic implications of contagion? To address this question, we resort to the following exercise. For each of the models featuring contagion – Models (1), (3), (5) and (7) – we compute the average decreases in spreads obtained after killing contagion effects – setting $c_i$ parameters to zero – all else being equal. For Models (1), (3), (5) and (7), we get average decreases of respectively 7%, 11%, 54% and 50%. Recalling that the last two models feature no frailty component, these results suggest that the (economic) importance of the contagion channel critically depends on the existence of a frailty component in the model.

### 4.8 Sovereign Defaults and Exchange Rates

The previous results are based on euro-denominated CDSs. However, CDS protection on many international corporations and on sovereign entities are available in euros and in U.S. dollars. (Data on both types of CDSs are collected by Thomson Reuters Datastream.) While most of European sovereign bonds are denominated in euros, a large share of European CDS are denominated in dollars. This is because the latter provides a better protection against a potential severe depreciation of the bond's currency in the case of a sovereign credit event (see Fontana and Scheicher, 2010; Augustin et al., 2020).[29] Here is the rationale behind the previous statement: the notional of a euro-denominated CDS is fixed in euros and that of a dollar-denominated CDS is fixed in dollars. Therefore, a euro depreciation leads to an increase of the notional of the dollar-denominated CDS expressed in euros. Formally, consider two CDS negotiated at date $t$: the first is a maturity-$h$ euro-denominated CDS and the second is a dollar-denominated one with the same maturity. At inception, we consider that both CDS have identical face values, say $N$ euros for the former and $N\exp(-s_t)$ dollars for the latter – denoting by $s_t$ the log of the exchange rate between the domestic and the foreign currency. Assume that entity $i$ defaults before the maturity of the contract and

[29]The study of the potential liquidity differences between euro-denominated and dollar-denominated bonds, mentioned e.g. in Credit Suisse (2010), is beyond the scope of this paper.

that the default triggers a euro depreciation: $s_{\tau_i} - s_t > 0$. Then, the payoff of the protection leg is higher for the dollar-denominated CDS than for the euro-denominated one. Indeed, we have $N(1 - \varrho_{i,\tau_i})\exp(s_{\tau_i} - s_t) > N(1 - \varrho_{i,\tau_i})$. ($\varrho_{i,\tau_i}$ is the recovery rate defined in Assumption 7.) Therefore, if the default by a euro-area member state is expected to be accompanied by a euro depreciation, dollar-denominated CDSs should have higher spreads than euro-denominated ones. The data are consistent with this view: the quanto CDS spreads„ defined by the deviation between a dollar-denominated CDS spread and a euro-denominated one, are mostly positive (see crosses in Figure 4).

In the following, we show that, once the previous model is augmented with the EURUSD exchange rate, it can capture the main fluctuations of the term structure of the quanto CDS spreads for the countries into consideration.

We assume that $\Delta s_t$, the one-period change in the logarithm of the real exchange rate, is given by:[30]

$$\Delta s_t = \chi + v_t + u_\delta' \delta_t, \tag{27}$$

where $v_t$ is an additional autonomous component of $y_t$, namely $y_t = (r_t, z_t, x_t', v_t)'$. If the elements of $u_\delta$ are positive, a sovereign default implies a depreciation of the euro with respect to the U.S. dollar.[31] We posit a Gamma distribution for the (i.i.d.) $v_t$ shocks; the associated scale and shape parameters are determined in such a way as to match the sample mean and variance of observed changes in real exchange rate.[32]

Estimated values of the elements of $u_\delta$ are obtained by minimizing a weighted sum of squared deviations between the observed and the model-implied quanto CDS spreads. The weights are the inverses of the sample means of squared quanto CDS spreads. (There is one weight for each

[30] Note that this specification implies that the logarithm of the real exchange rate is integrated of order one, which is consistent with the results of unit root tests carried out on exchange rate series.

[31] Because $v_t \geq 0$, $\chi$ has to be negative enough to allow for possible large euro appreciation (assuming the elements of $u_\delta$ are non-negative). We set $\chi = -0.5$. This implies that the lowest possible change in the real exchange rate is of about 40% (in one month), which seems to constitute a reasonable lower bound. The results are insensitive to the value of this parameter. For instance, replacing this value by $-1$ or $-0.10$ yields virtually identical results.

[32] Since $\Delta s_t$ and $\delta_t$ are observed, $v_t$ shocks are available.

country-maturity pair.)

According to the results, on average, sovereign defaults in France, Germany, Greece, Italy and Spain would respectively trigger euro depreciations of 15%, 20%, 0%, 6% and 8%.[33] As expected, these results suggest in particular that defaults by France and Germany, the two largest economies of this reduced euro-area, would have stronger impacts on the EURUSD exchange rate. Figure 4 compares observed and model-implied quanto CDS spreads. Let us stress that this result is obtained without introducing a novel latent variable ($v_t$ is indeed observed for $t \in [1, T]$). Except for Greece, the fit is surprisingly good. For the latter country, note that quanto CDSs correspond to a small part of observed CDSs (around 5%), and one may suspect that liquidity issues outweigh the exchange-rate-related spread differentials. Leaving Greece aside, this simple model extension accounts for two thirds of the variances of observed quanto CDS spreads, on average across countries and maturities.

Figure 4 also displays the (model-implied) quanto CDS spreads that would be observed if agents were risk-neutral. These spreads are obtained by applying the pricing formulas of Proposition 3.4 under the physical measure or, equivalently, after having set the prices of risk to zero. The differences between $\mathbb{Q}$ and $\mathbb{P}$ quanto CDS spreads can be interpreted as risk premia. Our results indicate that risk premia account for an important share of total quanto CDS spreads, especially for long maturities. This is consistent with the fact that quanto CDSs provide positive payoffs to the protection buyer in particularly bad states of the world (sovereign defaults).

Equation (27) assumes that credit risk affects the exchange rate through the credit-event variables only. We have considered a more general specification where $\Delta s_t$ is also allowed to depend on $z_t$ and $x_t$. That is, a term $u_z z_t + u_x' x_t$ is added on the right-hand side of Equation (27). This augmented flexibility hardly allows for an improvement of the fit. In addition, we have considered a specification where the term $u_z z_t + u_x' x_t$ is maintained in the specification of $\Delta s_t$, but where the

[33] These effects are deduced as follows from the components of $u_\delta$ (denoted by $u_{\delta,i}$). To begin with, note that as long as the intensity $\lambda^{\mathbb{P}}_{i,t}$ is small, conditional on having a default by country $i$ on date $t$ (i.e. conditional on $\mathsf{P}_{i,t} > 0$), the probability of having $\mathsf{P}_{i,t} = 1$ is close to one. It therefore comes that the distribution of $\delta_{i,t}$ on a default date is approximately $\Gamma_1(0.6)$ since $\mu_{\delta,i} = 0.6$ (see Subsection 4.4). The expected depreciation upon default is the expectation of $\exp(s_t) - 1$ conditional on $\delta_{i,t} > 0$. Given what precedes, we approximate this by the expectation of $\exp(u_{\delta,i}X) - 1$ where $X \sim \Gamma_1(0.6)$; this conditional expectation directly results from the knowledge of Laplace transform of a gamma distribution.

credit-event variables are removed (i.e. $u_\delta = 0$). The fit resulting from this alternative specification is very poor.[34] Altogether, these results suggest that it is the relationship between the exchange rate and the credit events *per se*, and less between the exchange rate and conditional default probabilities – driven by $(z_t, x_t')'$ – that is key to explain the fluctuation of quanto CDS spreads. These results are in line with those of Ehlers and Schonbucher (2004) and of Brigo et al. (2015).

# 5 Conclusion

We present a general affine positive credit-risk model able to simultaneously relax restrictive assumptions often employed in the reduced-form credit-risk literature while preserving tractability in the pricing of default-sensitive securities. Building on the recent non-negative affine Gamma-zero process, the model accommodates the presence of systemic risk (i.e. potential feedbacks from defaults towards common risk factors), contagion between entities, credit-event pricing (arising when the SDF directly depends on default events) and stochastic recovery rates. We provide explicit formulas to price defaultable securities such as defaultable bonds and CDS, for different recovery-rate conventions.

