A HAWKES MODEL OF THE TRANSMISSION OF EUROPEAN SOVEREIGN DEFAULT RISK

Ana-Maria H. Dumitru [a] and Tom D. Holden [a]

The run-up to the Greek default featured an increasing cost of insuring sovereign debt in Europe. Plausibly, market participants believed a default in one country might increase the risk of a future default in another. To test for such dynamic contagion between credit related events, we develop a procedure for tractably estimating high-dimensional Hawkes models using credit default swap prices, via maximum likelihood. We escape the curse of dimensionality by modelling a market portfolio of risk across countries. We find significant spillovers in credit risk between countries, with Spain, Portugal and Greece driving events in the other countries considered.

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

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 in a country such as Greece or Portugal. By studying the transmission of default risk between European countries in the run up to the Greek crisis, we can hopefully learn lessons about present and future frailties.

In this paper, we build and estimate a model to explore the dynamic transmission of shocks to credit risk across several European Union countries during the months preceding Greece's default. 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. By modelling the international portfolio of country risk, we are able to allow for a rich structure of instantaneous and dynamic spillovers, even with seven countries. We estimate the model via conditional maximum likelihood.[1] We do not impose stationarity in order to match the explosive path of Greece's CDS prices. Thanks to the maximum likelihood approach, we

[a] School of Economics, FASS, University of Surrey, Guildford, GU2 7XH, UK

[1] Conditioning on the initial observation.

are also able to incorporate the information concerning the non-occurrence of defaults. 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.

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,[2] and so changes in the riskiness of Greek debt only had a small impact on risk in other countries. For Portugal however, default was not inevitable according to our estimates, reducing these incentives for insulation. With Portugal having high debt to GDP, and being perceived as a substantial risk by the ratings agencies,[3] it naturally then became the key driver of the upward trend in sovereign risk amongst European countries. We also provide evidence that Portugal was the most systemically important debtor in the network of European debt holdings over the period, giving an additional explanation for its role in driving systemic risk.

After Portugal, Spain was probably the next most at risk of default. Indeed, due to its large size, such a default would have had major 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. Furthermore, our network analysis reveals that Spain was the most systemically important creditor over the period, providing further explanation of its prominence in our estimates.

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 Aït-Sahalia et al. (2014) points out that credit events[4] and other jumps

[2]This 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.

[3]See Table I.

[4]In accordance to the 2003 International Swaps and Derivatives Association, Inc. (ISDA hereafter) Credit

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.

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 also 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.

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 default 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 defaults 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, 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 derived from the implied shocks to the intensities of credit related events, 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

Derivatives Definitions, the term "credit event" for sovereign debt covers the following categories: failure to make payments, restructuring, repudiation and moratorium.

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.

The literature on sovereign credit risk has had various contributions, with a lot of the earlier work focusing on emerging economies. Edwards (1984) estimate the perceived probabilities of default for different countries relying on the spread between the interest rate charged to a particular country and the LIBOR[5] and find 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 1990ś 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. (2011) 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 using Russian dollar-denominated bonds. Later contributions involving the pricing of CDS spreads include Pan and Singleton (2008); Longstaff et al. (2005); Ang and Longstaff (2013). From the existing literature using reduced form default models to explain the dynamics of

[5]London Interbank Borrowing Rate

CDS spreads, two more recent contributions are particularly relevant to the present paper.

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.

As previously mentioned, Aït-Sahalia et al. (2014) were the first to use self-exciting point processes in the context of reduced form models. While building upon their contribution for pricing CDS spreads when intensities are Hawkes (1971) processes, our work substantially differs from theirs in all aspects: modelling, estimation and data. Firstly, we consider a marked point process, as opposed to an unmarked one. In this sense, our model is closer to the uni-variate model in Eymen Errais and Goldberg (2010), who consider a loss point process that jumps at the same time as a default counting process and is further used to price index and tranche market prices. Secondly, we impose a particular parameter structure on the multivariate 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. 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. 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. Finally, our data set covers many countries and maturities.

