

QUANTIFYING THE TRANSMISSION OF EUROPEAN SOVEREIGN DEFAULT RISK¹²

ANA-MARIA H. DUMITRU ^a AND TOM D. HOLDEN ^b

We build a non-stationary Hawkes model of sovereign credit risk for seven European countries, and estimate it on CDS data from the run-up to the Greek default. We model a country's credit risk as partly driven by a weighted combination of risks across countries. We find Spain and Portugal are the chief drivers of this component, with Greece's contribution also significant. Greece and Portugal are found to be particularly sensitive to external risk, with a Greek default 35% less likely in our period without shocks elsewhere. Our novel maximum-likelihood procedure permits tractable estimation of high-dimensional Hawkes models with unobserved events.

JEL CODES: G01, G12, G15, C32, C58

KEYWORDS: sovereign CDS spreads, credit risk, multivariate self-exciting point process, systemic risk.

¹This paper was previously circulated with the title: "A Hawkes model of the transmission of European sovereign default risk".

²The views expressed in this paper are those of the authors and do not represent the views of the Deutsche Bundesbank, the Eurosystem or its staff.

^aDeutsche Bundesbank and the University of Surrey. Corresponding author.

E-mail: a.h.dumitru@gmail.com

Web: <http://sites.google.com/site/anamariahdumitru/>

Post: Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany

^bDeutsche Bundesbank and the University of Surrey.

E-mail: thomas.holden@gmail.com

Web: <http://www.tholden.org/>

Post: Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany

1. INTRODUCTION

Credit default swap (CDS) data reveals that most European countries experienced a steady increase in the cost of insuring their debt against default from January 2010 until the restructuring of Greek debt in February 2012. While many countries faced few internal risks, all were exposed to the risk of contagion following a default elsewhere. Thus, news about one country's credit-worthiness could impact credit risk across Europe. In this paper, we trace the sources of this risk. We find that the chief drivers over this period were actually Spain and Portugal. Without shocks to these two countries, we estimate that Greece would have been around 34% less likely to default during our estimation period. Thus, Greek sensitivity to external credit risk is a key part of the narrative of the European sovereign debt crisis; one that has been overlooked until now.

EU countries are intertwined in a complex web of economic, financial and political relationships, which are likely to facilitate multi-directional credit risk spillovers between member states.¹ Hence, there is a need for a joint analysis of credit risk in multiple countries. Given this, our model of credit risk is based on a high-dimensional Hawkes (1971) point process, which we estimate from sovereign CDS prices at multiple maturities via conditional maximum likelihood.² We do not impose stationarity in order to match the explosive path of Greece's CDS prices.

To overcome the curse of dimensionality, we model an international portfolio of country risk as a weighted average of individual country risks. This specification allows for a rich structure of instantaneous and dynamic spillovers, even with seven countries. The weights are estimated and convey information on the sources of contagion at the European level.

¹The channels of contagion are discussed by Acharya et al. (2014), Benzoni et al. (2015) and De Grauwe and Ji (2013).

²Conditioning on the initial observation.

We find that the build-up in risk in most European countries was driven by Portuguese and Spanish risk rather than Greek. Our estimates suggest that Greek default risk was on an explosive path, meaning that default was inevitable, sooner or later. Given this, other European countries had no choice but to insulate themselves from the risk of a Greek default,³ and so changes in the riskiness of Greek debt only had a small impact on risk in other countries. For Portugal and Spain however, default was not inevitable according to our estimates, reducing these incentives for insulation. Portugal had high debt to GDP, and was perceived as a substantial risk by the ratings agencies,⁴ while Spain was experiencing extreme levels of unemployment. These countries naturally then became key drivers of the upward trend in sovereign risk amongst European countries.

Furthermore, our estimates suggest that Spain was the most important source of spill-over credit risk, with the probabilities of default in all other countries considerably decreasing when we run a counter-factual exercise suppressing shocks originating in Spain. Indeed, our estimates suggest that without Spanish shocks, the probability of a Greek default during our estimation period would have been around 23% lower. This fall becomes around 34% if we additionally switch off Portuguese shocks, and around 35% if all other countries shocks are switched off, except Greece's. This illustrates the pivotal role played by Spain and Portugal over the period.

As an additional explanation for their role in driving systemic risk, we provide evidence that Portugal was the most systemically important debtor in the network of European debt holdings over the period, while Spain was the most systemically important creditor. Due to Spain's large size and substantial debt holdings, a Spanish default would have had severe

³This insulation chiefly occurred through the pooling of risk that accompanied the movement of Greek debt away from particular banks and nations, towards large multi-national institutions, such as the IMF, the EU and the ECB.

⁴See Table I.

repercussions for the rest of the EU. Thus, while default in Spain might have always been quite a remote risk, it is unsurprising that other EU nations were still very sensitive to changes in Spanish sovereign risk over our estimation period.

As a further contribution of this paper, we provide a method for likelihood based inference for Hawkes models when events are not observed. Indeed, our data only contains observations of a regularly-sampled, non-linear function of the unknown parameters and the event intensity. Our paper is the first in the literature to successfully surmount the substantial computational and econometric challenges involved in estimating high-dimensional Hawkes processes from price data, without observing events. Thanks to our maximum likelihood approach, we are also able to incorporate information on the non-occurrence of defaults.

1.1. *Our credit risk model*

Our model is a reduced-form credit risk model. Background on such models is contained in Duffie and Singleton (2003). These models assume that CDS spreads are a function of the risk neutral default intensity (event arrival rate), usually taken as a doubly stochastic Poisson process (Pan and Singleton, 2008; Longstaff et al., 2005; Ang and Longstaff, 2013). A more recent paper by Ait-Sahalia et al. (2014) points out that credit events⁵ and other jumps in risk should have a feedback effect on the probabilities of default of all other countries. To incorporate this feedback, the authors model the intensity of jumps as self-exciting Hawkes (1971) processes. They further price the CDS spreads based on this new class of models and estimate bivariate self-exciting models for some pairs of Eurozone countries, by non-linear least squares.

⁵In accordance to the 2003 International Swaps and Derivatives Association, Inc. (ISDA hereafter) Credit Derivatives Definitions, the term “credit event” for sovereign debt covers the following categories: failure to make payments, restructuring, repudiation and moratorium.

To capture the full breadth of information that might affect default risk, we will build a model of the arrival of *credit-related events*. These will group together all events that could have an effect on credit risk, including not only official credit events, but any other shocks impacting default risk, such as changes in sovereign credit ratings, or the arrival of news on macroeconomic conditions. We do not assume that credit-related events are observed, though we will assume that defaults are observable. We model the intensity of credit-related events as a multivariate marked Hawkes (1971) point process. For a marked point process, the intensity is not only a function of the events themselves, but also of the variably sized jumps that occur at event times. Augustin et al. (2018) also use a broader notion of credit events in exploring default and currency devaluation in the Eurozone. They treat such events as observed, equating them with sufficiently extreme movements in the quanto spreads. By contrast, in our paper, we model such events as unobserved, inferring them within the estimation process.

In our set-up, CDS spreads are determined entirely by a pure jump process. This is in line with recent literature stating the importance of modelling jumps in the prices of financial securities, either along with a diffusion or as the unique component (Geman et al., 2002; Carr and Wu, 2004). Our model generates both regular small changes and occasional large jumps, thanks to our assumption of exponentially distributed jump sizes. It thus provides a parsimonious way of modelling data in continuous time in the presence of jumps. Using a pure jump process also removes the need for costly high dimensional integration, which would be unavoidable in the presence of additional diffusion terms.

In our model, the intensity of credit-related events for each country has two components. The first is country-specific, and would be the sole determinant of dynamics were the cross-country spillovers switched off. The second component captures the effects of the market

portfolio of country risk. It is modelled as a weighted sum of the event intensities of individual countries. As a result, the aggregate risk in our model is endogenous, and reflects the systemic forces that may cause a default in one country to trigger further shocks elsewhere. The weights in the market component are estimated and confer information about the relative systemic importance of each country. Much as in classical finance theory, we also allow for each country to have a different sensitivity to the market component. By reducing the dimensionality of the parameter space relative to an unrestricted model, this set-up greatly eases the econometric and computational burden of estimation.

We estimate the model via maximum likelihood on a three dimensional panel of weekly CDS spreads for seven EU countries and seven maturities. The likelihood is based upon the implied shocks to the event intensities where intensities are derived by inverting the pricing formula. This is a novel approach to estimating this complex type of point process, made necessary by the fact that we only observe CDS spreads and not intensities, event times or jumps. Following Pan and Singleton (2008), we assume that the 5-year maturity spread is observed with no measurement error, while the spreads for the rest of the maturities contain normally distributed measurement error. To faithfully mimic the observed CDS spreads, we do not impose a stationarity restriction on the intensity of credit-related events.

Our modelling and estimation choices, accompanied by the use of a global optimisation algorithm, enable the estimation of a very complex model, with a huge parameter space (132 parameters, of which 84 are profiled out), while using a large three-dimensional data set. To the best of our knowledge, this is the richest estimation problem that has been tackled in the credit risk literature. Relative to e.g. Azizpour et al. (2018), our task is made more difficult by the rarity of sovereign defaults, and the fact that we do not observe the news

about default risk that is driving our model.⁶

1.2. *Additional literature*

The literature on sovereign credit risk is extensive, although a lot of the earlier work focused on emerging economies. Edwards (1984) estimates the perceived probabilities of default for different countries relying on the spread between the interest rate charged to a particular country and the LIBOR⁷ and finds that some aspects are overlooked by banks when pricing sovereign debt for developing countries. Kamin and von Kleist (1999) develop credit spreads measurements for emergent economies for the 1990s and find significant regional differences in spreads across these economies after controlling for risk and maturity. Eichengreen and Mody (2000) use data on 1000 developing-country bonds and find that higher credit quality is related to a higher probability of bond issue and lower spreads.

As CDS spread data started to become available, several empirical developments followed. Remolona et al. (2008) use CDS data on emerging markets to show that country-specific fundamentals are the prime drivers of sovereign risk, while changes in the risk premia are driven by changes in risk aversion. Longstaff et al. (2011b) identify the principal components in sovereign CDS data and find that a single principal component accounts for 64% of the variation in spreads.

Duffie and Singleton (2003) describe the use of the doubly stochastic Poisson processes to model default intensities, as a base for pricing debt and debt related instruments. Several further contributions to the literature modelled default intensities using these processes. Duffie et al. (2003) build a new model to price sovereign debt, which they further estimate

⁶Azizpour et al. (2018) look at default contagion from observed corporate defaults in the United States using a model incorporating a Hawkes process.

⁷London Interbank Borrowing Rate.

using Russian dollar-denominated bonds. Later contributions involving the pricing of CDS spreads include Pan and Singleton (2008), Longstaff et al. (2005) and Ang and Longstaff (2013). Notably, Ang and Longstaff (2013) consider an international set-up and incorporate in their model two default intensities, a non-systemic one and a systemic one. In their set-up, the above intensities are modelled as square root processes driven by separate Brownian motions, without spillovers between the systemic and non-systemic component.

Other recent literature has stressed the importance of using credit risk models that allow credit events to instantaneously impact the probability of default. Gouriéroux et al. (2014), Bai et al. (2015) and Monfort et al. (2017a)⁸ observe that credit risk models in which the price is a function of both default events and intensities perform better than those in which the price is not a function of the events. Benzoni et al. (2015) provide a structural explanation for this, via a model in which investors do not know the underlying economic state or its probability distribution. Thus, when investors observe defaults, they assign higher probabilities to bad states, which increases their valuation of default insurance. Our model allows for events to impact prices, as we rely on a Hawkes (1971) point process, in which the intensities respond to events.

As previously mentioned, Aït-Sahalia et al. (2014) were the first to use self-exciting point processes to model sovereign credit risk. While building upon their contribution for pricing CDS spreads when intensities are Hawkes (1971) processes, our work substantially differs from theirs in all other aspects: modelling, estimation, data and results. Firstly, we consider a marked point process, as opposed to an unmarked one. Errais et al. (2010) also consider a marked Hawkes (1971) process, but only estimate a univariate model to price index and tranche swaps. Secondly, we impose a particular parameter structure on the multivari-

⁸See also Monfort et al. (2017b) which introduces the Gamma zero processes used in Monfort et al. (2017a).

ate point process used, integrating a market component and a country specific component. Thirdly, we do not impose stationarity on the Hawkes (1971) process we use. This assumption is important due to the explosive path of Greek CDS. Fourthly, we develop a novel procedure for tractably estimating high-dimensional Hawkes models using CDS prices. This enables us to simultaneously work with 7 EU countries, while Aït-Sahalia et al. (2014) are constrained to estimate only bivariate models. As previously emphasized, the complex links between the EU countries demand exploring credit risk contagion in a multivariate framework. Moreover, instead of nonlinear least squares (NLLS), we rely on maximum likelihood estimation (MLE), both for its greater efficiency, and since the dimensionality of the parameter space under NLLS would be impossibly large. In addition, our data set covers many countries and maturities. Finally, the estimates obtained by Aït-Sahalia et al. (2014) sometimes generate unrealistic probabilities of default,⁹ whereas our estimated model generates more sensible values for these probabilities.¹⁰

1.3. *Outline*

The rest of the paper is structured as follows. Section 2 describes our data. Section 3 presents the model and derives the pricing of CDS spreads within it. Section 4 explains our estimation method. Section 5 gives our results, and presents an analysis of the network of European debt holdings to aid their interpretation. Finally, section 6 concludes.

2. DATA

We use weekly data for the 1-, 2-, 3-, 4-, 5-, 7-, and 10- year CDS contracts from seven EU countries from the 11th of November 2008 to the 28th of February 2012, giving 173 observations. We stop in February 2012, as on the 9th of March, the ISDA announced that a

⁹See Section 3.3 below.

¹⁰See Section 5 below.

“Restructuring Credit Event” had occurred in Greece. This triggered payments in the CDS market and temporarily suspended the trading of Greek CDS contracts. The cross-section of countries includes France, Germany, Greece, Italy, Portugal, Spain and the United Kingdom. We include the UK in addition to the other six Eurozone members both since previous research suggests the UK is the major non-Eurozone EU country affected by the Euro debt crisis (Stracca, 2015), and since during the observed period, the UK is accumulating increasing levels of public debt, leading to higher levels of sovereign default risk.¹¹ Note that other large non-Eurozone EU members, like Sweden and Denmark, have much lower debt to GDP ratios (around 40% for Sweden and 45% for Denmark¹²). In addition, data on consolidated banking positions for the UK, Denmark and Sweden shows that Eurozone banks have much stronger links to UK banks than to banks in Denmark or Sweden.¹³ We do not include Ireland for two reasons: first, it is a small country, with half the population of Portugal or Greece; second, while it had very high levels of debt, its GDP per capita was one of the highest in Europe (EUR 38,200 in 2012¹⁴). As shown in Table I, Greece and Portugal, displaying both very high levels of debt and low GDP per capita, fit more into the usual category of “peripheral” Eurozone nations.

Table I presents key facts for the countries in our sample. France, Germany and the UK all had the highest possible credit ratings within the period. Spain and Italy have the next worse credit ratings, with the higher Italian GDP per capita partially counter-balancing its higher debt. Finally, we have the two countries judged by the ratings agencies to have significant credit risk, Portugal and Greece. Greece has the highest debt to GDP in our sample of countries, reaching a level of approximately 160% of GDP in February 2012, with

¹¹The UK loses its top AAA rating in February 2013, around one year after the end of our sample.

¹²Source:Eurostat.

¹³Source: Bank of International Settlements. For quarter 4, 2010, see: https://is.gd/bisB4_20104.

¹⁴Source:Eurostat.

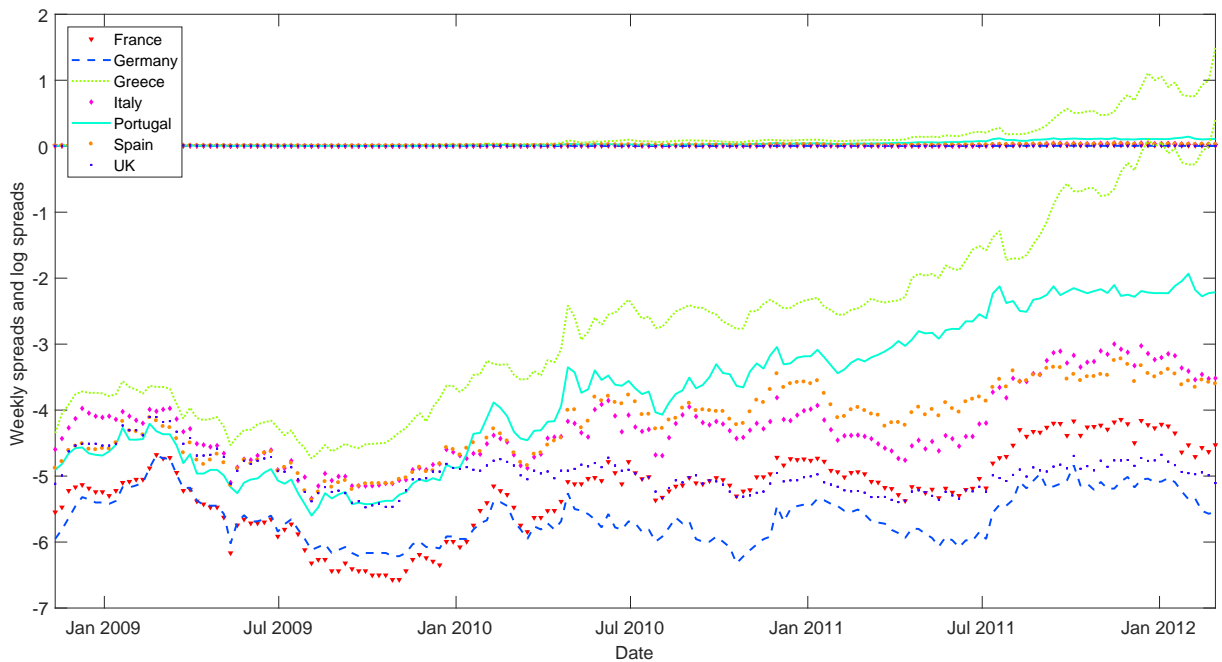
credit ratings suggesting that default was imminent.