We exploit this framework to investigate the pricing of sovereign credit risk using euro-area data. We show that one common factor and one country-specific factor for each country allow for a very good fit of CDS data. The estimation detects contagion effects, even when allowing for a frailty factor. Moreover, we find sizable credit-risk premia along the whole maturity spectrum. Typically, credit-risk premia account for more than half of CDS spreads at the five year maturity for France, Germany, Italy and Spain. Our findings also highlight the importance of credit-event pricing to allow for non-trivial short-term credit-risk premia. A simple extension of the model finally allows us to extract measures of expected (EURUSD) depreciations-at-default by jointly modeling term structures of sovereign CDSs denominated in euros and in U.S. dollars.

[34] For our the most complete specification (Model (1)), the ratios of mean squared pricing errors to mean squared quanto CDS spreads are of 23%, 21%, 14% and 12% on average across maturities for Germany, France, Italy and Spain, respectively. For the specification where credit-event variables cannot affect the exchange rate, the same ratios are 72%, 70%, 57% and 62%, respectively.

## References


Ait-Sahalia, Y., J. Cacho-Diaz, and R. J. A. Laeven (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics 117*, 585–606.

Ait-Sahalia, Y., R. J. A. Laeven, and L. Pelizzon (2014). Mutual excitation in eurozone sovereign CDS. *Journal of Econometrics 183*, 151–167.

Almeida, H. and T. Philippon (2007). The risk-adjusted cost of financial distress. *Journal of Finance 62*, 2557–2586.

Altman, E., B. Brady, A. Resti, and A. Sironi (2005). The link between default and recovery rates: Theory, empirical evidence, and implications. *The Journal of Business 78*(6).

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.

Ang, A. and F. Longstaff (2013). Systemic sovereign credit risk: Lessons from the U.S. and Europe. *Journal of Monetary Economics 60*(5).

Asonuma, T. and C. Trebesch (2016). Sovereign debt restructurings: Preemptive or post-default. *Journal of the European Economic Association 14*(1), 175–214.

Augustin, P., M. Chernov, and D. Song (2020). Sovereign credit risk and exchange rates: Evidence from CDS quanto spreads. *Journal of Financial Economics*.

Azizpour, S., K. Giesecke, and G. Schwenkler (2018). Exploring the sources of default clustering. *Journal of Financial Economics 129*(1), 154–183.

Bai, J., P. Collin-Dufresne, R. Goldstein, and J. Helwege (2015). On bounding credit-event risk premia. *Review of Financial Studies 28*(9), 2608–2642.

Bai, J., C. Julliard, and K. Yuan (2012). Eurozone sovereign bond crisis: Liquidity or fundamental contagion. Unpublished working paper, Federal Reserve Bank of New York.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Barro, R. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics 121*(3), 823–866.

Beirne, J. and M. Fratzscher (2013). The pricing of sovereign risk and contagion during the European sovereign debt crisis. *Journal of International Money and Finance 34*(C), 60–82.

Benzoni, L., P. Collin-Dufresne, R. Goldstein, and J. Helwege (2015). Modeling credit contagion via the updating of fragile beliefs. *Review of Financial Studies 28*(7), 1960–2008.

Brennan, M. and E. Schwartz (1980). Analyzing convertible bonds. *Journal of Financial and Quantitative Analysis 15*, 907–929.

Brigo, D., N. Pede, and A. Petrelli (2015). Multi currency credit default swaps: quanto effects and FX devaluation jumps. Working paper, Imperial College London.

Caporin, M., L. Pelizzon, F. Ravazzolo, and R. Rigobon (2018). Measuring sovereign contagion in Europe. *Journal of Financial Stability 34*(C), 150–181.

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, L. and D. Filipovic (2007). Credit derivatives in an affine framework. *Asia-Pacific Financial Markets 14*, 123–140.

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.

Cochrane, J. H. and J. Saa-Requejo (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy 108*(1), 79–119.

Collin-Dufresne, P., R. Goldstein, and J. Hugonnier (2004). A General formula for valuing defaultable securities. *Econometrica 72*(5), 1377–1407.

Coudert, V. and M. Gex (2013). Why the Greek CDS settlement did not lead to the feared meltdown. *Financial Stability Review* (17), 135–150.

Creal, D. and C. Wu (2015). Estimation of affine term structure models with spanned or unspanned stochastic volatility. *Journal of Econometrics 185*(2), 60–81.

Credit Suisse (2010). *Sovereign CDS Primer* (Fixed Income Research ed.).

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

D'Amato, J. and E. M. Remolona (2003). The credit spread puzzle. *BIS Quarterly Review*.

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

Das, Sanjiv R.and Hanouna, P. (2009). Implied recovery. *Journal of Economic Dynamics and Control 33*(11), 1837–1857.

Das, S. R., D. Duffie, N. Kapadia, and L. Saita (2007). Common failings: How corporate defaults are correlated. *Journal of Finance 62*(1), 93–117.

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

Driessen, J. (2005). Is the default event risk priced in corporate bonds? *Review of Financial Studies 18*(1), 165–195.

Duan, J.-C. and W. Miao (2016). Default correlations and large-portfolio credit analysis. *Journal of Business & Economic Statistics 34*(4), 536–546.

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, April). Sharpe ratios in term structure models. Economics Working Paper Archive 575, The Johns Hopkins University,Department of Economics.

Duffie, D. (1998). Defaultable term structure models with fractional recovery of par. Working paper, Graduate School of Business, Stanford University.

Duffie, D. (2001). *Dynamic asset pricing theory* (3. ed ed.). Princeton, NJ: Princeton Univ. Press.

Duffie, D. (2005). Credit risk modeling with affine processes. *Journal of Banking and Finance 29*.

Duffie, D., A. Eckner, G. Horel, and L. Saita (2009). Frailty correlated default. *Journal of Finance 64*(5), 2089–2123.

Duffie, D., M. Schroder, and C. Skiadas (1996). Recursive valuation of the defaultable securities and the timing of resolution of uncertainty. *The Annals of Applied Probability 6*(4), 1075–1090.

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

Ehlers, P. and P. Schonbucher (2004). The influence of FX risk on credit spreads. Working paper series, ETH Zurich.

Epstein, L. G. and S. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica 57*(4), 937–969.

Eraker, B. (2008). Affine general equilibrium models. *Management Science 54*(12), 2068–2080.

Fontana, A. and M. Scheicher (2010). An analysis of euro area sovereign CDS and their relation with government bonds. *Journal of Banking and Finance 62*, 126–140.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *The Quarterly Journal of Economics 127*(2), 645–700.

Giesecke, K., F. Longstaff, S. Schaefer, and I. Strebulaev (2011). Corporate bond default risk: A 150-year perspective. *Journal of Financial Economics 102*(2), 233–250.

Giesecke, K. and S. Zhu (2013). Transform analysis for point processes and applications in credit risk. *Mathematical Finance 23*(4), 742–762.

Gourieroux, C. and J. Jasiak (2006). Autoregressive gamma processes. *Journal of Forecasting 25*(2), 129–152.

Gourieroux, C., A. Monfort, and J.-P. Renne (2014). Pricing default events: Suprise, exogeneity and contagion. *Journal of Econometrics 182*, 397.

Guo, X., R. A. Jarrow, and Y. Zeng (2009). Modeling the recovery rate in a reduced form model. *Mathematical Finance 19*(1), 73.

Hansen, B. E. (1992). The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP. *Journal of Applied Econometrics 7*(S), 61–82.

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.

Hansen, L. P. and S. F. Richard (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica 55*(3), 587–613.

Huang, J.-Z. and M. Huang (2012). How much of the corporate Treasury yield spread is due to credit risk? *The Review of Asset Pricing Studies 2*(2), 153–202.

Jarrow, R. A., D. Lando, and F. Yu (2005). Default risk and diversification: Theory and empirical implications. *Mathematical Finance 15*(1), 1–26.

Jarrow, R. A. and S. Turnbull (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance 50*(1), 53–86.

Jarrow, R. A. and F. Yu (2001). Counterparty risk and the pricing of defaultable securities. *Journal of Finance 56*(5), 1765–1799.

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.

Kraft, H. and M. Steffensen (2007). Bankruptcy, counterparty risk, and contagion. *Review of Finance 11*, 209–252.

Lando, D. (1998). On Cox processes and credit risky securities. *The Review of Derivatives Research 2*, 99–120.