The rest of the paper is structured as follows. Section 2 reports a few non-structural empirical facts about the data. Section 3 presents the model and derives the pricing of CDS spreads within it. Section 4 describes at length our estimation method. Section 5 discusses the results obtained. Finally, section 6 concludes the paper.

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 “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.[6]

Table I presents key facts for the countries in our sample. France, Germany and the UK all had perfect 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 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

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

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.

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.

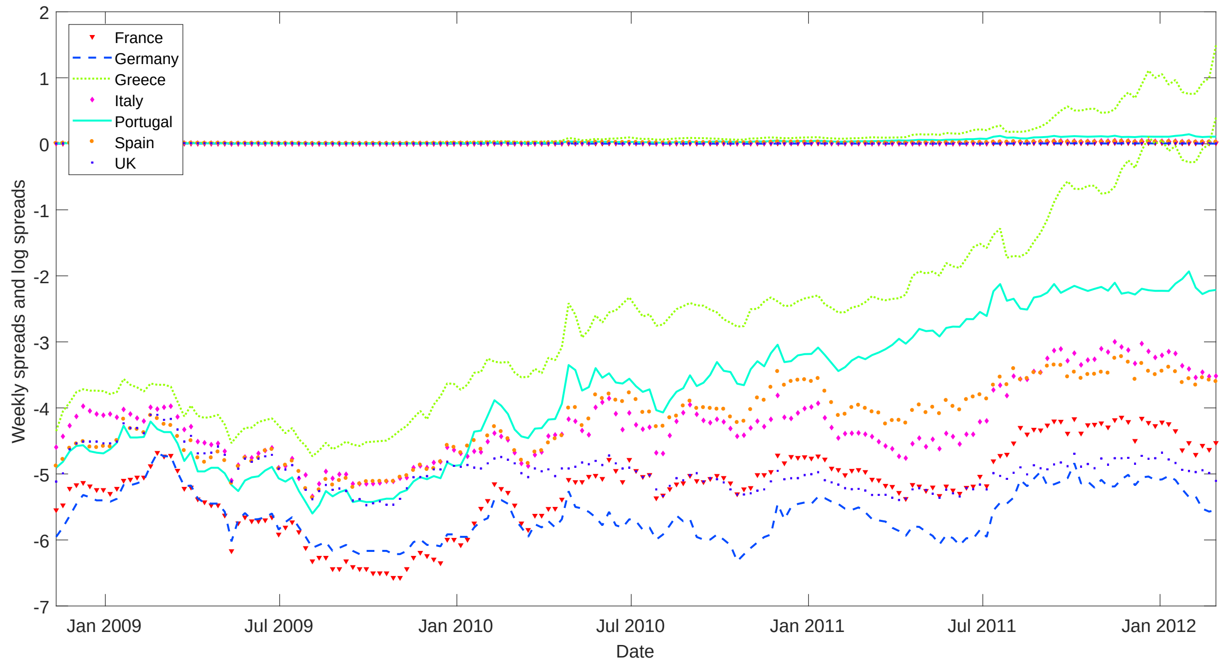
Both panels in Figure 1 also show that the ranking of countries based on credit scores is preserved for CDS spreads. Additionally, from the plots, all series appear non-stationary. To confirm this, for every country's log CDS price, we run (augmented) Dickey-Fuller and Philips-Perron tests with all numbers of lags from zero to ten, both with and without a trend included.[7] 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).