Country	GDP/ capita (EUR, average 2008-2012)	Debt (%GDP, average 2008- 2012)	Debt (%GDP, 2012)	Gross debt outstanding (bil EUR, average 2008-2012)	Credit Ratings, February 2012 (Moody's, Fitch, S&P)
France	31020	80.6	89.5	1628	Aaa, AAA, AA+
Germany	32480	75.5	79.9	1975	Aaa, AAA, AAA
Greece	19880	142.8	159.6	312	Ca, CCC, CC
Italy	26960	114	123.4	1838	A3, A+, BBB+
Portugal	16640	97.8	126.2	171	Ba3, BBB-, BB
Spain	23180	61.5	85.7	659	A3, AA-, A
UK	30080	71.5	85.1	1330	Aaa, AAA, AAA

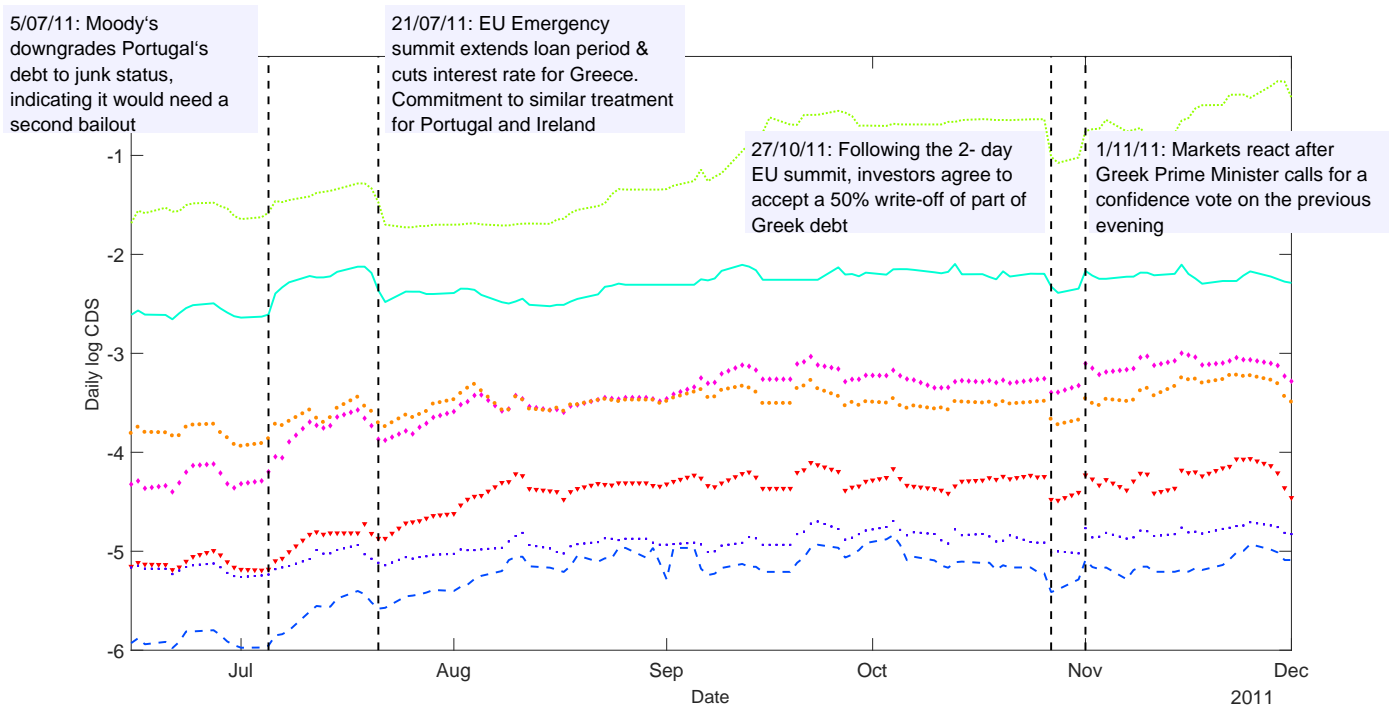
TABLE I

MACROECONOMIC SUMMARY FIGURES FOR THE SEVEN EU COUNTRIES. (SOURCE: EUROSTAT)

Figure 1 shows Euro-denominated 5-year CDS spreads in the run up to the Greek restructuring. The first panel, 1a, gives their dynamics both in levels, at the top of the plot, and in natural logarithm, at the bottom, over our estimation period. This plot reveals three facts. Firstly, Greek CDS spreads are much wider than those of other European countries. Secondly, given the very different magnitudes of CDS spreads across countries, the driving force for each country is likely to be its own credit risk. Finally, there are clear common patterns in CDS spreads across European countries. For instance, during and following the 2008 financial crisis, we see spreads widening across the board, before lowering again in 2009. Then, as the Greek debt crisis unfolded during 2010, similar patterns again emerge for all countries' spreads. Longstaff et al. (2011a) document this strong factor structure in sovereign CDS spreads.



(a) Euro denominated 5-year CDS spreads in levels (top) and logs (bottom) from November 2008 to February 2012 for 7 EU members.



(b) Log Euro denominated 5-year CDS spreads from the 15th of June 2011 to the 1st of December 2011 for the same 7 countries.

Figure 1: CDS spreads.

To get a better look at the joint dynamics in credit risk across countries, panel 1b in Figure 1 shows daily log 5-year CDS spreads for a shorter time window, from the 15th of June 2011 to the 1st of December 2011. In this zoomed figure, the commonalities across European countries are clearer. To illustrate this, we mark on the figure various events associated with a change in Greek or Portuguese risk. On each of these occasions, we observe corresponding movements in spreads in most countries. Two of the events marked on the panel represent good news. While in our model, risk can only jump upwards, there is still a natural interpretation for good news as the absence of bad news. With our estimates suggesting that over two hundred jumps are happening per week by the end of the sample, there is substantial scope for this channel to generate downward movements in risk.

Both panels in Figure 1 also show that the ranking of countries based on credit scores is preserved for CDS spreads. Additionally, from the plots, all series appear non-stationary. To confirm this, for every country's log CDS price, we run (augmented) Dickey-Fuller and Philips-Perron tests with all numbers of lags from zero to ten, both with and without a trend included.¹⁵ Of these 308 tests, not a single one had a p-value above 5%. Thus, there is no evidence that CDS spreads are stationary in any country over the period.

Our analysis also requires default free discount factors based on zero coupon bonds up to a 10-year maturity. Following Longstaff et al. (2005) and Aït-Sahalia et al. (2014), we use a cubic spline interpolation algorithm and rely on LIBOR rates (maturities 1-, 2-, 3-, 6- and 12- months) and Euro swap rates (maturities 2-, 3-, 5-, 7- and 10- years).

¹⁵In the case of the Philips-Perron test, the number of lags gives the number used in the Newey-West estimator of the long-run variance.

3. BUILDING A CREDIT RISK MODEL, STEP BY STEP

In this section, we describe our model of credit risk in EU countries. We begin with a more general specification then proceed to narrow it down in light of empirical, econometric and computational considerations.

3.1. *Self-excitation to capture systemic risk*

Systemic risk is the risk of collapse of an entire financial system. It is determined by the linkages between institutions, and, in a globalized economy, the linkages between countries. In our context, we are concerned with the systemic risk of defaults in all EU countries, triggered by one initial shock. With EU countries holding large amounts of each other's debt, and many sharing a currency, there is a legitimate concern of defaults being transmitted between countries. Thus, we seek to build a model of the impact of country specific shocks on the probability of default in nations across the EU. We do this by modelling the rate of arrival of credit-related events as a self-exciting (Hawkes, 1971) point process.

Let K be the number of countries. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space on which the credit-related events are defined. Then, for country $i = 1, \dots, K$, let $N_{i,t} \in \mathbb{N}$ be a stochastic process counting the number of credit-related events up to time t . In our model, each credit-related event will be associated with a jump in the event arrival rate. Define $J_{i,t} \in \mathbb{R}^{\geq 0}$ to be the sum, for country i , of all jump sizes up to time t . If a jump occurs at time t , both $N_{i,t}$ and $J_{i,t}$ change: $N_{i,t}$ by 1, $J_{i,t}$ by some real positive number. The vectors $N_t = (N_{i,t})_{i=1,\dots,K}$ and $J_t = (J_{i,t})_{i=1,\dots,K}$ group the counting and jump processes for all countries.

Now, let $\mathcal{F}_t \subseteq \mathcal{F}$ be the natural filtration for (N_t, J_t) . Then, the intensity of credit-related events at time t is defined as $\lambda_t := \lim_{h \downarrow 0} \mathbb{E} \left[\frac{N_{t+h} - N_t}{h} \middle| \mathcal{F}_t \right]$.

We suppose that λ_t follows the marked multivariate self exciting point process:

$$(1) \quad d\lambda_t = \check{\alpha} (\lambda_\infty - \lambda_t) dt + \check{\beta} dJ_t,$$

with solution $\lambda_t = e^{-\check{\alpha}t} (\lambda_0 - \lambda_\infty) + \lambda_\infty + \int_0^t e^{-\check{\alpha}(t-s)} \check{\beta} dJ_s$, where $\check{\alpha}$ and $\check{\beta}$ are $K \times K$ matrices. To ensure positiveness of the intensity, it is sufficient to assume that $\lambda_0 > \lambda_\infty$, $\check{\beta} > 0$ and that all off-diagonal elements of $\check{\alpha}$ are non-positive.¹⁶ Under these assumptions, $\lambda_t > \lambda_\infty$ for all t . If all eigenvalues of $\check{\beta} - \check{\alpha}$ are negative, then the process is stationary (Da Fonseca and Zaatour, 2015), and λ_∞ is the point to which the process would tend were there no more jumps.

The model in equation (1) shows that changes in the intensity of credit-related events in one country depend on both the intensity levels and the events in all other countries. The elements on the main diagonal of $\check{\beta}$ reflect the self-excitability of the model: shocks to a country lead to an increase in the arrival rate of further shocks to that country. This produces a clustering of shocks in time. The off-diagonal elements of $\check{\beta}$ reflect the cross-excitability between countries. Events in one country influence credit risk in all other countries, giving the system potentiality for a domino effect. As matrices $\check{\alpha}$ and $\check{\beta}$ do not need to be symmetric, the cross-country responses to other countries' intensities or shocks are not equal.

When matrices $\check{\alpha}$ and $\check{\beta}$ are diagonal, the event intensities in all countries are mutually independent. In this case, for every $i = 1, \dots, K$, there exist country-specific parameters α_i , β_i and $\lambda_{i,\infty}$ such that:

$$(2) \quad d\lambda_{i,t} = \alpha_i (\lambda_{i,\infty} - \lambda_{i,t}) dt + \beta_i dJ_{i,t}.$$

¹⁶See Lemma A.1 in Appendix A.

3.2. *Restricted model*

Figure 1 revealed co-movement across countries. This could be generated by a common “factor” driving all countries. There are two ways such a factor could be modelled. One would be to posit a purely exogenous driving process common to all countries. However, this would leave unclear the origins of this factor, and would, in any case, necessitate computationally expensive filtering based estimation. The second approach, that we adopt here, is to model the common factor as an endogenous object, driven purely by the shocks to individual countries.

Inspired by the use of stock exchange indices as a proxy for the market factor in the empirical classical finance literature, we build a “market” intensity of credit-related events defined as a weighted average of the intensities of all countries considered. Let: $w = [w_1, \dots, w_K]^\top$ be the vector of weights each country has within a “market” portfolio comprised of all countries, with $\sum_{i=1}^K w_i = 1$. The “market” factor at time t will be $w^\top \lambda_t$. Unlike in exchange indices, the weights here are not observed, but are parameters to estimate. They are designed to capture which countries’ dynamics matter most for explaining the commonalities between countries.

When a weight in one country is close to one, the relative importance of other countries would be poorly identified. Without loss of generality we may fix this by rescaling the weights as follows. For each country i , we define country-specific weights $W_{i,j}$, with $W_{i,i} = 0$ and $W_{i,j} = \frac{w_j}{\sum_{k \neq i} w_k}$. $W := [W_{i,j}]_{i,j=1,\dots,K}$ gives a matrix representation of these weights. This leads to a country i specific “market” intensity given by $W_i \lambda_t$, where W_i is the i -th row of matrix W .

In our model, we assume that changes in each country’s event intensity are driven by both country-specific and “market” components. Let δ_i and ϕ_i measure country i ’s sensitivity to,

respectively, changes and levels of the market intensity. The instantaneous change in the intensity for country i is given by the following stochastic differential equation:

$$(3) \quad d\lambda_{i,t} = \underbrace{\alpha_i (\lambda_{i,\infty} - \lambda_{i,t}) dt}_{\text{country-specific component}} + \underbrace{\beta_i dJ_{i,t}}_{\text{market change component}} + \underbrace{\delta_i W_i d\lambda_t}_{\text{market level component}} + \phi_i W_i (\lambda_t - \lambda_\infty) dt.$$

The first component is identical to the right hand side of equation (2), which gave the law of motion for $\lambda_{i,t}$ assuming independence across countries. Were this the only component, then countries would be independent. The second component captures country i 's response to changes in the “market”, and thus ensures a response of country i to shocks in other countries. The final component determines how country i responds to deviations in other countries' intensities from their long-run levels, λ_∞ . This component allows for spill-overs in risk across countries.

Regrouping the terms in the above equation and stacking together all countries leads to the following system of equations:

$$d\lambda_t = (\alpha - \phi W) (\lambda_\infty - \lambda_t) dt + \beta dJ_t + \delta W d\lambda_t,$$

where α , β , δ and ϕ are diagonal matrices with diagonals $\alpha_1, \dots, \alpha_K$, β_1, \dots, β_K , $\delta_1, \dots, \delta_K$ and ϕ_1, \dots, ϕ_K respectively. Further rearrangement of terms gives us:

$$(4) \quad d\lambda_t = \check{\alpha} (\lambda_\infty - \lambda_t) dt + \check{\beta} dJ_t,$$

$$\text{where } \check{\alpha} := (I - \delta W)^{-1} (\alpha - \phi W)$$

$$\text{and } \check{\beta} := (I - \delta W)^{-1} \beta.$$

While the restrictions imposed by (4) are intuitive, they also play an important role in dimensionality reduction for estimation purposes. If we consider the general multivariate self-exciting process in equation (1), matrices $\check{\alpha}$ and $\check{\beta}$ will have a total of $2K^2$ different elements.

With $K = 7$, this adds up to 98 parameters to estimate, whereas under our restrictions, these matrices contribute only 34 parameters. In any case, it is likely to be difficult to identify all parameters in the unrestricted system, as the identification must come from the non-linearity and non-Gaussianity of the system.¹⁷ Ait-Sahalia et al. (2014) estimate only bivariate models and impose a diagonal structure on the $\check{\alpha}$ matrix.

3.3. *Jump specification*

In order to attain a better fit of the original CDS data, we allow for variable sized jumps. With fixed size jumps, the distribution of next week's intensity conditional on this week's is multi-modal, with narrow support. By making jump sizes exponentially distributed, we ensure that the conditional intensity has full support, and a tractable distribution.

For country i , let $dJ_{i,t} = z_{i,t} dN_{i,t}$ be the jump occurring at time t , with a size equal to either 0, if no jump occurs, or a random variable $z_{i,t}$, if a jump occurs. We assume that $z_{i,t} \sim \text{Exp}(1)$, independent across time and countries. In vector form, we have $dJ_t = z_t \circ dN_t$, where $z_t = [z_{1,t}, \dots, z_{K,t}]^\top$.

We define $X_t := (N_t, J_t, \lambda_t)$, with dynamics given by equation (4) and the jump specification just given. By construction, X_t follows a Markov process affine in the state variables. Thanks to these properties, we may readily calculate the moments of X_t (see appendix B).

¹⁷In an equivalent linear-Gaussian framework, $\check{\beta}$ is not identified as there are K^2 elements in $\check{\beta}$, but only $\frac{K(K+1)}{2}$ elements in the covariance matrix from which $\check{\beta}$ must be inferred.

3.4. Pricing CDS contracts

We rely on the pricing formula derived by Ait-Sahalia et al. (2014):¹⁸

$$(5) \quad s_{i,\tau,t} = \frac{r \int_t^\tau D(t,s) \mathbb{E}[\gamma_i \lambda_{i,s} (1 - \gamma_i)^{N_{i,s} - N_{i,t}} \mid \mathcal{F}_t] ds}{\int_t^\tau D(t,s) \mathbb{E}[(1 - \gamma_i)^{N_{i,s} - N_{i,t}} \mid \mathcal{F}_t] ds},$$

where $s_{i,\tau,t}$ denotes the CDS spread at time t , for country i and for maturity τ , measured in years, $1 - r$ is the recovery rate (i.e. the fraction returned upon default), here assumed constant across countries, $D(t, s)$ is the discount factor at time t for a zero coupon bond with maturity s , and $0 < \gamma_i \leq 1$ is the probability of going into default upon the occurrence of a credit-related event. The pricing is under the risk-neutral measure, and thus λ_t is also under this measure. In practical terms, this means that changes in λ_t could be coming from either changes in the risk of a country's debt, or from changes in the market's appetite for this risk.