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

Longstaff, F. and E. Schwartz (1995). A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance 50*(3), 789–821.

Lucas, A., B. Schwaab, and X. Zhang (2014). Conditional euro area sovereign default risk. *Journal of Business & Economic Statistics 32*(2).

Ludwig, A. (2014). A unified approach to investigate pure and wake-up-call contagion: Evidence from the Eurozone's first financial crisis. *Journal of International Money and Finance 48*, 125–146.

Mendoza, E. G. and V. Z. Yue (2012). A general equilibrium model of sovereign default and business cycles. *The Quarterly Journal of Economics 127*(2), 889–946.

Monfort, A., F. Pegoraro, J.-P. Renne, and G. Roussellet (2017, December). Staying at zero with affine processes: An application to term structure modelling. *Journal of Econometrics 201*(2), 348–366.

Monfort, A., F. Pegoraro, J.-P. Renne, and G. Roussellet (2018). Affine modelling of credit risk, pricing of credit events and contagion. Working paper, CREST.

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

Na, S., S. Schmitt-Grohe, M. Uribe, and V. Z. Yue (2018). The twin Ds: Optimal default and devaluation. *American Economic Review 108*(7), 1773–1819.

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

Piazzesi, M. (2010). Affine term structure models. In *Handbook of Financial Econometrics, Volume 1* (Yacine Ait-Sahalia and Lars Peter Hansen North Holland ed.)., Chapter 12, pp. 691–766.

Reinhart, C. M. (2002). Default, currency crises, and sovereign credit ratings. *World Bank Economic Review 16*(2), 151–170.

Reinhart, C. M. and K. S. Rogoff (2011). The forgotten history of domestic debt. *Economic Journal 121*(552), 319–350.

Roussellet, G. (2019). The term structure of macroeconomic risk at the zero lower bound.

Trebesch, C. and M. Zabel (2017). The output costs of hard and soft sovereign default. *European Economic Review 92*, 416–432.

# Tables and Figures

Table 1: Summary of model parameterizations

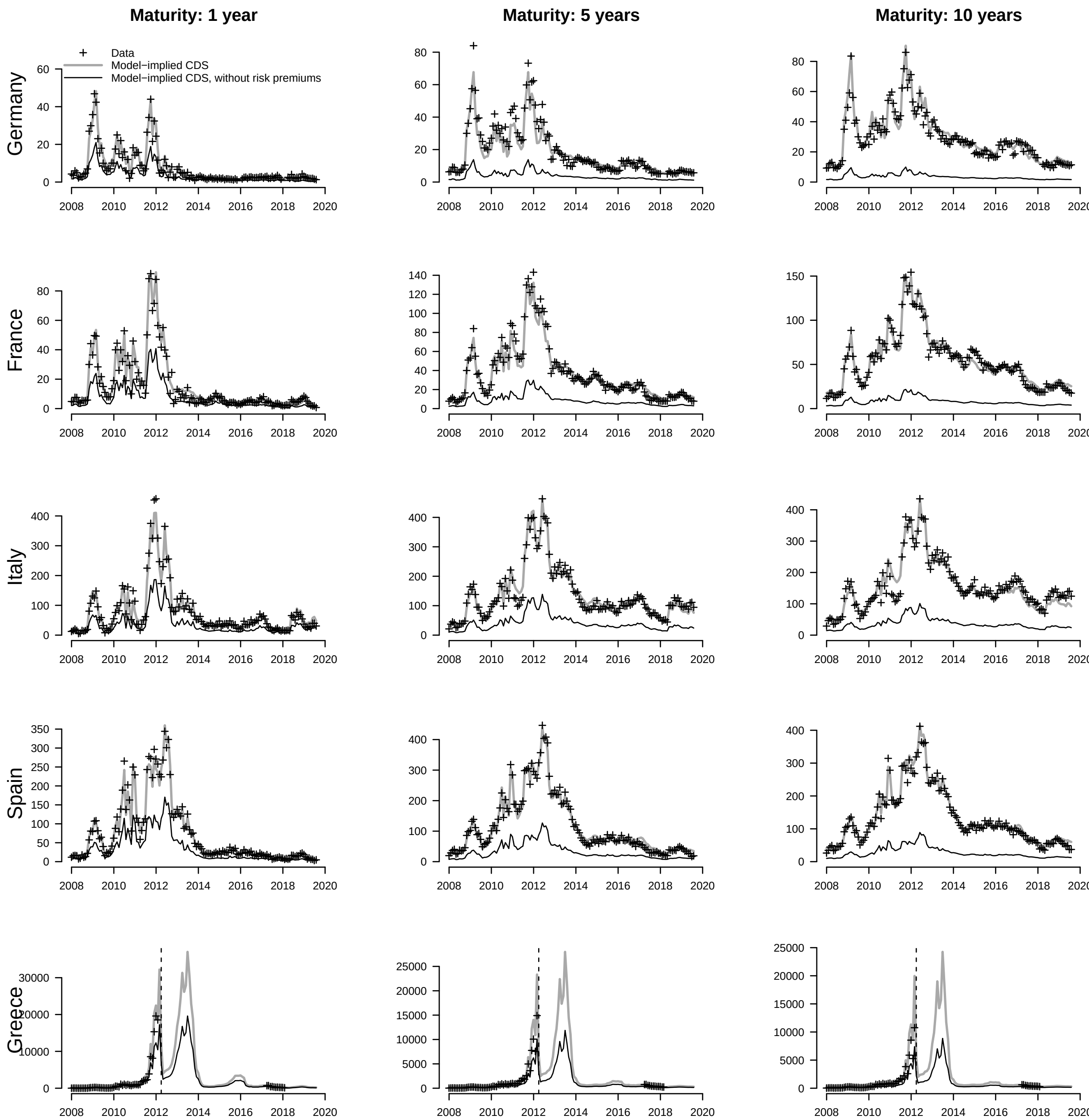
| Model | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Frailty | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ |
| Contagion | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ |
| Credit-event pricing | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ |

Table 2: Shares of CDS spreads corresponding to credit-risk premia

| Model | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | **Panel A – Horizon = 6 months** | | | | | | | |
| Germany | **0.510** | **0.409** | 0.118 | 0.158 | **0.289** | **0.587** | 0.086 | 0.122 |
| France | **0.518** | **0.409** | 0.112 | 0.158 | **0.268** | **0.571** | 0.068 | 0.125 |
| Italy | **0.503** | **0.409** | 0.110 | 0.157 | **0.209** | **0.558** | 0.067 | 0.130 |
| Spain | **0.502** | **0.409** | 0.105 | 0.157 | **0.208** | **0.540** | 0.062 | 0.128 |
| Greece | **0.388** | **0.400** | 0.087 | 0.141 | **0.142** | **0.458** | 0.043 | 0.109 |
| | **Panel B – Horizon = 60 months** | | | | | | | |
| Germany | **0.784** | **0.714** | 0.614 | 0.729 | **0.752** | **0.733** | 0.542 | 0.606 |
| France | **0.767** | **0.701** | 0.585 | 0.716 | **0.727** | **0.714** | 0.510 | 0.603 |
| Italy | **0.701** | **0.637** | 0.489 | 0.642 | **0.670** | **0.671** | 0.491 | 0.548 |
| Spain | **0.724** | **0.661** | 0.520 | 0.667 | **0.656** | **0.670** | 0.467 | 0.584 |
| Greece | **0.466** | **0.498** | 0.260 | 0.425 | **0.475** | **0.452** | 0.353 | 0.220 |