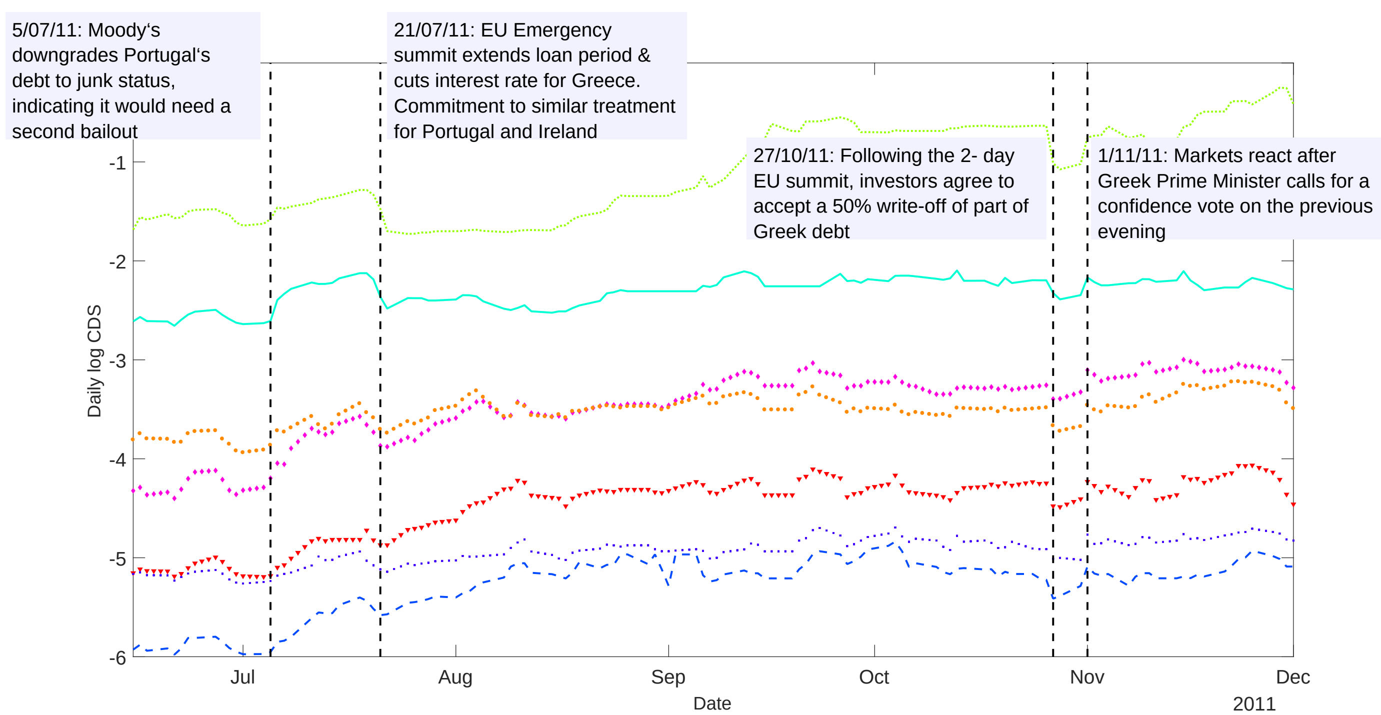
3. BUILDING A CREDIT RISK MODEL, STEP BY STEP

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.

[7]In 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.



(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.

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

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

$$\mathrm{d}\lambda_t = \breve{\alpha}\left(\lambda_\infty - \lambda_t\right)\mathrm{d}t + \breve{\beta}\,\mathrm{d}J_t, \tag{1}$$

with solution $\lambda_t = e^{-\breve{\alpha}t}\left(\lambda_0 - \lambda_\infty\right) + \lambda_\infty + \int_0^t e^{-\breve{\alpha}(t-s)}\breve{\beta}\,\mathrm{d}J_s$, where $\breve{\alpha}$ and $\breve{\beta}$ are $K \times K$ matrices. To ensure positiveness of the intensity, it is sufficient to assume that $\lambda_0 > \lambda_\infty$, $\breve{\beta} > 0$ and that all off-diagonal elements of $\breve{\alpha}$ are non-positive.[8] Under these assumptions, $\lambda_t > \lambda_\infty$ for all t. If all eigenvalues of $\breve{\beta} - \breve{\alpha}$ 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 $\breve{\beta}$ reflect the self-excitability of the model: shocks to a

[8]See Lemma A.1 in Appendix 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 $\breve{\beta}$ 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 $\breve{\alpha}$ and $\breve{\beta}$ do not need to be symmetric, the cross-country responses to other countries' intensities or shocks are not equal.