Looking at Figure 1, we observe a lot of variability in all CDS spreads. If we wish to describe this type of data using a pure jump process, which only moves when an event occurs, we will need a lot of jumps to faithfully mimic the variability in the data. This implies a very high intensity. For this not to generate counter-factually many defaults, this requires letting the parameter γ_i take a value lower than 1. Previous work (e.g. Longstaff et al. (2005); Pan and Singleton (2008)) fixed $\gamma_i = 1$, which is equivalent to assuming that all events generated by the model are proper credit events, as defined by ISDA. Most European countries considered had a good credit history in the recent past; the occurrence of credit events as in the ISDA definitions is very rare, and none occur in our sample.

Indeed, to reflect the observed data, γ_i should be close to 0. Consider the results of Ait-

¹⁸For estimation, the CDS data is converted from the original semi-annual to continuously compounded. We apply the formula: $s^{\text{continuous}} = 2 \log(1 + s^{\text{semi-annual}}/2)$.

Sahalia et al. (2014) for example, who estimate a bivariate model of the type specified in equations (1) with $dJ_t = dN_t$. Their estimated value of γ_i for Germany is 0.28, while a jump in Greece causes the intensity in Germany to increase by 1.3, under their estimates. This means that if a jump happens in Greece, then, per their estimates, the annual probability of default in Germany will increase by roughly $1.3 \times 0.28 = 0.36$, which seems somewhat improbable. To ensure reasonable estimates of γ_i , we include the series of observed non-default events within our estimation information set.

Computing the theoretical spreads $s_{i,\tau,t}$ involves the evaluation of the two expectations present in equation (5). Duffie et al. (2000) derive closed form expressions for expectations of this type for the class of affine jump diffusions. Appendix C offers details on the application of Duffie et al.'s (2000) results to our paper.

4. ESTIMATION

We estimate the model by maximum likelihood. Given that the state is not directly observed, the likelihood will be a product of the likelihood of the state, and a Jacobian term coming from the transformation mapping the state into the observed prices. Bowsher (2007) provides a comprehensive guide to the estimation of generalized Hawkes (1971) models when realisations of N_t are observed, which is not the case here.

The recent literature estimating credit risk models using CDS data has taken a least squares (or equivalently GMM) approach (Longstaff et al., 2005; Ang and Longstaff, 2013; Ait-Sahalia et al., 2014); it seeks to find values for the parameters and state-variables which minimise the gap between predicted and observed prices. This approach, while liberal in terms of error specification, requires considering the state variables (i.e. the intensities) as free parameters to estimate. This significantly increases the dimension of the parameter

space, rendering estimation of a model as large as ours essentially impossible.¹⁹ Additionally, the least squares approach throws away substantial quantities of information as it amounts to estimating the dynamics of market beliefs about the state variable, without using any information about the actual evolution of that state variable, e.g. auto-covariances.

We follow Pastorello et al. (2003), Pan (2002) and Pan and Singleton (2008) in assuming that, for all countries, all but one time series from our panel of financial instruments is observed with measurement error. Specifically, we assume that the 5-year CDS spread—the most liquid—is observed without true measurement error. Thanks to this, conditional on the values of the model parameters, the pricing formula for the 5-year CDS can be inverted to find λ_t at every observation, t . We perform the inversion sequentially, starting by finding λ_t in the first period, then in the next, and so on. In solving for $\lambda_{t+\Delta}$, where $\Delta := \frac{7}{365}$ is the sampling frequency, for each t , we impose the constraint that:

$$\lambda_{t+\Delta} \geq \mathbb{E}[\lambda_{t+\Delta} | \mathcal{F}_t, N_{t+\Delta} = N_t] = \lambda_\infty + e^{-\tilde{\alpha}\Delta} (\lambda_t - \lambda_\infty),$$

as any $\lambda_{t+\Delta}$ violating this condition has probability 0. Imposing this constraint also helps avoid multiple minima in the pricing error objective function. Moreover, when solving for $\lambda_{t+\Delta}$, we impose the constraint that there exist jump times $\tau_1^*, \dots, \tau_{L_{t,t+\Delta}}^* \in (t, t + \Delta]$ and jump amounts $z_1^*, \dots, z_{L_{t,t+\Delta}}^* \in \mathbb{R}_{>0}^K$ such that:

$$\lambda_{t+\Delta} = \mathbb{E}[\lambda_{t+\Delta} | \mathcal{F}_t, N_{t+\Delta} = N_t + L_{t,t+\Delta}, \forall l \in \{1, \dots, L_{t,t+\Delta}\} dJ_{\tau_l^*} = z_l^*].^{20}$$

Again, any violation of this constraint has probability 0 according to the model. Due to

¹⁹In our case, for 7 countries, we would have had an additional of 1211 new parameters to estimate.

²⁰For numerical reasons, in practice, this constraint is only imposed approximately. Further details of how it is imposed are given in Section 4.1.2.

misspecification, it may not always be possible to zero the error in pricing while imposing these constraints. In these cases, we minimise the pricing error conditional on imposing the constraints. This produces a residual pricing error for the 5-year CDS spread, a pseudo-measurement error. Since under correct specification, this pseudo-measurement error has standard deviation 0, and thus a singular likelihood, we do not attempt to estimate its standard deviation. Instead, we estimate the parameters governing the “pseudo-measurement” error for the 5-year CDS spread via shape-preserving piecewise cubic Hermite interpolation from the parameters governing other maturities.

To be concrete, we assume that the observed CDS spread at time t for country i and maturity τ , $\tilde{s}_{i,\tau,t}$, is given by

$$(6) \quad \tilde{s}_{i,\tau,t} = s_{i,\tau,t} + \varsigma_{i,\tau} e^{\varrho_{i,\tau} s_{i,\tau,t}} \varepsilon_{i,\tau,t},$$

where $\varepsilon_{i,\tau,t} \sim \text{NIID}(0, 1)$ and where $s_{i,\tau,t}$ is the CDS price at maturity τ implied by our λ_t . This form of heteroskedasticity gives substantial flexibility over how measurement error changes as prices change. Since $\varsigma_{i,5}$ and $\varrho_{i,5}$ are interpolated from $(\varsigma_{i,1}, \varsigma_{i,2}, \varsigma_{i,3}, \varsigma_{i,4}, \varsigma_{i,7}, \varsigma_{i,10})$ and $(\varrho_{i,1}, \varrho_{i,2}, \varrho_{i,3}, \varrho_{i,4}, \varrho_{i,7}, \varrho_{i,10})$ respectively, this gives $6 \times 7 \times 2 = 84$ measurement error parameters, all of which we are able to profile out of the likelihood.

One could potentially estimate our model via Gaussian quasi-maximum likelihood (QMLE). However, since we model non-stationary data, the standard consistency results (e.g. Bollerslev and Wooldridge (1992)) for QMLE do not apply as these require a finite unconditional variance. Consequently, we seek to estimate using the true likelihood, which should also be more efficient. In practice, some small approximations to the likelihood will still be necessary, but the error introduced by these approximations will disappear asymp-

totically as the sampling frequency goes to 0, giving consistency under the same conditions as true MLE.

4.1. Our likelihood

At $t + \Delta$, we observe two types of information. First, the CDS spread, which will provide information on the jumps between t and $t + \Delta$. Second, whether there was a default between t and $t + \Delta$, although in our sample, there are no defaults. Consequently, we want to calculate the joint likelihood of observing no default and a certain $\lambda_{t+\Delta}$, conditional on λ_t . This is computed as the product between the likelihood of observing $\lambda_{t+\Delta}$, conditional on λ_t , and the probability of no default conditional on both λ_t and $\lambda_{t+\Delta}$.

4.1.1. Likelihood of a certain jump size

For the computation of $f(\lambda_{t+\Delta} | \lambda_t)$, we approximate the true arrival rate λ_t over the time interval $(t, t + \Delta]$ by $\lambda_{t,t+\Delta}^u$, where $\lambda_{t,t+\Delta}^u$ is constant on $(t, t + \Delta]$. To minimise the approximation error, we define $\lambda_{t,t+\Delta}^u$ so that the expected number of jumps under the approximation (conditional on λ_t) is identical to the expected number of jumps under the true process (conditional on λ_t). This requires $\lambda_{t,t+\Delta}^u$ to be defined as the expected mean value of λ_t over the interval, conditional on λ_t , i.e.:²¹

$$\lambda_{t,t+\Delta}^u := \int_t^{t+\Delta} \mathbb{E}(\lambda_s | \lambda_t) ds = \bar{\lambda}_\infty + \Delta^{-1} (\check{\beta} - \check{\alpha})^{-1} \left[e^{(\check{\beta} - \check{\alpha})\Delta} - I \right] (\lambda_t - \bar{\lambda}_\infty),$$

where $\bar{\lambda}_\infty := -(\check{\beta} - \check{\alpha})^{-1} \check{\alpha} \lambda_\infty$. As $\Delta \rightarrow 0$, $\sup_{\tau \in (t, t+\Delta]} |\lambda_{t,t+\Delta}^u - \lambda_\tau| \xrightarrow{\text{a.s.}} 0$ and the error disappears.

Let $\chi_{t,t+\Delta} := \int_t^{t+\Delta} dJ_s = J_{t+\Delta} - J_t$. We will discuss the computation of this quantity in

²¹The solution follows from the calculation of $\mathbb{E}(\lambda_s | \mathcal{F}_t)$ in Lemma B.1 in Appendix B.

the next section. In Lemma D.1 in Appendix D, we show that under our constant arrival rate approximation, the probability distribution function (p.d.f.) for $\chi_{i,t,t+\Delta}$ is given by:

$$(7) \quad f(\chi_{i,t,t+\Delta}|\lambda_t) = \Delta \lambda_{i,t,t+\Delta}^u \frac{\mathcal{I}_1(2\sqrt{\Delta \lambda_{i,t,t+\Delta}^u \chi_{i,t,t+\Delta}})}{e^{\Delta \lambda_{i,t,t+\Delta}^u + \chi_{i,t,t+\Delta}} \sqrt{\Delta \lambda_{i,t,t+\Delta}^u \chi_{i,t,t+\Delta}}},$$

where $\mathcal{I}_1(\cdot)$ is the modified Bessel function of the first kind. We can then compute:

$$(8) \quad f(\lambda_{t+\Delta}|\lambda_t) = f(\chi_{t,t+\Delta}|\mathcal{F}_t) \left| \det \frac{d\chi_{t,t+\Delta}}{d\lambda_{t+\Delta}} \right|.$$

4.1.2. Inferred jump times and sizes

Although we observe λ_t and $\lambda_{t+\Delta}$, we do not observe J_t and $J_{t+\Delta}$, so we do not in fact know $\chi_{t,t+\Delta}$. A standard approach would be to take an Euler approximation. That is: $\chi_{t,t+\Delta} \approx \beta^{-1} [\lambda_{t+\Delta} - \mathbb{E}[\lambda_{t+\Delta}|\mathcal{F}_t, N_{t+\Delta} = N_t]]$. In the limit as $\Delta \rightarrow 0$, there would be at most a single jump in the interval $(t, t + \Delta]$, and the error in assuming it occurs at the end of the interval, as in the Euler approximation, would be $O_p(\Delta)$.²²

However, our sampling frequency (weekly) is relatively low compared to the estimated arrival rate (over two hundred per week by the end of the sample); thus, the approximation error from an Euler approximation can be substantial. This would manifest itself as upward bias in both α and β . To ameliorate this, within each interval, $(t, t + \Delta]$, we endeavour to find the jump times and jump sizes which maximise the likelihood over the interval. In the limit in which there is a single jump in the interval, we will be closer to its true arrival time, somewhat reducing error compared to the Euler approximation, and when there are many jumps in the interval, we will substantially reduce the error in the Euler approximation.

²²If the jump actually occurs at $t + s\Delta$, for $s \in [0, 1]$, then the error in using the Euler approximation is proportional to $\Delta(\lambda_{t+\Delta} - \lambda_t)(1 - s)$ plus higher order terms.

Solving for the maximum likelihood jump times and sizes can also be viewed as a “hard” expectation-maximization (EM) approach to dealing with the latent variables.²³

Let $\zeta_s := d\mathbb{E}[J_s | z_s, \lambda_s]/ds = z_s \circ \lambda_s$, where we treat z_s as a “noise” type process giving the jump size that would be observed at s were there indeed a jump. If we are free to optimise the likelihood over the number of jumps within an interval, then the maximum likelihood number of jumps is infinite if $\lambda_t > 1$ at some point in the interval (since in this case we may increase the likelihood without modifying the path by introducing a zero sized jump). In fact, solving this ill-posed optimisation problem is not necessary, since we are only concerned with $\chi_{t,t+\Delta}$, which is unaffected by zero sized jumps. Instead, we “integrate out” the unknown path of N_t , by maximising the expected likelihood over the interval, conditional on the entire path of λ_s and ζ_s for $s \in (t, t + \Delta]$ (which we are optimising over).

The expected maximum likelihood times and sizes of jumps occurring during $(t, t + \Delta]$ are the solution to the following maximization problem:

PROBLEM 4.1

$$\max_{\lambda_s, \zeta_s, s \in (t, t + \Delta]} \{ \mathbb{E} \log [f(\text{jump times in } (t, t + \Delta]) | \lambda_s, \zeta_s, s \in (t, t + \Delta)] \}$$

subject to:

$$(9) \quad \check{\alpha} (\lambda_\infty - \lambda_s) + \check{\beta} \zeta_s - \dot{\lambda}_s = 0,$$

$$\lambda_t, \lambda_{t+\Delta},$$

$$\zeta_s \geq 0 \quad \forall s \in (t, t + \Delta],$$

²³The canonical example of the hard-EM algorithm is the k-means clustering approach to estimating Gaussian mixture models.

where $f(\cdot)$ is the pdf of the jump times and equation (9) is obtained by taking expectations over equation (4) conditional on the paths of λ_s and ζ_s in $(t, t + \Delta]$. In practice, to keep the computational cost of solving this problem manageable, we take a first order Taylor approximation to the objective in λ_s , and we introduce a smoothing parameter “kappa” such that the solution to the problem with the linearised objective is recovered as $\kappa \rightarrow 0$. Further details on defining, approximating and solving this maximisation problem are given in appendix E. The optimal ζ_s for the approximated problem is:

$$(10) \quad \zeta_s = \max \left\{ 0, \frac{1}{\kappa} \left[\check{\beta}^T \left(e^{\check{\alpha}^T (s-t)} v + \check{\alpha}^{-T} \log \lambda_t \right) - 1_{K \times 1} \right] \right\},$$

where v is an integration constant, chosen to ensure $\lambda_{t+\Delta}$ takes the correct value. When $\kappa \rightarrow 0$, our solution is more accurate, and $\zeta_s \rightarrow \infty$ if the insides of the square brackets in equation (10) are positive, giving a jump at time s . In practice, for numerical reasons, we set $\kappa > 0$, but as small as is numerically feasible.²⁴

Further, we compute our estimate of the jump size over the interval $(t, t + \Delta]$ as:

$$(11) \quad \chi_{t,t+\Delta} = \int_t^{t+\Delta} \zeta_s \, ds.$$

4.1.3. Probability of no default conditional on a certain jump size.

Let $\Xi_{i,t,t+\Delta}$ denote the event that for country i , during the interval $(t, t + \Delta]$, no default occurred. For the computation of $\Pr(\Xi_{i,t,t+\Delta} | \lambda_t, \lambda_{t+\Delta})$, we approximate the true arrival rate λ_t over the time interval $(t, t + \Delta]$ by $\lambda_{t,t+\Delta}^c$, where $\lambda_{t,t+\Delta}^c$ is constant on $(t, t + \Delta]$. Much as

²⁴Estimations are obtained with $\kappa = 10^{-4}$.

before, to minimise the approximation error, we define $\lambda_{i,t+\Delta}^c$ so that the expected number of jumps under the approximation (conditional on λ_t and $\lambda_{t+\Delta}$) is identical to the expected number of jumps under the true process (conditional on λ_t and $\lambda_{t+\Delta}$). This requires $\lambda_{i,t+\Delta}^c$ to be defined as the expected mean value of λ_t over the interval, conditional on λ_t and $\lambda_{t+\Delta}$. The computation of this quantity is immediate from the solution of Problem 4.1.

Under the constant λ_t approximation, we prove in Lemma D.2 in Appendix D that

$$\begin{aligned}
 \Pr(\Xi_{i,t,t+\Delta} | \lambda_t, \lambda_{t+\Delta}) &= \mathbb{E} \left[(1 - \gamma_i)^k \middle| \chi_{i,t,t+\Delta} \right] \\
 (12) \qquad \qquad \qquad &= \frac{\sqrt{1 - \gamma_i} \mathcal{I}_1 \left(2\sqrt{\Delta(1 - \gamma_i)\lambda_{i,t,t+\Delta}^c} \chi_{i,t,t+\Delta} \right)}{\mathcal{I}_1 \left(2\sqrt{\Delta\lambda_{i,t,t+\Delta}^c} \chi_{i,t,t+\Delta} \right)}.
 \end{aligned}$$

4.1.4. The likelihood function

Combining the previous results, if we define the matrix of observations at t , $\tilde{s}_t := [\tilde{s}_{1,t}, \tilde{s}_{2,t}, \tilde{s}_{3,t}, \tilde{s}_{4,t}, \tilde{s}_{5,t}, \tilde{s}_{7,t}, \tilde{s}_{10,t}]$, we have that the conditional likelihood of the observation at $t + \Delta$ is given by:

$$(13) \quad f(\tilde{s}_{t+\Delta} | \tilde{s}_t) = f(\lambda_{t+\Delta} | \lambda_t) \Pr(\Xi_{i,t,t+\Delta} | \lambda_t, \lambda_{t+\Delta}) \left| \det \frac{d\lambda_{t+\Delta}}{d\tilde{s}_{5,t+\Delta}} \right| f(\tilde{s}_{t+\Delta} - s_{t+\Delta}),$$

where $f(\lambda_{t+\Delta} | \lambda_t)$ is as defined in equation (8), $\Pr(\Xi_{i,t,t+\Delta} | \lambda_t, \lambda_{t+\Delta})$ is as defined in equation (12), and $f(\tilde{s}_{t+\Delta} - s_{t+\Delta})$ is given by the error specification from equation (6).