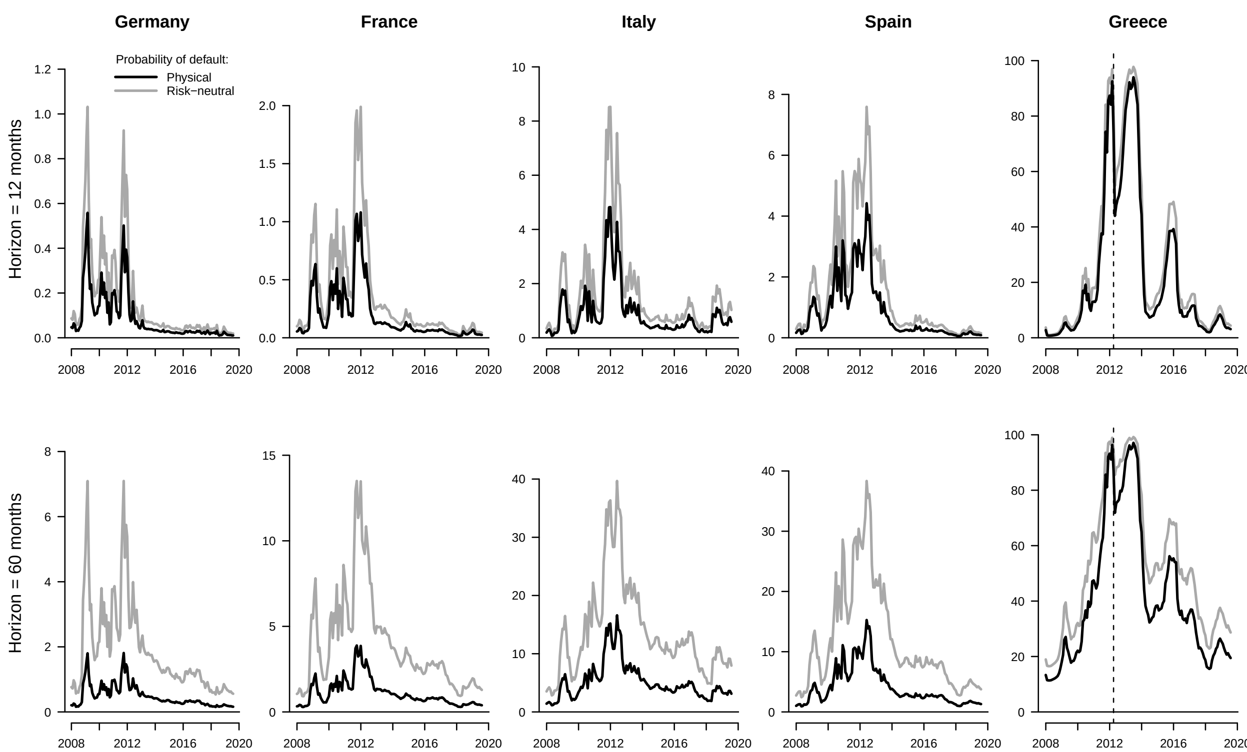
*Note*: This table reports the sample averages of the shares of CDS spreads corresponding to credit-risk premia. Credit-risk premia are defined as the difference between model-implied CDS spreads and the counterfactual CDS spreads obtained after having set the prices of risk to zero (the prices of risk are the components of $\theta_w$, see Equation 26). Figures in bold font are for models where $\mathbf{S} > 0$.

Figure 1: Observed vs model-implied CDSs



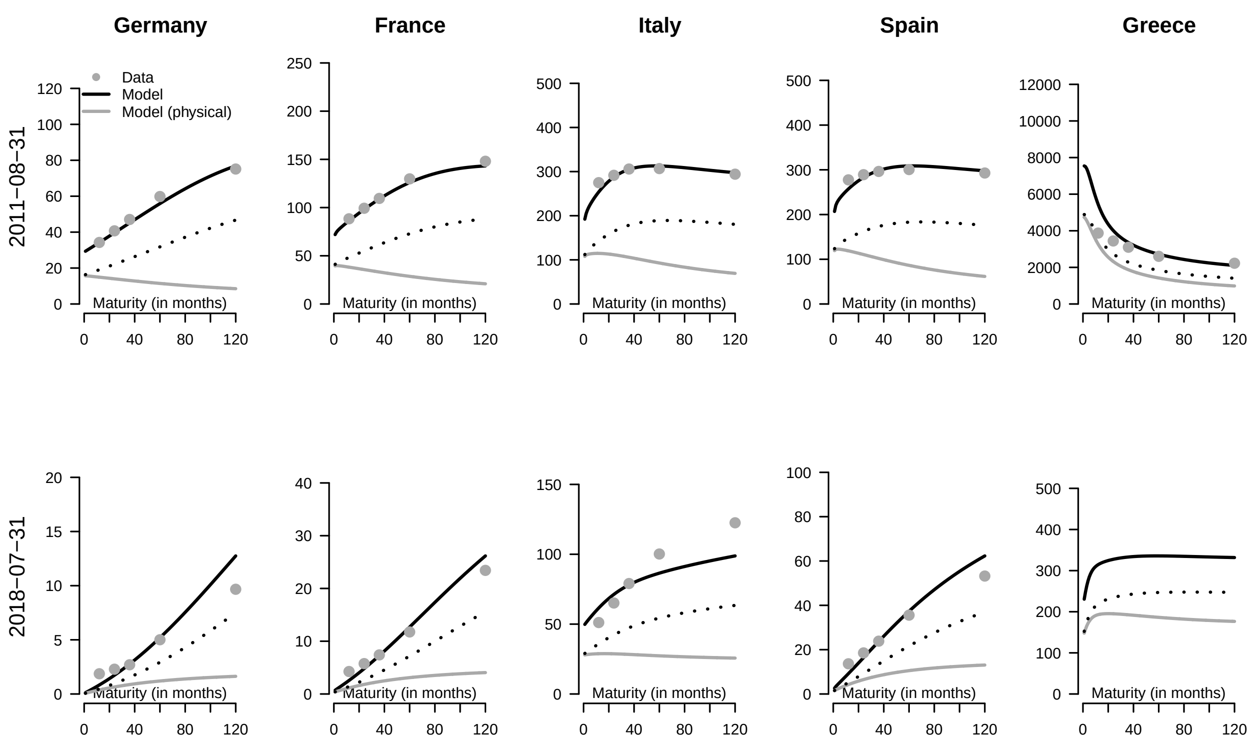

*Note:* The gray lines correspond to the model-implied CDS spreads, expressed in basis points. The data span the period from January 2007 to July 2019 at the monthly frequency. The thin black line corresponds to (model-implied) $\mathbb{P}$ CDS spreads, that are the spreads that would be observed if agents were not risk averse. The $\mathbb{P}$ CDS spreads are obtained by applying the CDS pricing formulas after having set the prices of risk ($\theta_x$, $\theta_y$, $\theta_r$ and $\mathbf{S}$) to zero. For Greece: the vertical dashed bar indicates the default period (March 2012).