When matrices $\breve{\alpha}$ and $\breve{\beta}$ 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:

$$\mathrm{d}\lambda_{i,t} = \alpha_i \left(\lambda_{i,\infty} - \lambda_{i,t}\right) \mathrm{d}t + \beta_i \, \mathrm{d}J_{i,t}. \tag{2}$$

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: $w = [w_1, \cdots, w_K]^\top$ 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:

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

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:

$$\mathrm{d}\lambda_t = \left(\alpha - \phi W\right) \left(\lambda_\infty - \lambda_t\right) \mathrm{d}t + \beta \,\mathrm{d}J_t + \delta W \,\mathrm{d}\lambda_t,$$

where α, β, δ and ϕ are diagonal matrices with diagonals $\alpha_1, \ldots, \alpha_K$, $\beta_1, \ldots, \beta_K$, $\delta_1, \ldots, \delta_K$ and $\phi_1, \ldots, \phi_K$ respectively. Further rearrangement of terms gives us:

$$\begin{aligned} \mathrm{d}\lambda_t &= \breve{\alpha} \left(\lambda_\infty - \lambda_t\right) \mathrm{d}t + \breve{\beta} \,\mathrm{d}J_t, \\ \text{where} \quad \breve{\alpha} &:= \left(I - \delta W\right)^{-1} \left(\alpha - \phi W\right) \\ \text{and} \quad \breve{\beta} &:= \left(I - \delta W\right)^{-1} \beta. \end{aligned} \tag{4}$$

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 $\breve{\alpha}$ and $\breve{\beta}$ 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.[9] Aït-Sahalia et al. (2014) estimate only bivariate models and impose a diagonal structure on the $\breve{\alpha}$ 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 $\mathrm{d}J_{i,t} = z_{i,t}\,\mathrm{d}N_{i,t}$ 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 $z_{i,t} \sim \mathrm{Exp}\,(1)$, independent across time and countries. In vector form, we have $\mathrm{d}J_t = z_t \circ \mathrm{d}N_t$, where $z_t = [z_{1,t}, \ldots, z_{K,t}]^\top$.

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).

3.4. *Pricing CDS contracts*

We rely on the pricing formula derived by Aït-Sahalia et al. (2014):[10]

$$s_{i,\tau,t} = \frac{r_i \int_t^\tau D(t,s)\,\mathbb{E}[\gamma_i \lambda_{i,s} (1-\gamma_i)^{N_{i,s}-N_{i,t}} \mid \mathscr{F}_t]\,\mathrm{d}s}{\int_t^\tau D(t,s)\,\mathbb{E}[(1-\gamma_i)^{N_{i,s}-N_{i,t}} \mid \mathscr{F}_t]\,\mathrm{d}s}, \tag{5}$$

where $s_{i,\tau,t}$ denotes the CDS spread at time t, for country i and for maturity τ, measured in years, $1 - r_i$ is the recovery rate (i.e. the fraction returned upon default), $D(t,s)$ 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.

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

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

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 very close to 0. Consider the results of Aït-Sahalia et al. (2014) for example, who estimate a bivariate model of the type specified in equations (1) with $\mathrm{d}J_t = \mathrm{d}N_t$. 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 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.