The likelihood function in equation (13) may be readily adapted to the estimation of other self-exciting point processes, when we observe a regularly sampled function of the intensity. The only requirement for estimation based on (13) is for the function relating the observables to the state variables to be invertible.

4.2. *More on estimation*

The model estimated here is, to the best of our knowledge, the richest in the credit risk literature. We estimate a total of 132 parameters (profiling out the 84 parameters associated with measurement error), where each evaluation of the likelihood function requires us to solve multi-variate non-linear equations for each observation, both to find the cumulated jump size given the intensities, and to invert the CDS prices into intensities. We consider 7 countries and allow for asymmetric interactions between their credit risk. Moreover, a long time span and a relatively large panel of maturities are used. An estimation exercise of large dimensions is also conducted by Ang and Longstaff (2013). In their paper, they estimate jointly 33 parameters, 30 describing the dynamics for 10 countries (or US states) and 3 corresponding to an independent market factor.

In order to maximize the log likelihood, we employ a global optimization algorithm. Specifically, we use the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) proposed by Hansen and Ostermeier (2001). As the algorithm employs a global search through the parameter space, the estimation takes a considerable amount of time and computing resources. There are two advantages to such a global, stochastic, search strategy. Firstly, it removes the need for computing derivatives of the likelihood, which, due to the nested optimisation problems can be unreliable. Secondly, it increases the chance that our final optimum truly represents a global optimum. Estimation took around two weeks on a 20 core, 40 thread, machine, starting from a point with equal parameters across countries which had performed well in prior estimation runs. We used Matlab Coder to generate C code for some of the more time consuming inner code.

During estimation we impose the sufficient conditions for positivity of λ derived in Appendix A. Furthermore, following Ang and Longstaff (2013) and Aït-Sahalia et al. (2014), r

in equation (5) is set equal to 0.5, which is approximately the loss suffered by the investors in Greek bonds due to the country’s default on its debt. Given that this parameter multiplies the CDS pricing equation, fixing it at this level is a simple solution for any identification issues that might otherwise emerge.

5. RESULTS

Table II reports, for each country, the estimated parameters followed by standard errors in brackets. The bold estimates are significant at a 1% significance level.

Country(<i>i</i>)	α_i	β_i	δ_i	ϕ_i	$\lambda_{i,\infty}$	γ_i	w_i
France	39.46 (0.136)	39.80 (0.101)	2.23E-04 (0.007)	0.0316 (0.000)	0.64 (0.003)	2.33E-05 (0.002)	0.0007 (0.004)
Germany	47.23 (0.158)	47.82 (0.156)	4.05E-04 (0.008)	0.0177 (0.001)	0.27 (0.001)	1.62E-05 (0.008)	0.0002 (0.003)
Greece	18.56 (0.095)	18.58 (0.084)	9.82E-05 (0.001)	0.1172 (0.001)	0.37 (0.001)	3.66E-04 (0.003)	0.0102 (0.001)
Italy	28.35 (0.231)	28.54 (0.152)	6.06E-05 (0.001)	0.0165 (0.002)	1.45 (0.006)	3.59E-05 (0.004)	0.0021 (0.002)
Portugal	17.54 (0.093)	17.58 (0.110)	1.85E-04 (0.014)	0.0269 (0.001)	0.86 (0.034)	0.0001 (0.001)	0.5012 (0.015)
Spain	19.42 (0.141)	19.51 (0.128)	1.30E-04 (0.007)	0.0135 (0.002)	1.42 (0.007)	4.55E-05 (0.001)	0.4844 (0.002)
UK	33.59 (0.283)	33.70 (0.266)	5.19E-05 (0.002)	0.0374 (0.001)	3.37 (0.012)	2.31E-05 (0.018)	0.0011 (0.012)

TABLE II

ESTIMATED PARAMETERS AND STANDARD ERRORS^{25,26}

The estimates for α_i and β_i are very high and statistically significant at 0.01% for all countries. α_i needs to be high to ensure that the model puts positive probability on the fastest drop in CDS prices observed in the data, and β_i needs to be broadly similar to explain the observed slow reversion to $\lambda_{i,\infty}$. The values of these parameters for Greece, Portugal and Spain are lower than for all other countries; this is natural as problems were more persistent in these countries. Interestingly, for all countries we have that $\beta_i > \alpha_i$, which supports the conclusion of non-stationarity we reached from our initial unit root tests.²⁷

While the values for δ_i are all very low and statistically non-significant even at 5%, the estimates of the other sensitivity parameter, ϕ_i , are much higher and even statistically significant at 0.01%. Thus, events intensities change more as a reaction to the levels of the “market” intensity rather than due to the changes in this factor. This constitutes evidence for a lagged response to changes in the “market” risk of default. Put another way, countries respond to the underlying risk of others, rather than directly responding to news about other countries.

The highest level of ϕ_i (≈ 0.12) is encountered for Greece, meaning that Greece’s credit risk is particularly sensitive to the credit risk levels of all other countries. This helps explain the country’s tumultuous dynamics in this period, and implies that Greece’s fate might have been better had other European countries not also been in difficulty at the same time. We further explore the effects of other countries’ shocks on Greece in Section 5.2.

²⁶We calculate the standard errors using the sample Fisher information matrix. To deal with the fat tails of this measure, we trim the 5 largest and 5 smallest values. In the absence of trimming, all coefficients are significant. Moreover, all coefficients are significant when using the Hessian or the sandwich matrix to calculate the standards errors, thus our trimming approach is conservative with respect to the significance test.

²⁶Given that our estimates suggest the process is explosive, there is a possibility that some of our estimators are super-consistent. As we rely on a combination of in-fill and long-time asymptotics, we expect to attain a mix of standard \sqrt{T} and faster convergence. Thus, our choice to perform standard \sqrt{T} inference will produce conservative standard errors.

²⁷It is also the case that all of the eigenvalues of $\check{\beta} - \check{\alpha}$ are positive, so the estimated model is non-stationary.

Apart from Greece, the levels of ϕ_i are moderate. This means that the credit risk of each European country is chiefly determined by the individual circumstances of that particular country, rather than by the circumstances in other countries. This is a natural finding, given the fact that the analysed European countries differ considerably in terms of debt, economic development and credit ratings, as previously shown in Table I.

The estimates for $\lambda_{i,\infty}$ are statistically significant even at 0.01% for all countries, implying that for no country in our sample will default risk ever hit zero. We observe the highest levels for the UK, followed by Italy and Spain. This indicates higher levels of intensities for these countries even in the absence of shocks, which might be related to their high levels of debt and thus, underlying risk. However, it is difficult to interpret this parameter in a nonstationary model.

The probability of default conditional on the occurrence of a credit-related event (γ_i) is estimated to be near zero for all countries. Since frequent events are needed to explain all of the movements in the observed CDS data, this is unsurprising. The events in the model correspond to the arrival of any news relevant to the countries' ability to repay their debt, and in times of crisis, almost any political, financial or economic development can fulfil this criterion.

5.1. *The estimated weights*

In our model, the “market” intensity represents, essentially, a portfolio of common default risk. The estimated weights give information on the sources of this common risk. Looking at the weights in table II, we notice that the main sources of risk are Portugal and Spain, with weights of 50% and 48% respectively. Greece comes next with a weight of only 1%, but still statistically significant at 0.01%. The weights for all other countries are very small and

non-significant even at 5%. The prevalence of Portugal, Spain and Greece in driving the risk during this period is natural, given that, from our cross-section, these three nations required international financial aid. Below, we give further intuition for the estimated values.

During the analysed period, we observe an extreme deterioration of the economic situation in Greece, with poor economic performance and rocketing debt to GDP, leading to a consistent downgrading of credit ratings (see Table I). In the face of an imminent default, investors sought insulation from the risk coming from Greece, as evidenced by the gradual change in the ownership of Greek debt to large multinational institutions, such as the IMF, the EU and the ECB. In addition, the two bailouts received by Greece in this period via the First and Second Economic Adjustment Programmes further helped alleviate risk coming from Greece for other countries. All of the above are reflected in the low weight estimated for Greece, as the country was perceived as a problem that was already being dealt with. Additionally, Greece is a relatively small economy by European standards, so it would have been surprising had its weight been much larger.

The situation is different for Portugal, as investors were still learning about its default risk over our period. For example, as shown in panel 1b in Figure 1, Moody's downgrading of Portuguese sovereign debt on the 5th of July, 2011 was accompanied by reactions in all observed sovereign CDS spreads. We believe the high weight estimated for Portugal reflects investors' worries that Portugal could be the next Greece, and that its collapse could take other countries with it.

Still, Portugal's high weight might be surprising given the country's small size. To further understand the systemic importance of Portugal, we examined the network structure of debt amongst our seven countries. It is plausible that if a country holds some of the debt of a risky country, then this increases the risk of the creditor nation. In this way, risk is transmitted

through the network of asset holdings. More concretely, we consider a network structure on our seven countries where the strength of the link from country A to country B is proportional to the amount that country B has lent to country A, divided by the GDP of country B, times the debt to GDP ratio in country A. We divide by the GDP of country B to capture the fact that a default in a small country is unlikely to lead to the default of a large one. The debt to GDP ratio in country A proxies the country-specific risk of country A. This gives a rough approximation to the strength of the transmission of risk from country A to country B. We use consolidated banking exposures from the Bank of International Settlements for quarter 4 of 2010.²⁸ In addition, we collect GDP and public debt data from Eurostat.

Using this network, we would like to assign a “debtor-importance” score and a “creditor-importance” score to each country. Plausibly, important creditors are those who lend to important debtors, and vice versa. This recursive definition maps onto the “hubs” and “authorities” of Kleinberg (1999b).²⁹ In particular, Kleinberg (1999a) shows that the weights (importance scores) of “authorities” (creditors in our set-up) and “hubs” (debtors) converge to the left and right singular vectors of the adjacency matrix associated with the network structure. The resulting debtor- and creditor- importance scores are reported in Table III below. We learn that the most systemically important debtor is Portugal, followed by the UK, and that the most systemically important creditor is Spain, followed by Germany.

With Spain being the most systemically important creditor according to our network analysis, this gives one explanation for its prominence in our estimates. Furthermore, while Spain did not have large amounts of public debt, our dataset covers a period of Spanish recession,

²⁸This was the only period between 2009 and 2011 for which data for Italy was available.

²⁹Kleinberg (1999b) develops an algorithm which classifies web pages as “authorities” (pages linked by many other pages) and “hubs” (pages that link to many other pages). Each page is assigned an authority weight, proportional to the sum of the hub weights of pages that link to it, as well as a hub weight, proportional to the sum of the authority weights of pages that it links to. Using the same method, we can define weights (importance scores) for debtors and creditors.

Country	France	Germany	Greece	Italy	Portugal	Spain	UK
Debtors	0.048	0.008	0.026	0.063	0.606	0.005	0.244
Creditors	0.130	0.136	0.017	0.023	0.016	0.580	0.098

TABLE III

IMPORTANCE SCORES FOR DEBTORS AND CREDITORS

characterized by unemployment rates of over 20%³⁰. This recession, following the collapse of the property boom, left the Spanish banks very fragile and led the Spanish Government to seek a bailout during the summer of 2012 to inject capital into banks.³¹ At the same time, Spain was the 5th largest economy in Europe³² and was perceived by markets as too big to fail. Given all of this, the high weight we estimate for Spain is unsurprising.

We note that France and Germany have very low and statistically insignificant weights, despite arguably being the most prominent powers in the EU. We offer two explanations for this. Firstly, $w^\top \lambda_t$ is a portfolio of risk and, as such, is mostly driven by countries with high variability in their spreads. Secondly, with default in France and Germany such a remote risk, investors are likely to allocate their attention towards higher-risk countries. This rational inattention³³ will decrease the sensitivity of CDS outside of the EU's core to events within it.

Ang and Longstaff (2013) and Augustin et al. (2018) proxy the systemic risk component by the German probability of default. This is imposed mainly to ease identification and is motivated by the common supposition that the Eurozone is as strong as its strongest economy, Germany. Implicitly, this assumes Germany is prepared to act as a lender-of-last-resort to the entire Eurozone, something that is called into question by the limited political

³⁰Source: Eurostat.

³¹See, for instance, <https://is.gd/bbc100612>.

³²During 2008-2012, Spain's total GDP (current prices) was ranked 5th after Germany, France, the UK and Italy. Source: Eurostat.

³³For more on rational inattention, see, for instance, Sims (2003).

support within Germany for the provision of aid to other Eurozone countries³⁴. By contrast, our estimated weights suggest that systemic risk originates in weaker countries, implying that markets do not believe Germany would provide indefinite backing for other Eurozone members.

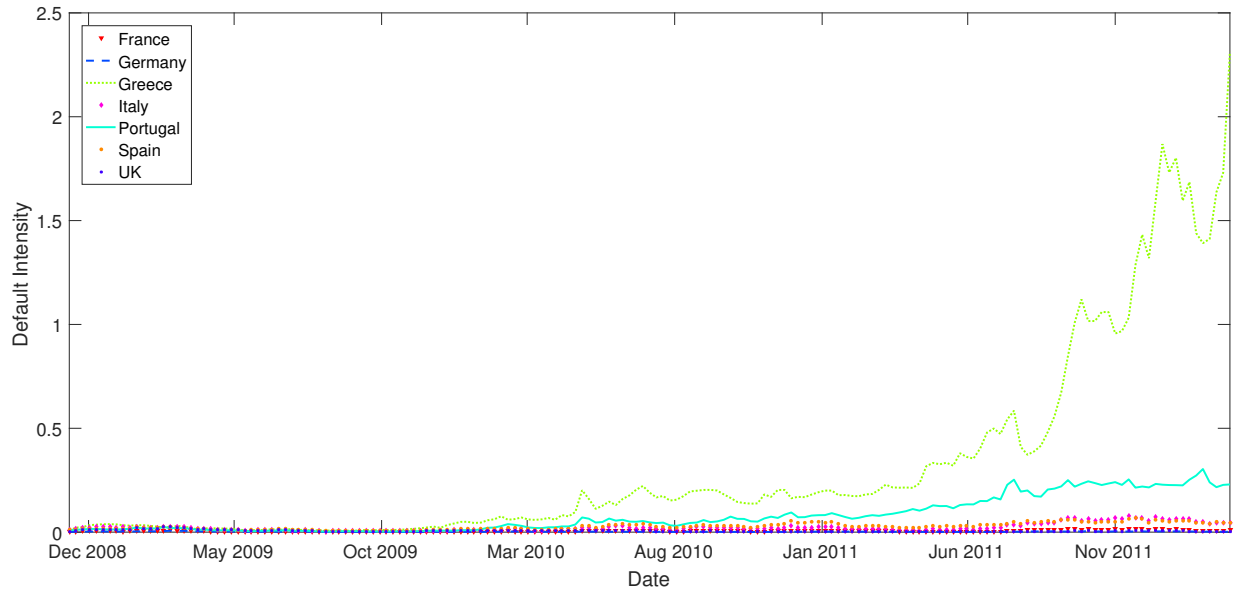
Panel 2a in figure 2 shows the estimated default intensities for all countries, obtained by multiplying country i 's event intensity by γ_i . We notice the similarity between this figure and the figure showing the original CDS spreads (figure 1). The intensity of default is extremely high for Greece, reaching a maximum of 2.31 expected defaults per year. Greece is followed by Portugal with a maximum intensity of default of 0.30 (expected defaults per year). Italy and Spain come next reaching maximums of 0.08 and 0.07 respectively. The lowest maximum default intensity was attained by Germany (0.007).

Panel 2b in figure 2 shows the evolution of the yearly “market” factor for the observed period together with the annual intensities of credit-related events for all countries. As we approach the Greek default, we notice the “market” intensity increasing to levels similar to the intensity levels in Spain and Portugal. This is expected, given that the highest weights in the market portfolio are taken by these countries. For Greece, we observe a maximum intensity of 6316, meaning 121 expected events per week.