Figure 2: Model-implied probabilities of default



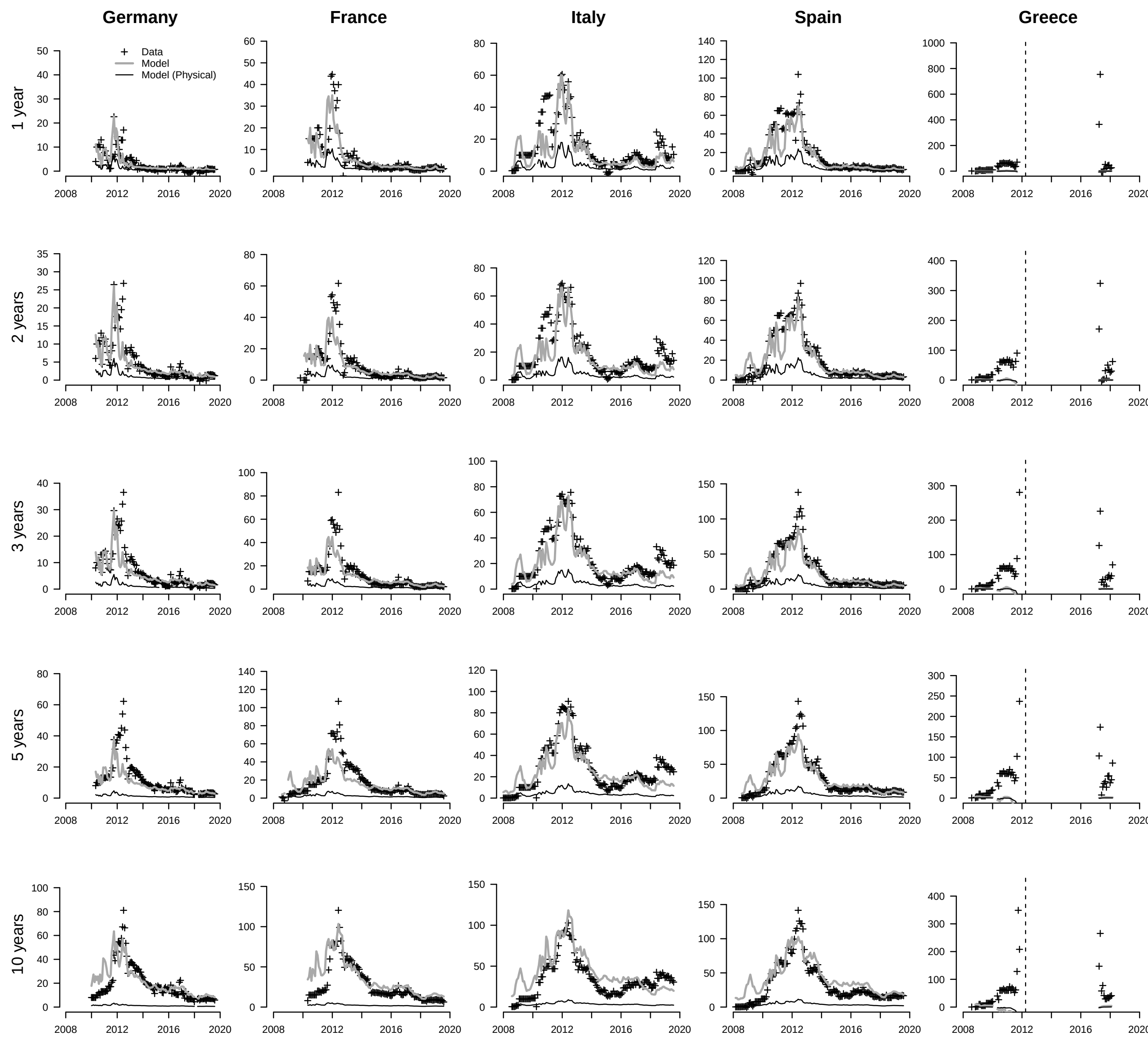

*Note:* This figure displays model-implied conditional probabilities of default for two horizons: 12 months and 60 months. The grey (respectively black) line corresponds to the physical (respectively risk-neutral) probability of default. The difference between the two curves reflects credit-risk premia. For Greece: the vertical dashed bar indicates the default period (March 2012).

Figure 3: Model-implied term structures of probabilities of default




*Note:* This figure displays model-implied term structures of CDS spreads (solid black lines) together with observed CDS spreads (grey dots) for two dates (under the complete model, i.e. Model (1) in Table A.10). The solid grey line represents the CDS spreads that would prevail if agents were risk-neutral or, equivalently, if the prices of risk ($\theta_w$, see Equation 26) were null. The spread differentials between the solid black line and the grey solid line therefore correspond to credit-risk premia. The dashed line corresponds to the CDS spreads that would prevail – everything else equal – if the credit-event pricing parameter $\mathbf{S}$ was equal to zero. The figure shows in particular that short-term credit-risk premia are small in the latter case.

Figure 4: Quanto CDS




*Note:* This figure compares observed and model-implied quanto CDS spreads (expressed in basis points). quanto CDS spreads are defined as the differences between dollar-denominated CDS premia and their euro-denominated counterparts. For some countries and maturities, Datastream-extracted CDS premia are the same for the euro- and dollar-denominated CDS; in these cases, the data are removed from the estimation sample. Data points are also removed when CDS premia do not change for three months in a row (which indicates illiquidity). The thin solid line corresponds to the (model-implied) quanto CDS spreads that would be observed if agents were risk-neutral. For Greece: the vertical dashed bar indicates the default period (March 2012).

# Online Appendix (not for publication)

**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:

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

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

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

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:

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

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:

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

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

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

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:

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

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

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

we obtain:

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

such that:

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

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 4.8).

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:

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

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:

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

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:

$$\Delta s_t = \chi + u_s' w_t\,. \tag{a.10}$$

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:*

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

*where:*

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

*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:

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

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:*

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

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

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

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

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

(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:*

$$
w_t = M_0 + M_1 w_{t-1} + \Sigma(w_{t-1})\xi_t, \tag{a.15}
$$

*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:*

$$
\left[\begin{array}{cc}
diag\left(M_2 + M_3 w_{t-1}\right) & diag\left(M_2 + M_3 w_{t-1}\right) M_4' \\
M_4 diag\left(M_2 + M_3 w_{t-1}\right) & M_4\, diag\left(M_2 + M_3 w_{t-1}\right) M_4' + diag\left(M_5 + M_6 w_{t-1}\right)
\end{array}\right],
$$

*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) \Big| \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] \Big| \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:

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

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

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

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

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

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

■

## 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 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:

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

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:

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

Last, the one-period SDF is given by:

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

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

### A.3.2 Calibration of the Illustrative Example

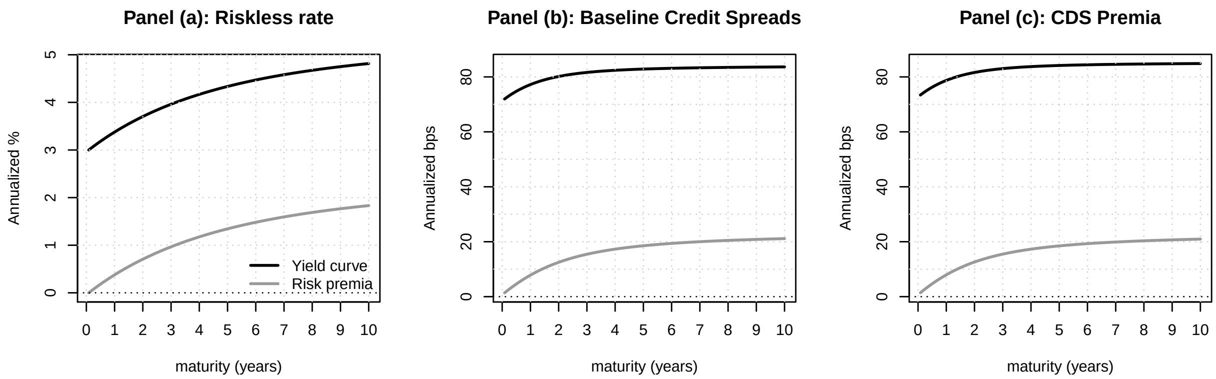
Our baseline model's calibration is at the monthly frequency and is presented in Table A.1. In order to avoid any discrepancies between recovery conventions, we impose that the recovery rate is zero ($\mu_\delta = 50$, the RR being defined by Equation 15). In this case, CDS spreads are virtually indistinguishable 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

| $\delta_t$ | | $y_t$ | | $r_t$ | | $M_{t-1,t}$ | |
|---|---|---|---|---|---|---|---|
| $\beta_\lambda^{(y)}$ | $5 \cdot 10^{-4}$ | $\beta_y^{(y)}$ | 0.95 | $\beta_r$ | 118,172.6 | $\theta_r$ | 0.05 |
| $\mu_\delta$ | 50 | $\mu_y$ | 1 | $\mu_r$ | $8.21 \cdot 10^{-6}$ | $\theta_y$ | 0.01 |
| $\mathbf{C}$ | 0 | $\nu_y$ | 0.06 | $\alpha_r$ | 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 $\delta_t$ 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.

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 $\delta_t$ are null so no defaults have happened.

### A.3.3 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 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 $\mathbf{S} = 3.5371 \cdot 10^{-3}$.