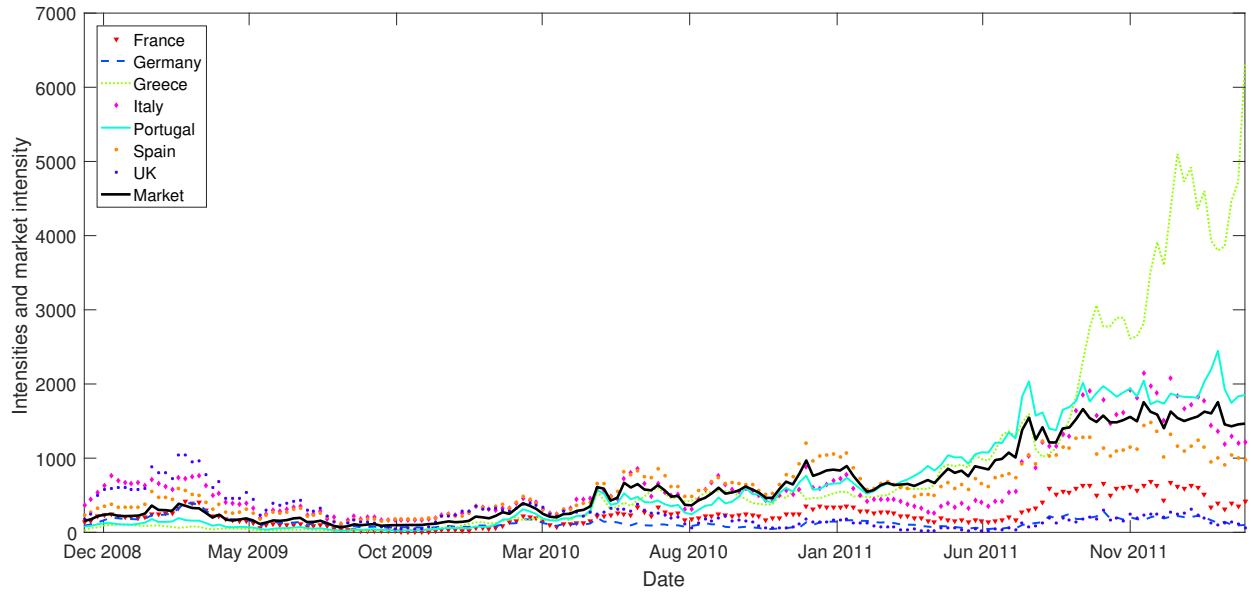
5.2. *Implications of our estimates*

This section quantifies the contagion channels implied by our estimated model. We present counter-factual results on probabilities of default for our analysed countries under assorted

³⁴For example, in August 2011, a poll of the members of Germany's main ruling parties found less than half of them supported the bailouts of Greece, Ireland and Portugal (see e.g. <https://is.gd/bbc100811>); in September 2011, Merkel stressed that she would never support a debt union (see e.g. <https://is.gd/bbc280911>); and, in February 2013 the “Alternative für Deutschland” party was founded on a strongly anti-bailout platform (see e.g. <https://is.gd/gu140413>).



(a) Estimated intensities of default for all the countries in the sample.



(b) Estimated annual event intensities and “market” event intensity during the observed period.

Figure 2: Intensities of default and credit-related events.

scenarios. We begin by focussing on the impact of the model’s parameters on the probability of a Greek default, as the value of ϕ_i there suggests Greece is particularly sensitive to risk elsewhere. The second part of this section explores the transmission of shocks between all countries. To perform these analyses, we recover the inferred shocks ζ_t from our estimation, and then re-generate the intensity paths of the model with certain parameters or shocks set to zero.

Table IV illustrates the effects of zeroing β_i , ϕ_i , $\lambda_{i,\infty}$ and w_i for one country at a time.³⁵

Country(<i>i</i>)	France	Germany	Greece	Italy	Portugal	Spain	UK
β_i	-0.013	-0.005	-99.049	-0.103	-8.662	-23.418	-0.048
ϕ_i	-0.001	0.000	-35.002	-0.001	-0.856	-0.092	-0.001
$\lambda_{i,\infty}$	-0.003	-0.001	-7.089	-0.008	-1.526	-1.400	-0.008
w_i	0.008	0.001	0.012	-0.042	10.491	-13.282	-0.015

TABLE IV

PERCENTAGE CHANGE³⁶ IN THE PROBABILITY OF THERE BEING AT LEAST ONE GREEK DEFAULT DURING OUR OBSERVATION PERIOD WHEN THE PARAMETER GIVEN IN THE FIRST COLUMN IS ZEROED FOR THE COUNTRY GIVEN IN THE FIRST ROW.

Switching off β_i for a certain country is equivalent to switching off the shocks from that country. Doing this for Greece renders a Greek default almost impossible. More importantly, we notice a substantial impact from turning off β_i for Portugal and Spain. While the estimated weight for Portugal is slightly higher than the estimated weight for Spain, this exercise shows us that shocks from Spain actually matter considerably more; we observe a drop of approximately 23% in the Greek default probability when switching off shocks from Spain, as opposed to a drop of approximately 9% when switching off shocks from Portugal.³⁷ This

³⁵Results for switching off α_i , δ_i and γ_i are available upon request. α_i controls mean reversion; hence, when it is fixed to zero in some country, that country experiences an explosive path and, via contagion, Greek default becomes a virtual certainty. As δ_i and γ_i are very small to start with, zeroing them does not have any great impact.

³⁶To be clear: an increase in default probability from 25% to 50% would appear here as 100%.

³⁷Mathematically, this is down to Spain’s higher long-run level in the absence of shocks, $\lambda_{i,\infty}$, its greater

is consistent with Spain's greater size, meaning its default would be more devastating than a default in a smaller country such as Portugal.

Repeating the above analysis while switching off shocks originating in both Spain and Portugal leads to a decrease of approximately 34.3% in the Greek probability of default. In addition, if shocks from all other countries are set to 0, the probability of default in Greece decreases by 34.6%, indicating again that from the pool of considered countries, Spain and Portugal are the chief driving forces of contagion risk. These results imply that a sizeable part of Greece's default risk was down to the difficulties experienced by other countries at the same time. It is thus reasonable to believe that Greece's fate could have been better had other European countries not also been in difficulty at the same time.

When switching off ϕ_i for each country in turn, the largest impact is unsurprisingly for Greece itself (around 35%), as in all other cases the effect is indirect.³⁸ This parameter controls the change in the intensity of credit-related events in response to changes in the level of the "market" intensity, so the high change in default probability we find here reinforces our prior finding of the importance of contagion for the Greek default. Indeed, the effect of setting both ϕ_i and δ_i to zero for Greece is the same as the effect of switching off shocks in all other countries, combined with the effect of setting all of their initial intensities ($\lambda_{i,0}$) to their long-run levels in the absence of shocks ($\lambda_{i,\infty}$). However, since δ_i is so low, ϕ_i is the main channel for contagion in practice. This finding also implies that around 65% of the Greek probability of default was generated at a purely domestic level.

Zeroing $\lambda_{i,\infty}$ in Greece generates a 7% decrease in its own default probability. In addition,

degree of intrinsic explosiveness, as evidenced by the higher $\beta_i - \alpha_i$, and also the higher implied value of $\lambda_{i,0}$ for Spain as opposed to Portugal.

³⁸Of the other countries, the effect of setting ϕ_i to zero is largest for Portugal. This reflects the combination of Portugal's high weight (w_i), and high sensitivity to market risk (ϕ_i). Note in particular that Portugal has a much higher sensitivity to market risk than Spain.

zeroing this for Portugal and Spain has a moderate effect on the Greek default probability. These effects stem from the fact that lowering $\lambda_{i,\infty}$ tends to lower the level of the event intensity in a country, combined with the large weights of Portugal and Spain.

Finally, zeroing the weights for Portugal and Spain has a considerable impact on the default probability in Greece. We note that other weights are always rescaled so that the weights add up to one. This explains why reducing Portugal's weight actually increases the chance of a Greek default. With Portugal's weight set to zero, almost all weight is on Spain, which has particularly explosive dynamics (i.e. $\beta_i - \alpha_i$ is higher for Spain). Since Greece responds to the market component, and Spain responds to Greece (now even more strongly than before due to the zeroing of Portugal's weight), Greece inherits Spain's more explosive dynamics, leading to a higher default risk. Conversely, when we set Spain's weight to zero, the relatively unexplosive Portugal has most of the weight, leading to a lower chance of Greek default.

Figure 3 shows the impact of turning off shocks for one country at a time on the probabilities of default of each other country. Just as before, the main sources of credit risk spillovers for the countries considered are Portugal and Spain. Moreover, responses to shocks from Spain are bigger in absolute value. We again attribute this result to the fact that Spain is a much bigger country than Portugal. This makes Spain the most important source of sovereign credit risk during our period.

5.3. *More on contagion*

Our model enables us to explore the effects of an event in a particular country on CDS prices in all other countries. We do this with an impulse response analysis of the estimated model. In particular, we analyse the median effects of an event occurring at the start of the sample (the 11th of November 2008), for each country in turn. We focus on the median

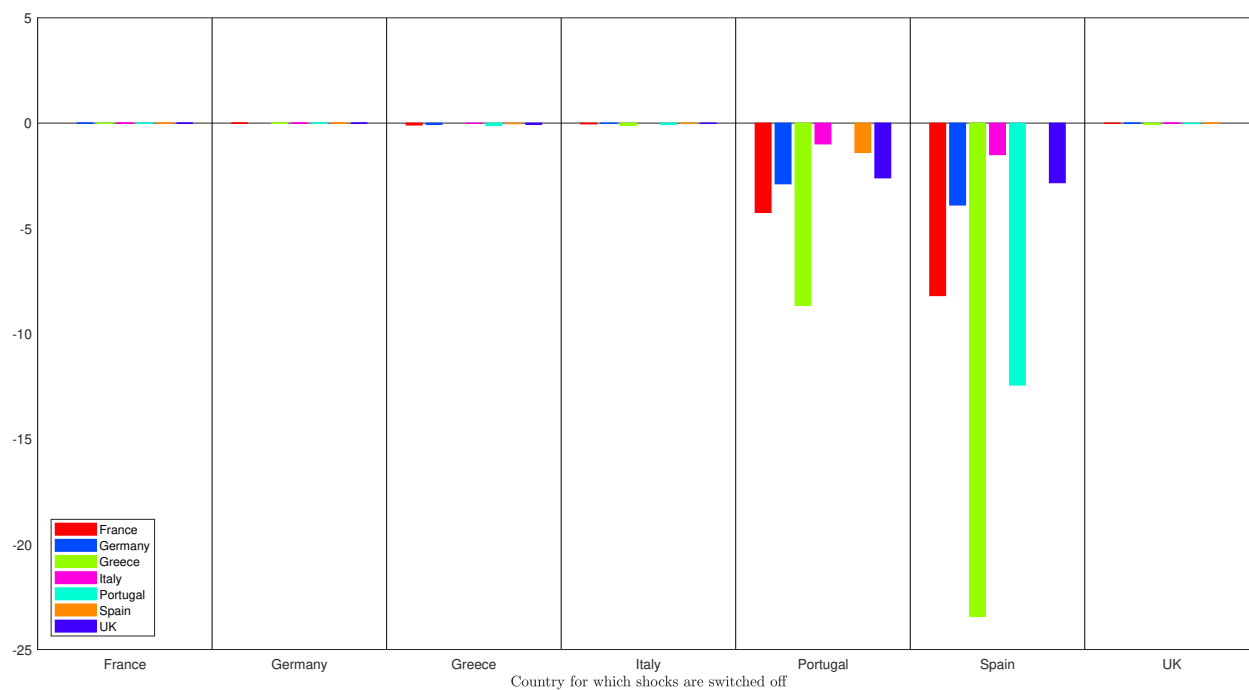


Figure 3: Percentage change in the probability of there being at least one default for a given country during our observation period when switching off shocks for one country at a time.

response, both since the non-stationarity of the point process implies that the mean response is likely to diverge, and since the fat-tails of the process render accurate Monte Carlo means very computationally expensive. To capture the potentially different responses to large and small shocks, we also report the quartiles of the responses.

Figures 4 and 5 show the percentage quartile responses of CDS prices in all countries to events in Portugal and Spain (first figure) and Greece (second figure). Responses to events in all other countries are included in Appendix F. All figures are generated by taking the Monte Carlo quartiles of the difference between simulations starting from the estimated λ_t at the start of the sample with an additional event in that week, and simulations without such an additional event.³⁹

In all figures, the bigger plots on the left are dominated by the response of each country to its own event, but other countries' responses are also visible. The 25% quartile and the median show a fast, exponential decay in all cases. However, for the 75% quartile, we observe greater persistence, with only linear decay. Thus the non-linearities of the model's solution ensure greater persistence to larger shocks.

The smaller figures show the separate responses in the other countries. As with the own country response, in all cases, the 25% and 50% quartiles show that shocks usually dissipate within the simulated window. However, the upper quartiles are more persistent, and hump shaped in some cases, again illustrating the effects of the solution's non-linearity.

Given their high weights, it is unsurprising that events in Portugal and Spain have a larger impact on the other countries than events elsewhere. Greece and Portugal appear the most

³⁹For precision, data is created at an intraday level, with 10 observations per day. To draw the next intraday observation, we assume that the event intensity is constant over the interval, and that all events happen simultaneously at the end of the interval. Under this approximation, stepping the simulation forward requires one draw from a Poisson (to determine how many events occurred), and one draw from a Gamma (to determine the cumulated jump).

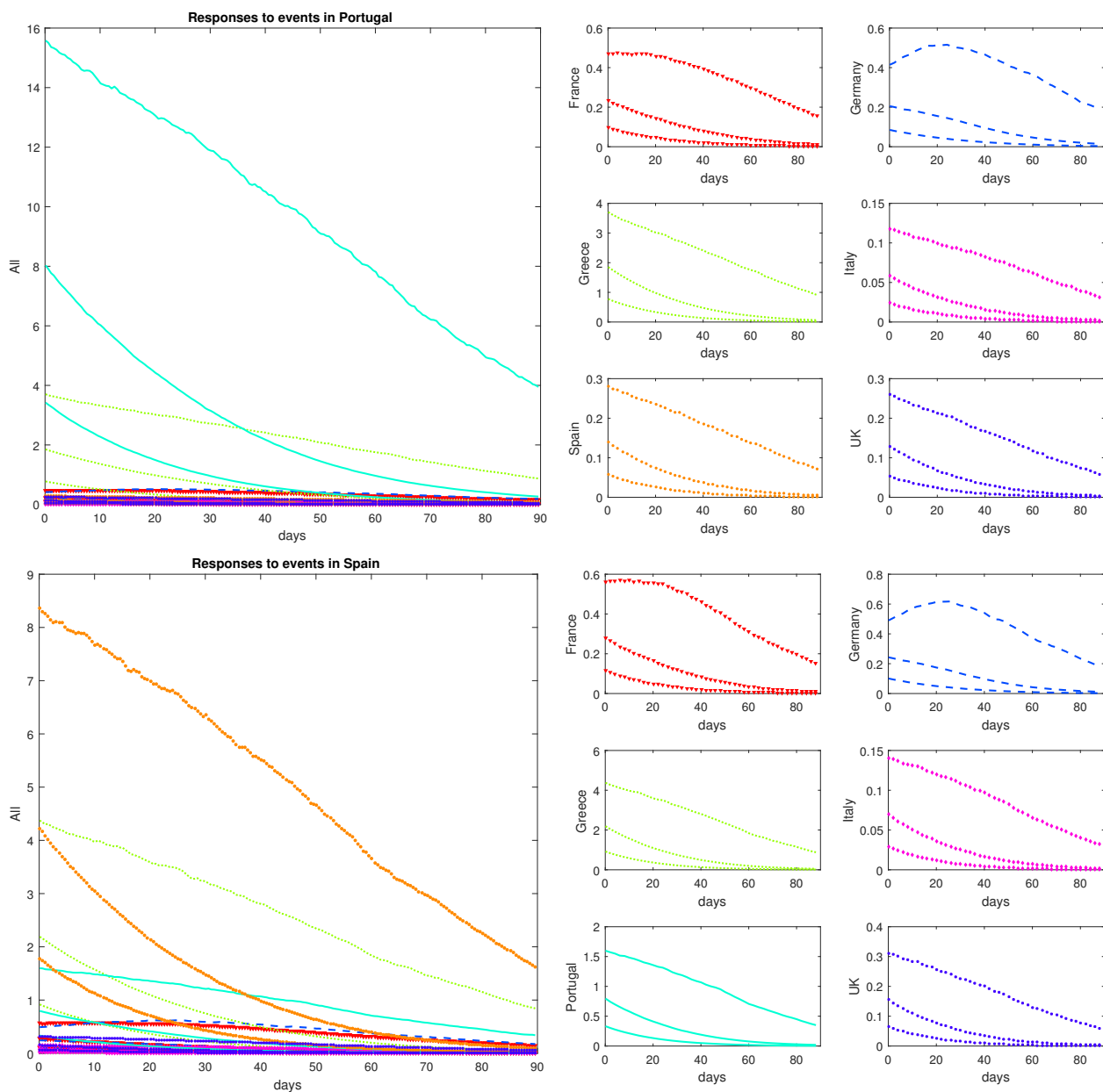


Figure 4: Percentage effect of one event at the start of the sample in Portugal (upper panel) and Spain (lower panel) on the CDS spreads of all countries (bigger figure) and the rest of countries (smaller figures)

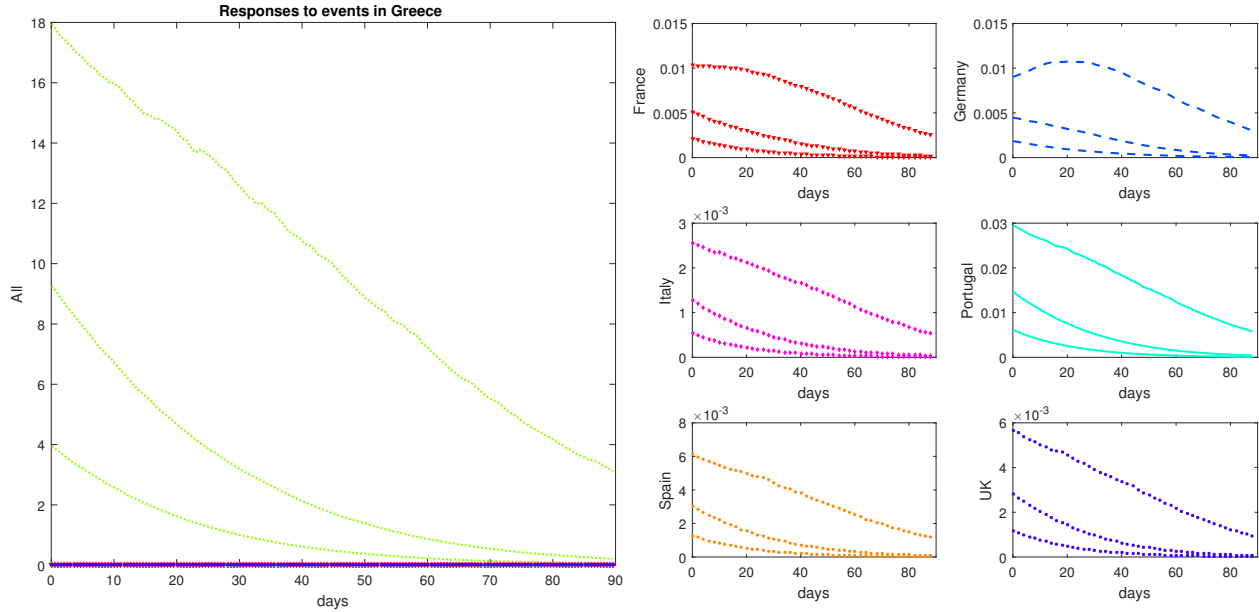


Figure 5: Percentage effect of one event at the start of the sample in Greece on the CDS spreads of all countries (bigger figure) and the rest of countries (smaller figures)

sensitive to these events, followed by France and Germany. The sensitivity of Greece and Portugal is natural given their precarious status. The impact on France and Germany is also quite natural, given that these nations lend most to the other countries in our sample.⁴⁰

6. CONCLUSION

In this paper, we model the intensities of sovereign credit-related events for a group of seven EU countries as a multi-variate self-exciting marked point process with both country-specific and market components. We allow each country to respond differently to both the level and the change in this market component. This structure permits feedback between risks in different countries, allowing us to discover evidence of substantial contagion between countries. While evidence of contagion has been found by other authors,⁴¹ our results on the

⁴⁰This statement is supported by the data used in Section 5.1 for the network analysis.