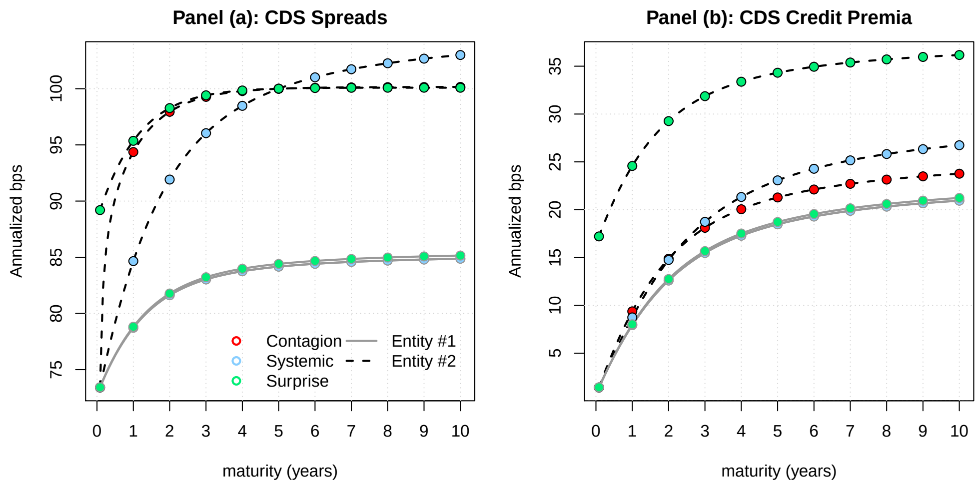
The resulting yield curves are presented on Figure A.2. The solid grey lines present the results obtained for entity 1, which are virtually insignificant compared to the baseline. In contrast, the contagion, systemic and surprise scenarios propose three distinct term structures of CDS spreads on the 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.3.4 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.[35]

[35] We use the same shocks across scenarios for the simulation of $\mathsf{P}_{\delta,t}$, $\delta_t$, $\mathsf{P}_{y,t}$, $y_t$, $\mathsf{P}_{r,t}$ and $r_t$. Since Gamma processes are conditionally heteroskedastic weak AR processes, we simulate uniform distributions and use inverse cumulative

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 $\delta_t$ 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.

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%.[36]

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

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.

[36]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 $\delta_t$* | | | | | |
|---|---|---|---|---|---|---|
| | 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 $\delta_t$ 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.

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* 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.

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.

## 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 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 $\delta_t$ is observable in real-time. She has also access to the term structures of bond credit spreads $\{\mathrm{CS}_i(t,h)\}_{h\in H_i}$, where $H_i$ is the discrete set of observable maturities and $i$ refers to the defaultable entities.[37] More precisely, for any entity $i$ and any maturity $h$, the bond 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 $\rho_\delta := \mu_\delta \cdot \beta_\lambda^{(y)}$.[38] These measurement equations are accompanied with transition equations defining the VARG joint dynamics of $y_t$ and $\delta_t$ as functions of $\Theta$, 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:

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

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

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 $w_t$:

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

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

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

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

[37] 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 $(y_t, \delta_t)$. This forces the econometrician to use a non-linear filter as the extended Kalman filter. We adopt such a procedure in our real-data application in Section 4.

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

**Filtering-based Estimation:** The most standard approach for estimating term structure models with unobserved factors is based on approximate Kalman filtering (see e.g. de Jong, 2000). Compared with GMM-based methods, these methods allow to estimate the parameters and back out $y_t$ at the 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: $\zeta_t$ 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)* $\zeta_t$ 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).[39] 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 described 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 $\delta_t$ that are observable to the econometrician. We are looking for the one-period conditional log-likelihood function $\mathcal{L}\left(\text{Obs}_t \middle| \underline{\text{Obs}_{t-1}}\right)$. We also denote by $\text{Obs}_t^*$ 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 $\text{Obs}_t = g_t\left(\text{Obs}_t^*, y_t\right)$.

[39] 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.

The conditional quasi log-likelihood can be written in terms of both $\mathrm{Obs}_t^*$ and $y_t$:

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

For all dimensions but one, the function $g_t$ is equal to identity since it transforms elements of $\mathrm{Obs}_t$ 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:

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

The first term of the log-likelihood represents the joint Gaussian distribution of the measurement errors $\varepsilon_t$ and $\eta_t$. 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).[40]

**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.[41] 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 $\Theta$:

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

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:

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

[40] 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.

[41] 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.

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

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 $\Theta$. 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 $\delta_t$ 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 $\delta_t$ 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 $\delta_t$, 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 $\delta_t$. 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 $\delta_t$ 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 $\delta_t$ 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 $\rho_\delta$ 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 $\delta_t$ 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 $\delta_t$ 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 $\delta_t$ 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 $\delta_t$ in the set of observable

| | | | $\nu_y$ | $\beta_y$ | $\rho_\delta$ | $\theta_y$ | $\mathbf{I}$ | $\mathbf{C} \cdot 10^3$ | $\mathbf{S} \cdot 10^3$ | $\sigma_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 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 $\delta_t$ in the set of observable

| | | | $\nu_y$ | $\beta_y$ | $\rho_\delta$ | $\theta_y$ | $\mathbf{I}$ | $\mathbf{C} \cdot 10^3$ | $\mathbf{S} \cdot 10^3$ | $\sigma_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 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 $\delta_t$ in the set of observable

| | | | $\nu_y$ | $\beta_y$ | $\rho_\delta$ | $\theta_y$ | $\mathbf{I}$ | $\mathbf{C} \cdot 10^3$ | $\mathbf{S} \cdot 10^3$ | $\sigma_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 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 $\delta_t$ in the set of observable

| | | | $\nu_y$ | $\beta_y$ | $\rho_\delta$ | $\theta_y$ | $\mathbf{I}$ | $\mathbf{C}\cdot 10^3$ | $\mathbf{S}\cdot 10^3$ | $\sigma_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 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 $\delta_t$, 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/ $\delta_t$) | 143.63 | 120.68 | 53.96 | 110.7 | 326.35 |
| | | | (w/ $\delta_t$) | 535.29 | 630.11 | 80.79 | 246.28 | 1985.27 |
| | | $\widehat{y}_t - y_t$ | (wo/ $\delta_t$) | 0.018 | 0.118 | -0.071 | 0.002 | 0.169 |
| | | | (w/ $\delta_t$) | -0.047 | 0.271 | -0.499 | 0.001 | 0.189 |
| | | $\widehat{\delta}_{1,t} - \delta_{1,t}$ | (wo/ $\delta_t$) | 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/ $\delta_t$) | 103.84 | 94.95 | 49.09 | 79.53 | 237.41 |
| | | | (w/ $\delta_t$) | 334.72 | 475.33 | 54.49 | 164.83 | 1516.16 |
| | | $\widehat{y}_t - y_t$ | (wo/ $\delta_t$) | -0.01 | 0.113 | -0.177 | 0.001 | 0.12 |
| | | | (w/ $\delta_t$) | -0.262 | 1.604 | -2.948 | 0.025 | 1.459 |
| | | $\widehat{\delta}_{1,t} - \delta_{1,t}$ | (wo/ $\delta_t$) | -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/ $\delta_t$) | 175.49 | 199.65 | 65.72 | 131.08 | 349.53 |
| | | | (w/ $\delta_t$) | 843.21 | 768.66 | 78.25 | 435.3 | 2217.19 |
| | | $\widehat{y}_t - y_t$ | (wo/ $\delta_t$) | -0.057 | 0.186 | -0.402 | -0.001 | 0.042 |
| | | | (w/ $\delta_t$) | 0.01 | 0.143 | -0.124 | 0.001 | 0.183 |
| | | $\widehat{\delta}_{1,t} - \delta_{1,t}$ | (wo/ $\delta_t$) | 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/ $\delta_t$) | 204.87 | 217.6 | 67.33 | 140.11 | 634.86 |
| | | | (w/ $\delta_t$) | 345.95 | 415.41 | 58.52 | 169.2 | 1284.22 |
| | | $\widehat{y}_t - y_t$ | (wo/ $\delta_t$) | -0.163 | 0.384 | -0.983 | -0.004 | 0.029 |
| | | | (w/ $\delta_t$) | 0.131 | 0.346 | -0.201 | 0.095 | 0.715 |
| | | $\widehat{\delta}_{1,t} - \delta_{1,t}$ | (wo/ $\delta_t$) | 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/ $\delta_t$) | 303.18 | 339.04 | 75.56 | 175.94 | 1106.98 |
| | | | (w/ $\delta_t$) | 597.7 | 641.31 | 69.36 | 296.12 | 1963.39 |
| | | $\widehat{y}_t - y_t$ | (wo/ $\delta_t$) | 0.064 | 0.447 | -0.457 | 0.05 | 0.612 |
| | | | (w/ $\delta_t$) | 0.016 | 0.088 | -0.05 | 0.001 | 0.142 |
| | | $\widehat{\delta}_{1,t} - \delta_{1,t}$ | (wo/ $\delta_t$) | 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/ $\delta_t$) | 224.16 | 331.7 | 58.32 | 114.32 | 824.34 |
| | | | (w/ $\delta_t$) | 281.62 | 394.55 | 48.72 | 145.86 | 1163.66 |
| | | $\widehat{y}_t - y_t$ | (wo/ $\delta_t$) | 0.035 | 1.901 | -2.13 | 0.06 | 1.738 |
| | | | (w/ $\delta_t$) | -0.298 | 3.569 | -5.238 | -0.103 | 4.618 |
| | | $\widehat{\delta}_{1,t} - \delta_{1,t}$ | (wo/ $\delta_t$) | 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/ $\delta_t$) | 177.85 | 184.27 | 67.54 | 129.83 | 387.26 |
| | | | (w/ $\delta_t$) | 1016.89 | 757.9 | 212.24 | 602.95 | 2281.57 |
| | | $\widehat{y}_t - y_t$ | (wo/ $\delta_t$) | 0.005 | 0.068 | -0.063 | 0.001 | 0.085 |
| | | | (w/ $\delta_t$) | 0.026 | 0.139 | -0.068 | 0.002 | 0.226 |
| | | $\widehat{\delta}_{1,t} - \delta_{1,t}$ | (wo/ $\delta_t$) | 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/ $\delta_t$) | 167.14 | 153.83 | 62.26 | 117.63 | 412.91 |
| | | | (w/ $\delta_t$) | 428.1 | 473.33 | 74.17 | 277.57 | 1594.72 |
| | | $\widehat{y}_t - y_t$ | (wo/ $\delta_t$) | -0.012 | 0.071 | -0.132 | 0.001 | 0.06 |
| | | | (w/ $\delta_t$) | 0.058 | 0.463 | -0.539 | 0.045 | 0.712 |
| | | $\widehat{\delta}_{1,t} - \delta_{1,t}$ | (wo/ $\delta_t$) | -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 $\delta_t$ 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/ $\delta_t$) | 295.96 | 230.18 | 80.94 | 222.65 | 778.9 |
| | | (w/ $\delta_t$) | 204.56 | 151.86 | 73.98 | 140.32 | 518.84 |
| | $\delta_t > 0$ | (wo/ $\delta_t$) | 244.57 | 181.95 | 100.49 | 181.28 | 556.46 |
| | | (w/ $\delta_t$) | 138.16 | 79.24 | 73.42 | 121.21 | 250.22 |
| Contagion | $\delta_t = 0$ | (wo/ $\delta_t$) | 259.1 | 210.83 | 63.38 | 209.06 | 614.02 |
| | | (w/ $\delta_t$) | 191.59 | 142.99 | 78.45 | 143.34 | 532.8 |
| | $\delta_t > 0$ | (wo/ $\delta_t$) | 271.07 | 204.78 | 79.78 | 197.42 | 676.03 |
| | | (w/ $\delta_t$) | 142.26 | 72.48 | 75.78 | 127.19 | 272.14 |
| Systemic | $\delta_t = 0$ | (wo/ $\delta_t$) | 220.68 | 157.16 | 56.6 | 189.96 | 559.98 |
| | | (w/ $\delta_t$) | 170.17 | 76.61 | 78.67 | 163.92 | 330.99 |
| | $\delta_t > 0$ | (wo/ $\delta_t$) | 209.96 | 169.93 | 56.57 | 161.95 | 508.72 |
| | | (w/ $\delta_t$) | 141.67 | 59.51 | 75.76 | 137.55 | 239.16 |
| Surprise | $\delta_t = 0$ | (wo/ $\delta_t$) | 252.56 | 208.05 | 60.16 | 208.3 | 599.67 |
| | | (w/ $\delta_t$) | 192.9 | 148.03 | 79.33 | 142.25 | 534.37 |
| | $\delta_t > 0$ | (wo/ $\delta_t$) | 208.51 | 148.09 | 70.6 | 159.27 | 482.75 |
| | | (w/ $\delta_t$) | 143.26 | 69.59 | 78.12 | 130.85 | 257.65 |

*Notes:* In the case where $\delta_t$ 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 $\mu_{\delta_i}$ in the Sovereign Credit Risk Application

Data on 1983-2015 sovereign defaults – and more specifically to the associated recovery rates – are used to calibrate $\mu_{\delta_i}$, 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 $\mu_{\delta_i}$. (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 $\mu_{\delta_i}$. 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).

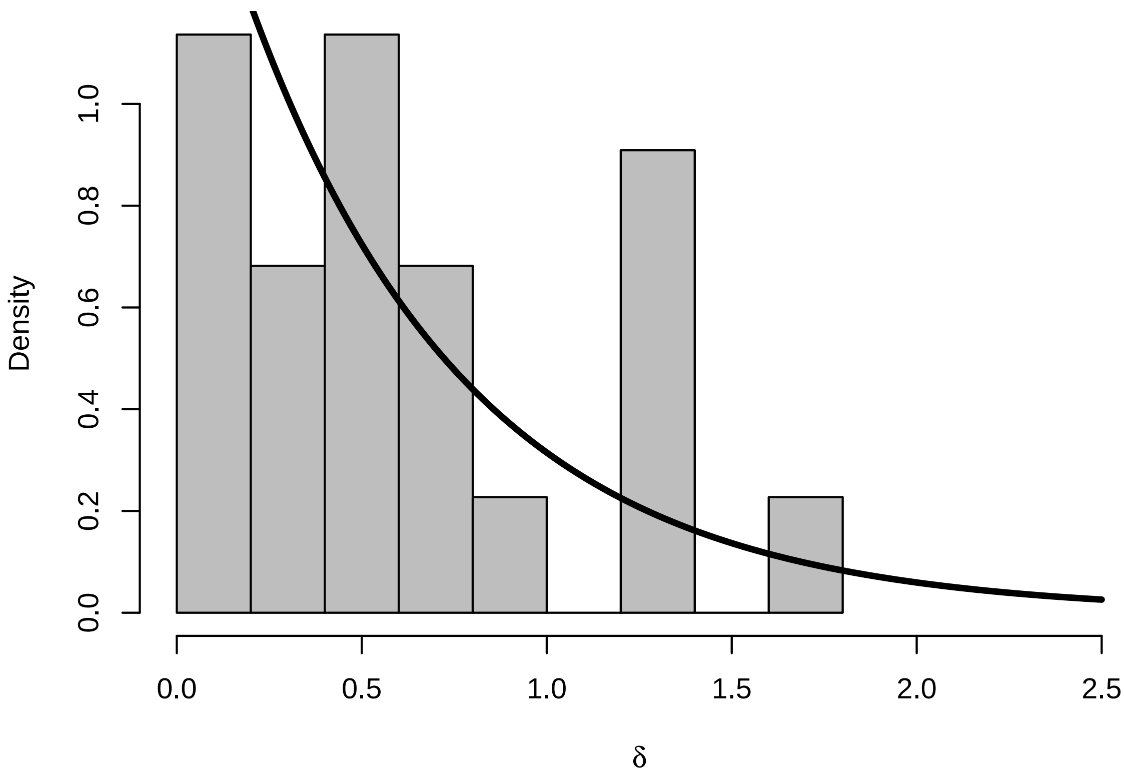
### A.5.3 Tables and Figures

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 | – |
| $\nu_z$ | $\times 10^2$ | 44.014 | 64.854 | 60.875 | 50.276 | – | – | – | – |
| $1-\beta^{(z)}_z$ | $\times 10^3$ | 26.434 | 25.576 | 26.125 | 38.538 | – | – | – | – |
| $\nu_x$ | $\times 10^1$ | 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 |
| $\alpha_r$ | $\times 10^2$ | 0.463 | 0.457 | 0.446 | 0.466 | 0.437 | 3.939 | 0.478 | 3.461 |
| $\mu_r$ | $\times 10^5$ | 0.548 | 0.556 | 0.561 | 0.574 | 0.539 | 0.570 | 0.665 | 0.919 |
| $\beta_r$ | $\times 10^{-5}$ | 1.826 | 1.797 | 1.784 | 1.741 | 1.856 | 1.752 | 1.503 | 1.088 |
| $\theta_z$ | $\times 10^3$ | 6.572 | 7.348 | 6.648 | 14.511 | – | – | – | – |
| $\theta_x$ | $\times 10^2$ | 8.639 | 7.594 | 8.848 | 11.999 | 9.698 | 4.849 | 6.629 | 10.995 |
| $\theta_r$ | $\times 10^{-2}$ | 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 | – | – |
| $\ell$ | | 0.109 | 0.000 | 0.055 | 0.000 | 0.353 | 0.106 | 0.588 | 0.000 |
| $\sigma_{RF}$ | | 0.291 | 0.291 | 0.291 | 0.291 | 0.287 | 0.295 | 0.291 | 0.292 |
| $\eta_{CDS}$ | | 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 |