⁴¹See Augustin (2014) for a review of the sovereign CDS literature.

sources and propagation of credit risk are novel.

We find that Portugal and Spain are the chief sources of credit risk for Europe in this period, with Spain being particularly important. This is in line with the great damage that would be inflicted upon the European economy were a larger nation such as Spain to default. We also find a statistically significant, but small role for Greece as a source of credit risk. As the situation in Greece was already well understood during the period, it appears that investors' focus was shifted toward future dangers, of which the ones from Portugal and Spain were more imminent.

Our impulse response analysis shows that Greece and Portugal are particularly responsive to shocks from elsewhere. Indeed, unlike some other work in the literature,⁴² our findings portray Greece mostly as a receiver of spillover risk, rather than as a source of such risk for other countries. This reflects Greece's precarious state in the period. In quantifying these effects, we find that Greece would have been around 35% less likely to default during the period we consider had Portugal and Spain not been in trouble at the same time.

To capture the rich dynamics of sovereign CDS over the period, our econometric model has substantial flexibility. We allow for random sized jumps, and we do not impose stationarity of the intensities of credit-related events. Moreover, these intensities receive asymmetric feedback both from contemporaneous shocks, as well as from the intensity levels in other countries. In estimation, we use a novel technique to back out implied cumulated shocks to the intensities of credit-related events, which permits a maximum-likelihood approach.

Our modelling and estimation choices, accompanied by the use of a global optimisation algorithm, enable us to estimate the many parameters of this model on a large data set. To the best of our knowledge, this is one of the hardest estimation problems dealt with in the

⁴²See e.g. Ang and Longstaff (2013) and Ait-Sahalia et al. (2014).

credit risk literature.

REFERENCES

- Acharya, V., I. Drechsler, and P. Schnabl (2014). A pyrrhic victory? Bank bailouts and sovereign credit risk. *The Journal of Finance* 69(6), 2689–2739.
- Aït-Sahalia, Y., R. J. Laeven, and L. Pelizzon (2014). Mutual excitation in Eurozone sovereign CDS. *Journal of Econometrics* 183(2), 151 – 167.
- Ang, A. and F. A. Longstaff (2013). Systemic sovereign credit risk: Lessons from the U.S. and Europe. *Journal of Monetary Economics* 60(5), 493 – 510.
- Augustin, P. (2014). Sovereign credit default swap premia. Technical report, Forthcoming, Journal of Investment Management.
- Augustin, P., M. Chernov, and D. Song (2018, April). Sovereign credit risk and exchange rates: Evidence from CDS quanto spreads. Working Paper 24506, National Bureau of Economic Research.
- 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. S. Goldstein, and J. Helwege (2015). On bounding credit-event risk premia. *The Review of Financial Studies* 28(9), 2608–2642.
- Benzoni, L., P. Collin-Dufresne, R. S. Goldstein, and J. Helwege (2015). Modeling credit contagion via the updating of fragile beliefs. *The Review of Financial Studies* 28(7), 1960–2008.
- Bollerslev, T. and J. M. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11(2), 143–172.
- Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics* 141(2), 876 – 912.
- Carr, P. and L. Wu (2004). Time-changed Lévy processes and option pricing. *Journal of Financial Economics* 71(1), 113 – 141.
- Da Fonseca, J. and R. Zaatour (2015). Clustering and mean reversion in a Hawkes microstructure model. *Journal of Futures Markets* 35(9), 813–838.
- Daley, D. J. and D. Vere-Jones (2003). *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods, 2nd ed.* New-York: Springer.

- De Grauwe, P. and Y. Ji (2013). Self-fulfilling crises in the Eurozone: An empirical test. *Journal of International Money and Finance* 34, 15 – 36.
- Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68(6), 1343–1376.
- Duffie, D., L. H. Pedersen, and K. J. Singleton (2003). Modeling sovereign yield spreads: A case study of Russian debt. *The Journal of Finance* 58(1), 119–159.
- Duffie, D. and K. J. Singleton (2003). *Credit Risk*. Princeton, NJ: Princeton University Press.
- Edwards, S. (1984). LDC foreign borrowing and default risk: An empirical investigation, 1976-80. *The American Economic Review* 74(4), 726–734.
- Eichengreen, B. and A. Mody (2000). *What Explains Changing Spreads on Emerging Market Debt?*, pp. 107–134. University of Chicago Press.
- Errais, E., K. Giesecke, and L. R. Goldberg (2010). Affine point processes and portfolio credit risk. *SIAM Journal of Financial Mathematics* 1(1), 642–665.
- Geman, H., P. Carr, D. B. Madan, and M. Yor (2002). The fine structure of asset returns: An empirical investigation. *Journal of Business* 75(2), 305 – 332.
- Gouriéroux, C., A. Monfort, and J. Renne (2014). Pricing default events: Surprise, exogeneity and contagion. *Journal of Econometrics* 182(2), 397 – 411.
- Hansen, N. and A. Ostermeier (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation* 9(2), 159–195.
- Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58(1), 83–90.
- Kamin, S. B. and K. von Kleist (1999). The evolution and determinants of emerging markets credit spreads in the 1990s. Technical Report 68, BIS.
- Kleinberg, J. M. (1999a). Authoritative sources in a hyperlinked environment. *Journal of the ACM* 46(5), 604–632.
- Kleinberg, J. M. (1999b). Hubs, authorities, and communities. *ACM computing surveys (CSUR)* 31(4es), 5.
- Longstaff, F., S. Mithal, and E. Neis (2005). Corporate yield spreads: Default risk or liquidity? new evidence from the credit default swap market. *Journal of Finance* 60(5), 2213–2253.
- Longstaff, F., J. Pan, L. Pedersen, and K. Singleton (2011a). How sovereign is sovereign credit risk? *American*

Economic Journal: Macroeconomics 3(2), 75–103.

Longstaff, F. A., J. Pan, L. H. Pedersen, and K. J. Singleton (2011b, April). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics* 3(2), 75–103.

Monfort, A., F. Pegoraro, J.-P. Renne, and G. Roussellet (2017a). Affine modelling of credit risk, pricing of credit events and contagion. Working paper.

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

Pan, J. (2002). The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics* 63(1), 3 – 50.

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

Pastorello, S., V. Patilea, and E. Renault (2003). Iterative and recursive estimation in structural nonadaptive models. *Journal of Business & Economic Statistics* 21(4), 449–509.

Remolona, E. M., M. Scatigna, and E. Wu (2008). The dynamic pricing of sovereign risk in emerging markets: fundamentals and risk aversion. *Journal of Fixed Income* 17(4), 57–71.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics* 50(3), 665 – 690.

Stracca, L. (2015). Our currency, your problem? The global effects of the euro debt crisis. *European Economic Review* 74, 1 – 13.

FOR ONLINE PUBLICATION ONLY

APPENDIX A: CONDITIONS FOR POSITIVITY

We begin with a helpful Lemma, before introducing the conditions we impose throughout in order to ensure positivity.

LEMMA A.1 *If $\tilde{\alpha}$ is a real square matrix for which all off-diagonal elements are non-positive, then all elements of $e^{-\tilde{\alpha}}$ are real and non-negative.*

PROOF: Let $A := -\tilde{\alpha}$, so matrix A has non-negative off-diagonal entries. Let $B := A + cI$, where c is a positive real scalar such that the diagonal elements of B are non-negative.

From the definition of the matrix exponential, we have that $e^B = I + B + \frac{1}{2!}B^2 + \dots$. As B has only non-negative entries, all powers of B also have only non-negative entries, and consequently e^B has only non-negative entries.

Since the identity matrix commutes with anything, by the standard properties of matrix exponentials, we have that $e^A = e^B e^{-cI} = e^{-c} e^B$. As e^{-c} is a positive scalar and e^B has only non-negative elements, so does e^A . Q.E.D.

We impose the following conditions:

CONDITION A.1 *All off-diagonal elements of $\tilde{\alpha}$ are non-positive. All elements of $\check{\beta}$ are positive.*

This condition is sufficient for positivity of λ , by Lemma A.1, and the solution to the Hawkes process, equation (1).

APPENDIX B: CONDITIONAL AND UNCONDITIONAL MOMENTS

$X_t = (N_t, J_t, \lambda_t)$ is a Markov process defined on the space $D := \mathbb{N}^K \times \mathbb{R}_{\geq 0}^K \times \mathbb{R}_{\geq 0}^K$. Let $d := 3K$ be the overall dimension of X_t . Let $f: D \rightarrow \mathbb{R}$ be a function of X_t and let ∇f be its gradient (a row vector). The infinitesimal generator for $f(x)$, defined as $Af(x) = \lim_{t \downarrow 0} \frac{\mathbb{E}(f(X_{t+dt}) - f(x))}{dt}$, is given by:

$$(14) \quad Af(x) = \begin{bmatrix} 0_{K \times 1} \\ 0_{K \times 1} \\ \check{\alpha}(\lambda_\infty - \lambda) \end{bmatrix}^\top \nabla^\top f + \sum_{i=1}^K \lambda_{i,\cdot} \int_{\mathbb{R}_{\geq 0}^d} [f(x + Z) - f(x)] d\nu^i(Z),$$

where $x := [\lambda^\top, J^\top, N^\top]^\top$, $Z := [Z_1, \dots, Z_d]^\top$ is the generic jump in D for the state process X , and $\nu^i(Z)$ is the cumulative distribution function corresponding to the i -th jump. When a credit-related event occurs in country i , $N_{i,\cdot}$ increases by 1, $J_{i,\cdot}$ increases by $z_{i,\cdot} \sim \text{Exp}(1)$ and finally, because of the cross-excitation in the model, the intensities of all countries will increase proportionally with $z_{i,\cdot}$. Thus, $\nu^i(Z)$ is given by:

$$\begin{aligned} d\nu^i \left(\left[Z_1 \cdots Z_i \cdots Z_K \quad Z_{K+1} \cdots Z_{2K} \quad Z_{2K+1} \cdots Z_d \right]^\top \right) = \\ \left[\delta(Z_1 - 0) \cdots \delta(Z_i - 1) \cdots \delta(Z_K - 0) \quad \delta(Z_{K+1} - 0) \cdots \exp(-Z_{K+i}) \cdots \delta(Z_{2K} - 0) \right. \\ \left. \delta(Z_{2K+1} - Z_{K+i}\check{\beta}_{1i}) \cdots \delta(Z_{2k+i} - Z_{K+i}\check{\beta}_{ii}) \cdots \delta(Z_d - Z_{K+i}\check{\beta}_{Ki}) \right]^\top \circ dZ, \end{aligned}$$

where $Z_{K+i} = z_{i,\cdot}$, $\delta(\cdot)$ is the Dirac delta function and $\check{\beta}_{ij}$, denotes the i, j^{th} element of the $\check{\beta}$ matrix.

To further help understand equation (14), we note that the first term of the sum in that equation is given

$$\begin{aligned} \text{by:} \\ \lambda_{1,\cdot} \int_{\mathbb{R}_{\geq 0}^d} \left[f \left(\left[N_{1,\cdot} + 1 \quad \cdots \quad N_{K,\cdot} + 0 \quad J_{1,\cdot} + z_{1,\cdot} \quad \cdots \quad J_{K,\cdot} + 0 \quad \lambda_{1,\cdot} + \check{\beta}_{11}z_{1,\cdot} \quad \cdots \quad \lambda_{K,\cdot} + \check{\beta}_{K1}z_{1,\cdot} \right]^\top \right) - \right. \\ \left. f \left(\left[N_{1,\cdot} \quad \cdots \quad N_{K,\cdot} \quad J_{1,\cdot} \quad \cdots \quad J_{K,\cdot} \quad \lambda_{1,\cdot} \quad \cdots \quad \lambda_{K,\cdot} \right]^\top \right) \right] d\nu^1(Z). \end{aligned}$$

LEMMA B.1 *Conditional expectation.* *The conditional expected intensity satisfies the following ordinary differential equation (ODE hereafter):⁴³*

$$d\mathbb{E}\lambda_t = \check{\alpha}(\lambda_\infty - \mathbb{E}\lambda_t) dt + \check{\beta}\mathbb{E}\lambda_t dt,$$

with initial value λ_0 and solution:

$$\mathbb{E}\lambda_t = e^{(\check{\beta}-\check{\alpha})t}\lambda_0 - (\check{\beta} - \check{\alpha})^{-1}\check{\alpha}\lambda_\infty + (\check{\beta} - \check{\alpha})^{-1}e^{(\check{\beta}-\check{\alpha})t}\check{\alpha}\lambda_\infty.$$

⁴³Errais et al. (2010) use the same technique based on the infinitesimal generator to derive moments for a univariate marked self exciting point process.

PROOF: The proof is a simple extension to marked processes of Lemma 1 on page 818 in Da Fonseca and Zaatour (2015). Q.E.D.

LEMMA B.2 *Unconditional expectation.* *If it exists, the long-run expected intensity is given by:*

$$\lim_{n \rightarrow \infty} \mathbb{E}\lambda_t = \bar{\lambda}_\infty = -(\check{\beta} - \check{\alpha})^{-1} \check{\alpha} \lambda_\infty.$$

PROOF: Apply Lemma B.1 with $t \rightarrow \infty$. Q.E.D.

For use in the following, we define $m_1 = \mathbb{E}z_{i,t} = 1$ and $m_2 = \mathbb{E}z_{i,t}^2 = 2$ (for any t and i).

LEMMA B.3 *Conditional second moment.* *The conditional second moment of the intensity, $\mathbb{E}(\lambda_t \lambda_t^\top)$, satisfies the following ODE:*

$$(15) \quad \frac{\mathbb{E}\lambda_t \lambda_t^\top}{dt} = \check{\alpha} \lambda_\infty \mathbb{E}\lambda_t^\top + \mathbb{E}\lambda_t \lambda_\infty^\top \check{\alpha}^\top + (m_1 \check{\beta} - \check{\alpha}) \mathbb{E}(\lambda_t \lambda_t^\top) + \mathbb{E}(\lambda_t \lambda_t^\top) (m_1 \check{\beta} - \check{\alpha})^\top + m_2 \left\{ \check{\beta} \text{diag} [\mathbb{E}\lambda_t] \check{\beta}^\top \right\}.$$

PROOF: Let $f(x) = \lambda_{1,\cdot}^2$. Then:

$$\begin{aligned} Af(x) &= \begin{bmatrix} 0_{K \times 1} \\ 0_{K \times 1} \\ 0_{K \times 1} \\ \check{\alpha}(\lambda_\infty - \lambda) \end{bmatrix}^\top \begin{bmatrix} 0_{K \times 1} \\ 0_{K \times 1} \\ 2\lambda_{1,\cdot} \\ \vdots \\ 0 \end{bmatrix} + \lambda_{1,\cdot} \int_{\mathbb{R}_{\geq 0}^d} [(\lambda_{1,\cdot} + \check{\beta}_{11} z_{1,\cdot})^2 - \lambda_{1,\cdot}^2] d\nu^1(Z) + \dots \\ &+ \lambda_{K,\cdot} \int_{\mathbb{R}_{\geq 0}^d} [(\lambda_{1,\cdot} + \check{\beta}_{1K} z_{K,\cdot})^2 - \lambda_{1,\cdot}^2] d\nu^K(Z) \\ &= 2\lambda_{1,\cdot} \check{\alpha}_1^\top (\lambda_\infty - \lambda) + 2m_1 \lambda_{1,\cdot} \sum_i \check{\beta}_{1i} \lambda_{i,\cdot} + m_2 \sum_i \check{\beta}_{1i}^2 \lambda_{i,\cdot}. \end{aligned}$$

Using the martingale property of $f(X_t) - f(X_0) - \int_0^t Af(X_s) ds$, we follow Da Fonseca and Zaatour (2015) in using the following formula to compute conditional expectations:

$$\mathbb{E}[f(X_t)] = f(X_0) + \mathbb{E} \left[\int_0^t Af(X_s) ds \right].$$

Let $\check{\alpha}_i$ and $\check{\beta}_i$ denote the i -th lines of matrices $\check{\alpha}$ and $\check{\beta}$. We have:

$$\mathbb{E}[\lambda_{1,t}^2] = \lambda_{1,0}^2 + \mathbb{E} \left[\int_0^t Af(X_s) ds \right] \implies$$

$$\begin{aligned} \frac{d\mathbb{E}[\lambda_{1,t}^2]}{dt} &= 2\check{\alpha}_1^\top \lambda_\infty \mathbb{E}(\lambda_{1,t}) - 2\check{\alpha}_1^\top \mathbb{E}(\lambda_{1,t} \lambda_t) + 2m_1 \check{\beta}_1^\top \mathbb{E}(\lambda_{1,t} \lambda_t) + m_2 \check{\beta}_1^\top \text{diag}[\mathbb{E}\lambda_t] \check{\beta}_1. \\ &= 2\check{\alpha}_1^\top \lambda_\infty \mathbb{E}(\lambda_{1,t}) + 2(\check{\beta}_1 m_1 - \check{\alpha}_1)^\top \mathbb{E}(\lambda_{1,t} \lambda_t) + m_2 \check{\beta}_1^\top \text{diag}[\mathbb{E}\lambda_t] \check{\beta}_1. \end{aligned}$$

Let $f(x) = \lambda_{1,\cdot}, \lambda_{2,\cdot}, \dots$. Then:

$$\begin{aligned}
 A f(x) &= \begin{bmatrix} 0_{K \times 1} \\ 0_{K \times 1} \\ \check{\alpha}(\lambda_\infty - \lambda_\cdot) \end{bmatrix}^\top \begin{bmatrix} 0_{K \times 1} \\ 0_{K \times 1} \\ \lambda_{2,\cdot} \\ \lambda_{1,\cdot} \\ \vdots \\ 0 \end{bmatrix} + \lambda_{1,\cdot} \int_{\mathbb{R}_{\geq 0}^d} [(\lambda_{1,\cdot} + \check{\beta}_{11} z_{1,\cdot}) (\lambda_{2,\cdot} + \check{\beta}_{21} z_{1,\cdot}) - \lambda_{1,\cdot} \lambda_{2,\cdot}] d\nu^1(Z) \\
 &+ \dots + \lambda_{K,\cdot} \int_{\mathbb{R}_{\geq 0}^d} [(\lambda_{1,\cdot} + \check{\beta}_{1K} z_{K,\cdot}) (\lambda_{2,\cdot} + \check{\beta}_{2K} z_{K,\cdot}) - \lambda_{1,\cdot} \lambda_{2,\cdot}] d\nu^K(Z) \\
 &= \lambda_{2,\cdot} \check{\alpha}_1^\top (\lambda_\infty - \lambda_\cdot) + \lambda_{1,\cdot} \check{\alpha}_2^\top (\lambda_\infty - \lambda_\cdot) + m_1 \lambda_{1,\cdot} \sum_i \check{\beta}_{2i} \lambda_{i,\cdot} + m_1 \lambda_{2,\cdot} \sum_i \check{\beta}_{1i} \lambda_{i,\cdot} + m_2 \sum_i \check{\beta}_{1i} \check{\beta}_{2i} \lambda_{i,\cdot}.
 \end{aligned}$$

In terms of expectations, we have:

$$\begin{aligned}
 \frac{d \mathbb{E}[\lambda_{1,t} \lambda_{2,t}]}{dt} &= \check{\alpha}_1^\top \lambda_\infty \mathbb{E}(\lambda_{2,t}) + \check{\alpha}_2^\top \lambda_\infty \mathbb{E}(\lambda_{1,t}) - \check{\alpha}_1^\top \mathbb{E}(\lambda_{2,t} \lambda_t) - \check{\alpha}_2^\top \mathbb{E}(\lambda_{1,t} \lambda_t) \\
 &+ m_1 \check{\beta}_1^\top \mathbb{E}(\lambda_{2,t} \lambda_t) + m_1 \check{\beta}_2^\top \mathbb{E}(\lambda_{1,t} \lambda_t) + m_2 \check{\beta}_1^\top \text{diag}[\mathbb{E} \lambda_t] \check{\beta}_2. \\
 &= \check{\alpha}_1^\top \lambda_\infty \mathbb{E}(\lambda_{2,t}) + \check{\alpha}_2^\top \lambda_\infty \mathbb{E}(\lambda_{1,t}) + (\check{\beta}_1 \cdot m_1 - \check{\alpha}_1)^\top \mathbb{E}(\lambda_{2,t} \lambda_t) \\
 &+ (\check{\beta}_2 \cdot m_1 - \check{\alpha}_2)^\top \mathbb{E}(\lambda_{1,t} \lambda_t) + m_2 \check{\beta}_1^\top \text{diag}[\mathbb{E} \lambda_t] \check{\beta}_2.
 \end{aligned}$$

Given the above ODEs for $\mathbb{E}(\lambda_{1,t})$ and $\mathbb{E}[\lambda_{1,t} \lambda_{2,t}]$, the generalizations stated in Lemma B.3 can be obtained. Q.E.D.

LEMMA B.4 Long term covariance. *If it exists, the long term covariance matrix of λ_t , defined as $\bar{\Lambda}_\infty = \lim_{n \rightarrow \infty} \mathbb{E}(\lambda_t \lambda_t^\top) - \bar{\lambda}_\infty \bar{\lambda}_\infty^\top$, solves the following algebraic equation:*

$$(m_1 \check{\beta} - \check{\alpha}) \bar{\Lambda}_\infty + \bar{\Lambda}_\infty (m_1 \check{\beta} - \check{\alpha})^\top + m_2 \check{\beta} \text{diag}(\bar{\lambda}_\infty) \check{\beta}^\top = 0.$$

PROOF: Apply Lemma B.3 with $t \rightarrow \infty$.

Q.E.D.

APPENDIX C: TRANSFORMS

We maintain the set-up and definitions from Appendix B.

PROPOSITION 1 *For each $i \in \{1, \dots, K\}$, and $T \geq t$,*

$$(16) \quad \mathbb{E} \left[(1 - \gamma_i)^{N_{i,T}} \middle| \mathcal{F}_t \right] = \exp \left(a^i(t) + b_i^i(t) N_{i,t} + b_{2K+1}^i(t) \lambda_{1,t} + \dots + b_{3K}^i(t) \lambda_{K,t} \right)$$

where for $i \in \{1, \dots, K\}$, the coefficients $a^i(t)$ and $\mathbf{b}^i(t) = (b_1^i(t), \dots, b_{3K}^i(t))^\top$ are solutions to the ODEs given below.

a.

$$\dot{\mathbf{b}}_{2K+1:3K}^i(t) = \tilde{\alpha}^\top \mathbf{b}_{2K+1:3K}^i(t) + \mathbf{1}_{K \times 1} - \theta(\mathbf{b}^i(t)),$$

and $\dot{\mathbf{b}}_{1:2K}^i(t) = 0$, where for $j_1 < j_2$, $\mathbf{b}_{j_1:j_2}^i(t) = (b_{j_1}^i(t), \dots, b_{j_2}^i(t))^\top$, and where $\theta : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}^K$ is defined by:

$$\theta_i(c) = \int_{\mathbb{R}_{\geq 0}^d} e^{(c \cdot Z)} d\nu^i(Z), \quad i = 1, \dots, K,$$

for all $c \in \mathbb{R}^d, i \in \{1, \dots, K\}$.

This ODE has the terminal condition $\mathbf{b}^i(T) = (0, \dots, \log(1 - \gamma_i), \dots, 0, 0, \dots, 0, 0, \dots, 0)^\top$, for all $i \in \{1, \dots, K\}$.

b.

$$\dot{a}^i(t) = -(\tilde{\alpha} \lambda_\infty) \cdot \mathbf{b}_{2K+1:3K}^i(t),$$

with terminal condition $a^i(T) = 0$, for $i \in \{1, \dots, K\}$.

PROOF: The result is obtained through a direct application of Proposition 1 and Appendix B in Duffie et al. (2000) or the results in Section 4 of Errais et al. (2010). Using the notations in the aforementioned paper, we have the drift of X_t is given by $\mu(X_t) = K_0 + K_1 x$ where the two matrices are defined as follows: $K_0 := \begin{pmatrix} 0_{K \times 1} \\ 0_{K \times 1} \\ \check{\alpha} \lambda_\infty \end{pmatrix}$; $K_1 := \begin{pmatrix} 0_{K \times K} & 0_{K \times K} & 0_{K \times K} \\ 0_{K \times K} & 0_{K \times K} & 0_{K \times K} \\ 0_{K \times K} & 0_{K \times K} & -\check{\alpha} \end{pmatrix}$. As we have no diffusion component in our process, $H_0 \equiv 0$ and $H_1 \equiv 0$. Moreover, as we assume independence between the state vector and the discount factor: $\rho_0 = 0$; $\rho_1 = 0_{d \times 1}$. (Discounting will drop out of the expectation we wish to evaluate.) We have:

$$\begin{aligned} dX_t &= \begin{bmatrix} dN_{1,t} & \cdots & dN_{K,t} & dJ_{1,t} & \cdots & dJ_{K,t} & d\lambda_{1,t} & \cdots & d\lambda_{K,t} \end{bmatrix} \\ &= \begin{bmatrix} 0_{2K \times 1} \\ \check{\alpha}(\lambda_\infty - \lambda_t) \end{bmatrix} dt + \begin{bmatrix} 1 & \cdots & 0 & Z_{K+1} & \cdots & 0 & Z_{K+1}\check{\beta}_{11} & \cdots & Z_{K+1}\check{\beta}_{K1} \end{bmatrix} dN_{1,t} + \cdots \\ &+ \begin{bmatrix} 0 & \cdots & 1 & 0 & \cdots & Z_{2K} & Z_{2K}\check{\beta}_{1K} & \cdots & Z_{2K}\check{\beta}_{KK} \end{bmatrix} dN_{K,t} \end{aligned}$$

The default intensities satisfy the equations $\lambda_i(x) = l_0^i + l_1^i x$, for $i \in \{1, \dots, K\}$, where $l_0^i = 0$ for all $i \in \{1, \dots, K\}$ and:

$$\begin{aligned} l_1^1 &= (0, \dots, 0, 0, \dots, 0, 1, \dots, 0)^\top \\ &\vdots \\ l_1^K &= (0, \dots, 0, 0, \dots, 0, 0, \dots, 1)^\top. \end{aligned}$$

We proceed by integrating terms in the jump transform one by one:

$$\begin{aligned}
\theta_i(c) &= \int_{\mathbb{R}_{\geq 0}^d} \exp\left(\sum_{i=1}^d c_i Z_i\right) d\nu^i(Z) \\
&= \int_0^\infty \exp(c_i + c_{K+i} z_{i,\cdot} + c_{2K+1} \check{\beta}_{1i} z_{i,\cdot} + \dots + c_{3K} \check{\beta}_{Ki} z_{i,\cdot}) \exp(-z_{i,\cdot}) dz_{i,\cdot} \\
&= \frac{\exp(c_i + c_{K+i} z_{i,\cdot} + c_{2K+1} \check{\beta}_{1i} z_{i,\cdot} + \dots + c_{3K} \check{\beta}_{Ki} z_{i,\cdot} - z_{i,\cdot})}{c_{K+i} + c_{2K+1} \check{\beta}_{1i} + \dots + c_{3K} \check{\beta}_{Ki} - 1} \Bigg|_{z_{i,\cdot}=0}^\infty \\
&= \frac{\exp(c_i)}{\max\{0, 1 - c_{K+i} - c_{2K+1} \check{\beta}_{1i} - \dots - c_{3K} \check{\beta}_{Ki}\}}.
\end{aligned}$$

Applying Proposition 1 in Duffie et al. (2000) leads to:

$$\dot{\mathbf{b}}^i(t) = -K_1^\top \mathbf{b}^i(t) - \sum_{j=1}^K l_1^j [\theta_j(\mathbf{b}^i(t)) - 1],$$

with $\mathbf{b}^i(T) = (0, \dots, \log(1 - \gamma_i), \dots, 0, 0, \dots, 0, 0, \dots, 0)^\top$. So:

$$\dot{\mathbf{b}}_{2K+1:3K}^i(t) = \check{\alpha}^\top \mathbf{b}_{2K+1:3K}^i(t) + \mathbf{1}_{K \times 1} - \theta(\mathbf{b}^i(t))$$

and $\dot{\mathbf{b}}_{1:2K}^i(t) = 0, \forall t$.

To show that this ODE is well-defined despite the pole in θ , first define $\psi : \mathbb{R}^K \rightarrow \mathbb{R}^k$ by $\psi(c) = -\log(\mathbf{1}_{K \times 1} - \check{\beta}^\top c)$, for all $c \in \mathbb{R}^K$, where the logarithm is elementwise. Now, note that $\theta(\mathbf{b}^i(t)) = (\mathbf{1}_{K \times 1} - \gamma_i e_i) \circ \exp \psi(\mathbf{b}_{2K+1:3K}^i(t))$, and that $\psi^{-1}(d) = \check{\beta}^{-\top} (\mathbf{1}_{K \times 1} - \exp(-d))$ for all $d \in \mathbb{R}^K$, where the exponentiation is element-wise in both cases, and where e_i is the i^{th} column of the identity matrix. Hence:

$$\begin{aligned}
\frac{d}{dt} \psi(\mathbf{b}_{2K+1:3K}^i(t)) &= \exp \psi(\mathbf{b}_{2K+1:3K}^i(t)) \\
&\circ \check{\beta}^\top \left[\check{\alpha}^\top \check{\beta}^{-\top} (\mathbf{1}_{K \times 1} - \exp(-\psi(\mathbf{b}_{2K+1:3K}^i(t)))) + \mathbf{1}_{K \times 1} - (\mathbf{1}_{K \times 1} - \gamma_i e_i) \circ \exp \psi(\mathbf{b}_{2K+1:3K}^i(t)) \right].
\end{aligned}$$

This gives an ODE in $\psi(\mathbf{b}_{2K+1:3K}^i(t))$ without poles on the right hand side. From the solution of this ODE, we can then back-out a solution for $\mathbf{b}^i(t)$.

Moreover, we have:

$$\dot{a}^i(t) = -K_0 \cdot \mathbf{b}^i(t) = -(\check{\alpha}\lambda_\infty) \cdot \mathbf{b}_{2K+1:3K}^i(t),$$

where $a_i(T) = 0$ for $i \in \{1, \dots, K\}$.

Q.E.D.

PROPOSITION 2 For each $i \in \{1, \dots, K\}$ and $T \geq t$,

$$\begin{aligned} \mathbb{E} \left[\gamma_i \lambda_{i,T} (1 - \gamma_i)^{N_{i,T}} \middle| \mathcal{F}_t \right] &= \exp \left(a^i(t) + b_i^i(t) N_{i,t} + b_{2K+1}^i(t) \lambda_{1,t} + \dots + b_{3K}^i(t) \lambda_{K,t} \right) \\ &\quad \left(A^i(t) + B_{2K+1}^i(t) \lambda_{1,t} + \dots + B_{3K}^i(t) \lambda_{K,t} \right) \end{aligned}$$

where for $i \in \{1, \dots, K\}$, the coefficients $A^i(t)$ and $\mathbf{B}^i(t) = (B_1^i(t), \dots, B_{3K}^i(t))^\top$ are solutions to the ODEs given below.

a.

$$-\dot{\mathbf{B}}_{2K+1:3K}^i(t) = -\check{\alpha}^\top \mathbf{B}_{2K+1:3K}^i(t) + \nabla \theta(\mathbf{b}^i(t)) \mathbf{B}^i(t),$$

and $\dot{\mathbf{B}}_{1:2K}^i(t) = 0$, where for $j_1 < j_2$, $\mathbf{B}_{j_1:j_2}^i(t) = (B_{j_1}^i(t), \dots, B_{j_2}^i(t))^\top$, and where $\nabla \theta(\cdot)$ is the Jacobian of $\theta(c)$.

This ODE has the terminal condition $\mathbf{B}^i(T) = (0, \dots, 0, 0, \dots, 0, 0, \dots, \gamma_i, \dots, 0)^\top$, for all $i \in \{1, \dots, K\}$.

b.

$$-\dot{A}^i(t) = (\check{\alpha}\lambda_\infty) \cdot \mathbf{B}_{2K+1:3K}^i(t),$$

with $\dot{A}^i(T) = 0$.

PROOF: The result is obtained through a direct application of Proposition 3 and Appendix B in Duffie et al. (2000) or the results in Section 4 of Errais et al. (2010). In addition to all the notation introduced in the proof of Proposition 1, we define $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}^K$ by:

$$\varphi_i(c) = \frac{1}{1 - c_{K+i} - c_{2K+1}\check{\beta}_{1i} - \cdots - c_{3K}\check{\beta}_{Ki}},$$

for all $c \in \mathbb{R}^d$ and $i \in \{1, \dots, K\}$. Using this, we have that the Jacobian of θ is given by:

$$\nabla\theta(c) = \left[\text{diag } \theta(c) \quad \text{diag } (\theta(c) \circ \varphi(c)) \quad \text{diag } (\theta(c) \circ \varphi(c))\check{\beta}^\top \right].$$

Applying Proposition 3 in Duffie et al. (2000), we get:

$$-\dot{\mathbf{B}}^i(t) = K_1^\top \mathbf{B}^i(t) + \sum_{j=1}^K l_1^j \nabla\theta_j(\mathbf{b}^i(t)) \mathbf{B}^i(t),$$

with $\mathbf{B}^i(t) = (0, \dots, 0, 0, \dots, 0, 0, \dots, \gamma_i, \dots, 0)^\top$. So:

$$-\dot{\mathbf{B}}_{2K+1:3K}^i(t) = -\check{\alpha}^\top \mathbf{B}_{2K+1:3K}^i(t) + \nabla\theta(\mathbf{b}^i(t)) \mathbf{B}^i(t),$$

and $\dot{\mathbf{B}}_{1:2K}^i(t) = 0$. Moreover, we have: $-\dot{A}^i(t) = K_0 \cdot \mathbf{B}^i(t) = (\check{\alpha}\lambda_\infty) \cdot \mathbf{B}_{2K+1:3K}^i(t)$, with $\dot{A}^i(t) = 0$.