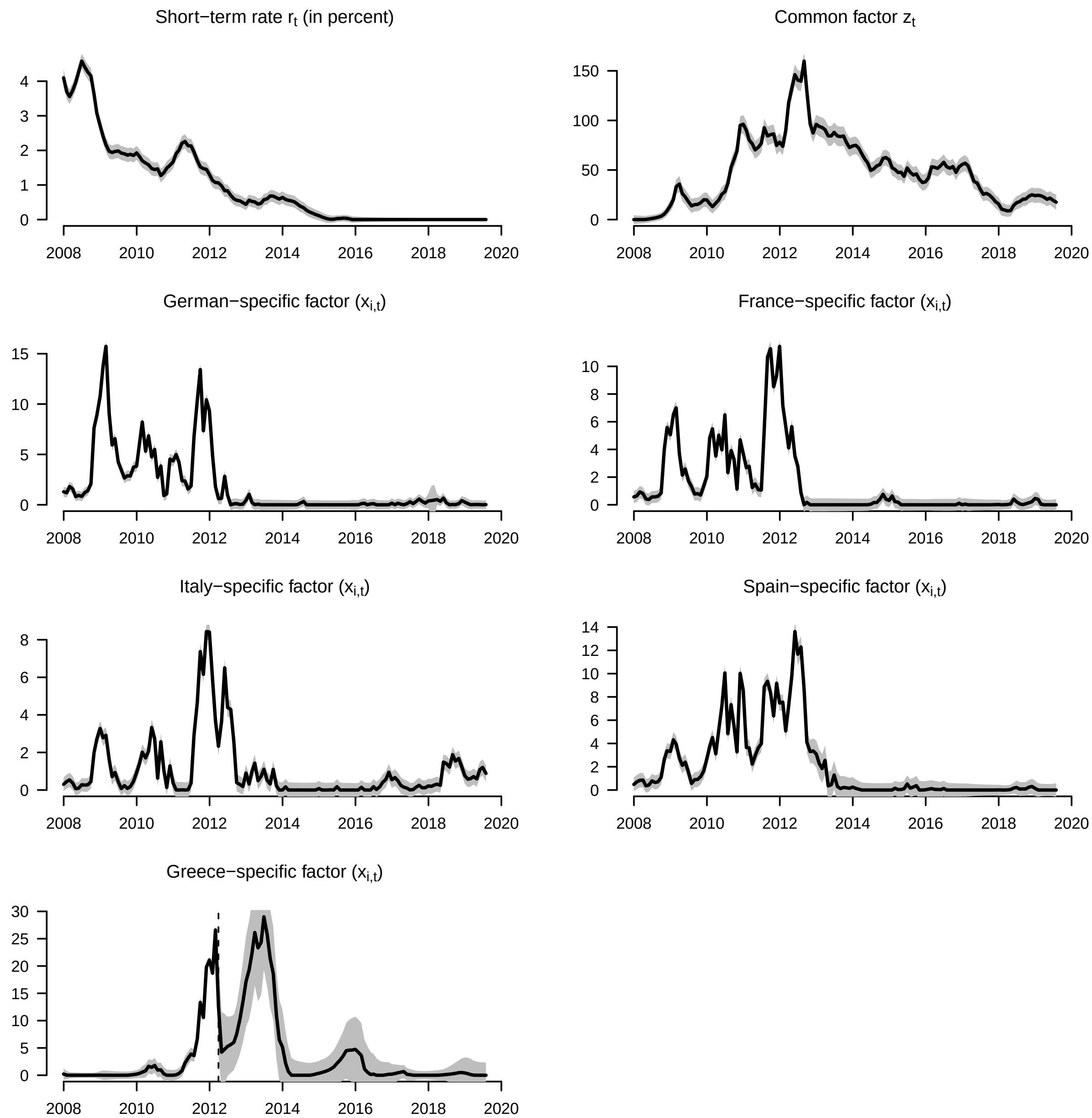
*Note*: Models (Eqs. 24, 25 and 26) are estimated by MLE. "–" indicates parameters that are constrained to be equal to zero. Model (1) is the baseline model; Models (5) to (8) feature no frailty factor ($z_t$) and Models (2), (4), (6) and (8) feature no contagion. $\sigma_{RF}$ is the standard deviation 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 $\eta_{CDS}$ 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 $\ell$. "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



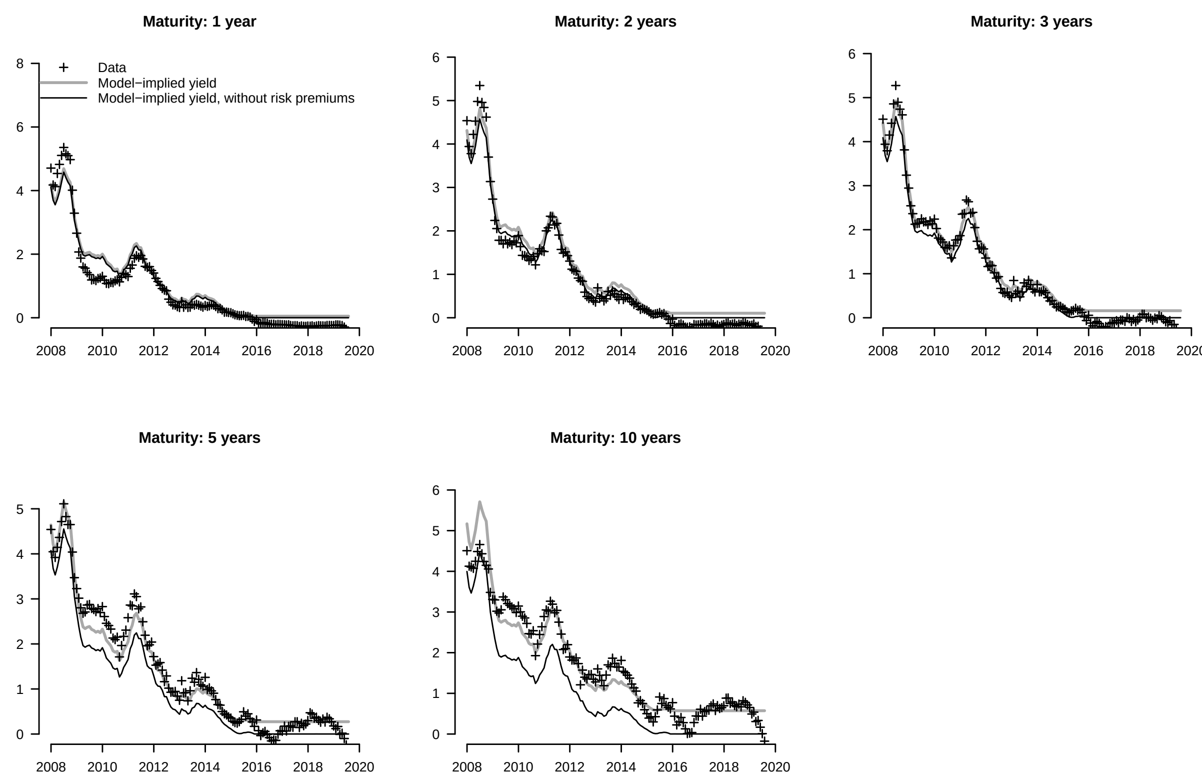

*Note:* This figure displays an histogram of $-\log(\overline{\varrho}_i)$, where $\overline{\varrho}_i$, $i \in 1, \dots, 22$, are estimates of the recovery rates associated with sovereign defaults that took place over the last thirty years (Moody's, 2016). In the RMV specification, $-\log(\overline{RR})$ is identical to the credit-event variable $\delta$. 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 4.2 for a description of these factors). The grey areas are two-standard-deviation bands, reflecting Kalman-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. $\theta_x$, $\theta_y$, $\theta_r$ and $\mathbf{S}$, to zero).