Q.E.D.

APPENDIX D: DERIVING THE LIKELIHOOD OF OBSERVING A CERTAIN JUMP SIZE AND NO DEFAULT OVER A TIME INTERVAL

LEMMA D.1 *Suppose $d\tilde{J}_t = \tilde{z} d\tilde{N}_t$ where \tilde{N}_t is a univariate Poisson process with constant intensity $\tilde{\lambda}$, and where $\tilde{z} \sim \text{Exp}(1)$. Assume $\tilde{J}_0 = 0$. Then the p.d.f. of \tilde{J}_Δ is given by:*

$$f(\tilde{J}_\Delta) = \Delta \tilde{\lambda} \frac{\mathcal{I}_1\left(2\sqrt{\Delta \tilde{\lambda} \tilde{J}_\Delta}\right)}{e^{\Delta \tilde{\lambda} + \tilde{J}_\Delta} \sqrt{\Delta \tilde{\lambda} \tilde{J}_\Delta}}.$$

PROOF: Note that \tilde{N}_Δ is Poisson distributed with parameter $\Delta \tilde{\lambda}$. Furthermore, conditional on having k jumps, \tilde{J}_Δ is Gamma distributed with shape parameter k and scale parameter 1. Therefore:

$$\begin{aligned} f(\tilde{J}_\Delta) &= \sum_{k=0}^{\infty} \Pr(\tilde{N}_\Delta = k) f(\tilde{J}_\Delta | \tilde{N}_\Delta = k) = \sum_{k=0}^{\infty} \frac{(\Delta \tilde{\lambda})^k e^{-\Delta \tilde{\lambda}}}{k!} \frac{\tilde{J}_\Delta^{k-1} e^{-\tilde{J}_\Delta}}{\Gamma(k)} \\ &= \Delta \tilde{\lambda} \frac{\mathcal{I}_1\left(2\sqrt{\Delta \tilde{\lambda} \tilde{J}_\Delta}\right)}{e^{\Delta \tilde{\lambda} + \tilde{J}_\Delta} \sqrt{\Delta \tilde{\lambda} \tilde{J}_\Delta}}. \end{aligned}$$

Q.E.D.

LEMMA D.2 *Suppose $d\tilde{J}_t = \tilde{z} d\tilde{N}_t$ where \tilde{N}_t is a univariate Poisson process with constant intensity $\tilde{\lambda}$, and where $\tilde{z} \sim \text{Exp}(1)$. Assume $\tilde{J}_0 = 0$. Then for all $\gamma \in [0, 1]$:*

$$\mathbb{E}\left[(1-\gamma)^{\tilde{N}_\Delta} \mid \tilde{J}_\Delta\right] = \frac{\sqrt{1-\gamma} \mathcal{I}_1\left(2\sqrt{\Delta(1-\gamma)\tilde{\lambda}\tilde{J}_\Delta}\right)}{\mathcal{I}_1\left(2\sqrt{\Delta\tilde{\lambda}\tilde{J}_\Delta}\right)}.$$

PROOF:

$$\mathbb{E}\left[(1-\gamma)^{\tilde{N}_\Delta} \mid \tilde{J}_\Delta\right] = \sum_{k=0}^{\infty} (1-\gamma)^k \Pr\left(\tilde{N}_\Delta = k \mid \tilde{J}_\Delta\right) = \sum_{k=0}^{\infty} \frac{(1-\gamma)^k f(\tilde{J}_\Delta | \tilde{N}_\Delta = k) \Pr(\tilde{N}_\Delta = k)}{f(\tilde{J}_\Delta)},$$

where we applied Bayes's theorem.

Moreover, $f(\tilde{J}_\Delta)$ is given in Lemma D.1 above, $\tilde{J}_\Delta | \tilde{N}_\Delta = k \sim \Gamma(k, 1)$ and $\Pr(\tilde{N}_\Delta = k)$ is Poisson with parameter $\Delta\tilde{\lambda}$. Hence, we have:

$$\begin{aligned} \mathbb{E} \left[(1-\gamma)^k | \tilde{J}_\Delta \right] &= \sum_{k=0}^{\infty} (1-\gamma)^k \frac{\tilde{J}_\Delta^{k-1} e^{-\tilde{J}_\Delta}}{\Gamma(k)} \frac{(\Delta\tilde{\lambda})^k e^{-\Delta\tilde{\lambda}}}{k!} \left[\frac{\Delta\tilde{\lambda} \mathcal{I}_1 \left(2\sqrt{\Delta\tilde{\lambda}\tilde{J}_\Delta} \right)}{e^{\Delta\tilde{\lambda} + \tilde{J}_\Delta} \sqrt{\Delta\tilde{\lambda}\tilde{J}_\Delta}} \right]^{-1} \\ &= \frac{\sqrt{1-\gamma} \mathcal{I}_1 \left(2\sqrt{\Delta(1-\gamma)\tilde{\lambda}\tilde{J}_\Delta} \right)}{\mathcal{I}_1 \left(2\sqrt{\Delta\tilde{\lambda}\tilde{J}_\Delta} \right)}. \end{aligned}$$

Q.E.D.

APPENDIX E: MAXIMUM LIKELIHOOD JUMP TIMES AND SIZES

Over the time interval $(t, t + \Delta]$, the log-likelihood of observing $L_{t,t+\Delta}$ jumps at times $t < \tau_1^* \leq \dots \leq \tau_{L_{t,t+\Delta}}^* \leq t + \Delta$ is given by:⁴⁴ $\log f(\text{jump times in } (t, t + \Delta]) = \int_t^{t+\Delta} \left[-1_{K \times 1}^\top \lambda_s + \dot{N}_s^\top (\log \lambda_s - z_s) \right] ds$.

Conditioning on the paths of λ_s and ζ_s and taking expectations leads to:

$$\mathbb{E} \log [f(\text{jump times in } (t, t + \Delta)) | \lambda_s, \zeta_s, s \in (t, t + \Delta)] = \int_t^{t+\Delta} \left[-1_{K \times 1}^\top \lambda_s + \lambda_s^\top \log \lambda_s - 1_{K \times 1}^\top \zeta_s \right] ds.$$

We define the following maximization problem:

PROBLEM

$$(17) \quad \max_{\lambda_s, \zeta_s, \omega_s, s \in (t, t + \Delta)} \left\{ \int_t^{t+\Delta} \left[-1_{K \times 1}^\top \lambda_s + \lambda_s^\top \log \lambda_s - 1_{K \times 1}^\top \zeta_s + \omega_s^\top \zeta_s \right] ds \right\}$$

subject to:

$$\check{\alpha} (\lambda_\infty - \lambda_s) + \check{\beta} \zeta_s - \dot{\lambda}_s = 0,$$

⁴⁴For more information on how to build the likelihood for marked point processes, see section 7.3 in Daley and Vere-Jones (2003).

$$\lambda_t, \lambda_{t+\Delta},$$

$$\omega_s \geq 0, \zeta_s \geq 0, \omega_s^\top \zeta_s = 0, \quad \forall s \in (t, t + \Delta],$$

where $\omega_s, s \in (t, t + \Delta]$ are Lagrange multipliers on the positivity constraint on ζ_s .

To further simplify our maximisation problem, in equation (17), we replace $-1_{K \times 1}^\top \lambda_s + \lambda_s^\top \log \lambda_s$ with a first order Taylor approximation of this function in λ_s around λ_t , its value at the lower bound of our interval.

Due to the linearity in the objective function in ζ_s , the solution to the problem above features jumps in λ_s . To avoid the numerical difficulties deriving from this, we add a smoothing quadratic term to the terms in (17), pre-multiplied by a smoothing constant set to $-\frac{\kappa}{2}$.

This results in the following approximated maximisation problem:

PROBLEM E.1

$$\max_{\lambda_s, \zeta_s, \omega_s, s \in (t, t + \Delta]} \left\{ \int_t^{t+\Delta} \left[-1_{K \times 1}^\top \lambda_t + \lambda_s^\top \log \lambda_t - 1_{K \times 1}^\top \zeta_s + \omega_s^\top \zeta_s - \frac{\kappa}{2} \zeta_s^\top \zeta_s \right] ds \right\}$$

subject to:

$$\check{\alpha} (\lambda_\infty - \lambda_s) + \check{\beta} \zeta_s - \dot{\lambda}_s = 0.$$

$$\lambda_t, \lambda_{t+\Delta},$$

$$\omega_s \geq 0, \zeta_s \geq 0, \omega_s^\top \zeta_s = 0, \quad \forall s \in (t, t + \Delta],$$

Solution. Let $\mu_s, s \in (t, t + \Delta]$ be the co-state variable of the optimisation problem E.1. The first

order conditions are given in the following system of equations:

$$\begin{cases} \dot{\mu}_\tau = \check{\alpha}^\top \mu_\tau - \log \lambda_t \\ \dot{\lambda}_\tau = \check{\alpha} (\lambda_\infty - \lambda_\tau) + \check{\beta} \zeta_\tau \\ \omega_s + \check{\beta}^\top \mu_s - 1_{K \times 1} = \kappa \zeta_s \end{cases} \Rightarrow \begin{cases} \mu_s = e^{\check{\alpha}^\top (s-t)} v + (\check{\alpha}^\top)^{-1} \log \lambda_t \\ \lambda_s = \lambda_\infty + e^{-\check{\alpha}(s-t)} (\lambda_t - \lambda_\infty) + \int_t^s e^{-\check{\alpha}(s-\tau)} \check{\beta} \zeta_\tau d\tau \\ \zeta_s = \max \left\{ 0, \frac{1}{\kappa} [\check{\beta}^\top \mu_s - 1_{K \times 1}] \right\}, \end{cases}$$

with $\tau \in (t, s]$ and v an integration constant that must be chosen to ensure that $\lambda_{t+\Delta}$ takes the correct value. In practice, this requires solving a non-linear equation.

APPENDIX F: QUARTILE RESPONSES IN THE CDS SPREADS TO EVENTS IN ONE COUNTRY
AT A TIME

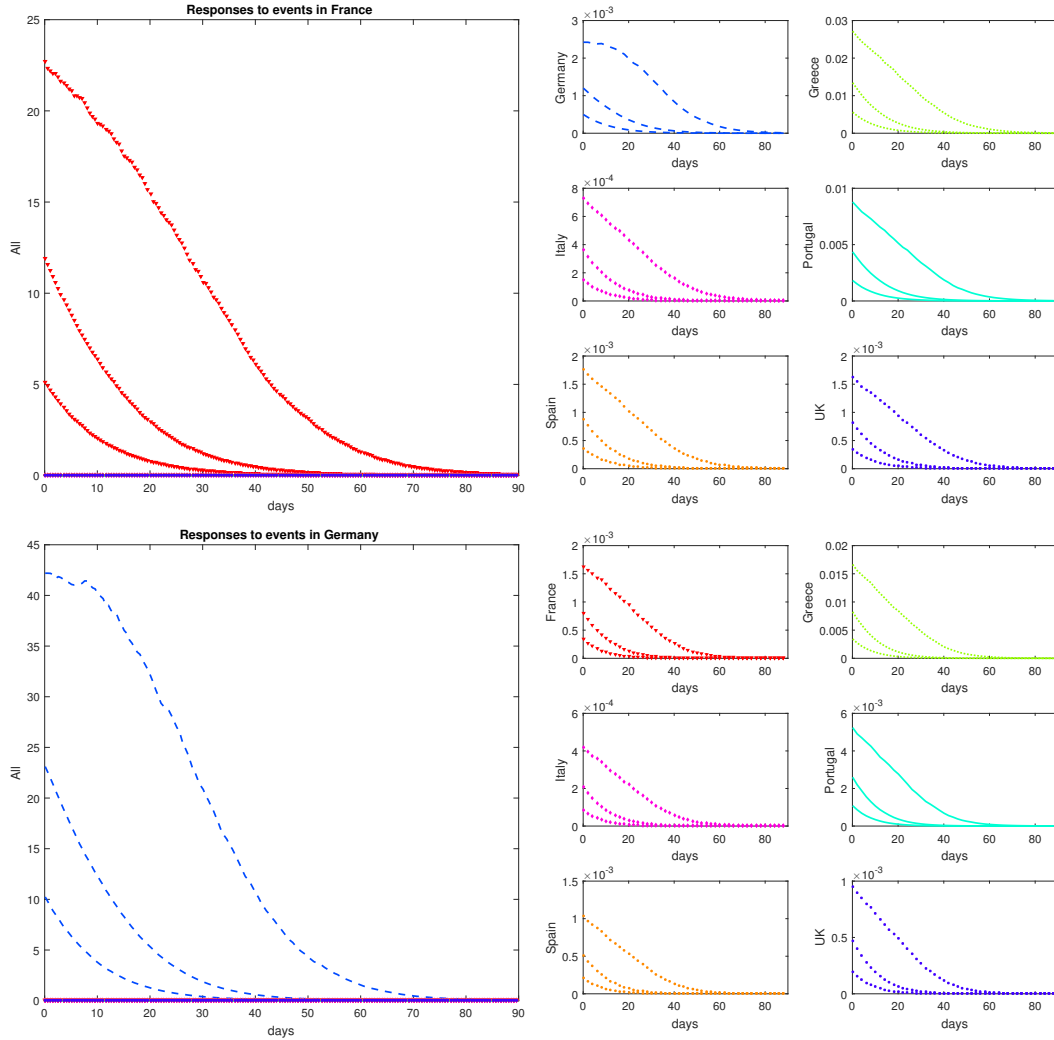


Figure 6: Percentage effect of one event at the start of the sample in France (upper panel) and Germany (lower panel) on the CDS spreads of all countries (bigger figure) and the rest of countries (smaller figures)

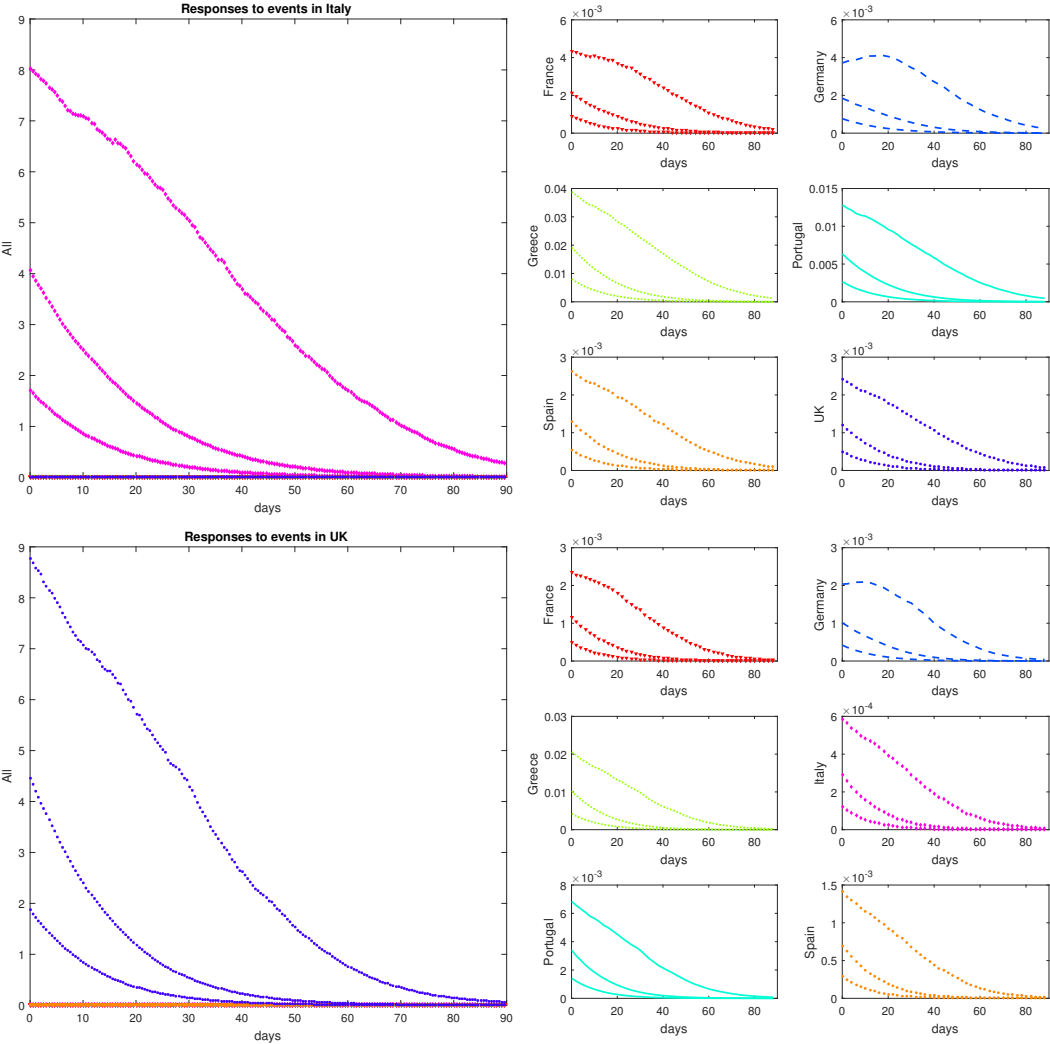


Figure 7: Percentage effect of one event at the start of the sample in Italy (upper panel) and the UK (lower panel) on the CDS spreads of all countries (bigger figure) and the rest of countries (smaller figures)