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; Aït-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.[11] 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:

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

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:

$$\lambda_{t+\Delta} = \mathbb{E}\left[\lambda_{t+\Delta} | \mathscr{F}_t,\ N_{t+\Delta} = N_t + L_{t,t+\Delta},\ \forall l \in \{1, \dots, L_{t,t+\Delta}\}\ \mathrm{d}J_{\tau_l^*} = z_l^*\right].$$ [12]

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

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

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

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

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

where $\varepsilon_{i,\tau,t} \sim \mathrm{NIID}(0,1)$ and where $s_{i,\tau,t}$ is the CDS price at maturity τ implied by our λ_t. Allowing for this form of heteroskedasticity nests having measurement error in levels and in logs. 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 asymptotically 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 $\lambda^{\mathrm{u}}_{t,t+\Delta}$, where $\lambda^{\mathrm{u}}_{t,t+\Delta}$ is constant on $(t, t+\Delta]$. To minimise the approximation error, we define $\lambda^{\mathrm{u}}_{t,t+\Delta}$ 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 $\lambda^{\mathrm{u}}_{t,t+\Delta}$ to be defined as the expected mean value of λ_t over the interval, conditional on λ_t, i.e.:[13]

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

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

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

$$f\left(\chi_{i,t,t+\Delta} | \lambda_t\right) = \Delta\lambda^{\mathrm{u}}_{i,t,t+\Delta} \frac{\mathscr{I}_1\left(2\sqrt{\Delta\lambda^{\mathrm{u}}_{i,t,t+\Delta}\chi_{i,t,t+\Delta}}\right)}{e^{\Delta\lambda^{\mathrm{u}}_{i,t,t+\Delta} + \chi_{i,t,t+\Delta}}\sqrt{\Delta\lambda^{\mathrm{u}}_{i,t,t+\Delta}\chi_{i,t,t+\Delta}}}, \tag{7}$$

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

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

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: $\chi_{t,t+\Delta} \approx \beta^{-1}\left[\lambda_{t+\Delta} - \mathbb{E}\left[\lambda_{t+\Delta} | \mathscr{F}_t, N_{t+\Delta} = N_t\right]\right]$. In the limit as $\Delta \to 0$, 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 $O_p(\Delta)$.[14]

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

[13] The solution follows from the calculation of $\mathbb{E}\left(\lambda_s | \mathscr{F}_t\right)$ in Lemma B.1 in Appendix B.

[14] If 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.

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. 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.[15]

Let $\zeta_s := \mathbb{E}[\,\mathrm{d}J_s \mid z_s, \lambda_s] = z_s \circ \lambda_s$, 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

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

subject to:

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

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

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

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

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

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 $\kappa \to 0$. Further details on defining, approximating and solving this maximisation problem are given in appendix E. The optimal ζ_s for the approximated problem is:

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

where υ is an integration constant, chosen to ensure $\lambda_{t+\Delta}$ takes the correct value. When $\kappa \to 0$, our solution is more accurate, and $\zeta_s \to \infty$ 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.[16]

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

$$\chi_{t,t+\Delta} = \int_t^{t+\Delta} \zeta_s \, \mathrm{d}s. \tag{11}$$

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 $\Pr\left(\Xi_{i,t,t+\Delta}|\lambda_t, \lambda_{t+\Delta}\right)$, we approximate the true arrival rate λ_t over the time interval $(t, t+\Delta]$ by $\lambda^{\mathrm{c}}_{t,t+\Delta}$, where $\lambda^{\mathrm{c}}_{t,t+\Delta}$ is constant on $(t, t+\Delta]$. Much as before, to minimise the approximation error, we define $\lambda^{\mathrm{c}}_{t,t+\Delta}$ 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 $\lambda^{\mathrm{c}}_{t,t+\Delta}$ 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

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

[16]Estimations are obtained with $\kappa = 10^{-4}$.

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:

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

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

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.[17] 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.

[17]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.

During estimation we impose the sufficient conditions for positivity of λ derived in Appendix A.

r_i 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.

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

TABLE II
ESTIMATED PARAMETERS AND STANDARD ERRORS[18]

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

[18]We 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 respected to significance.

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.[19]

While the values for δ_i are all very low and statistically non-significant at 5%, the estimates of the other sensitivity parameter, ϕ_i, are all higher and 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 others.

The highest level of ϕ_i (≈ 0.12) is encountered for Greece, suggesting a higher sensitivity of Greece's credit risk 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 could have been better had other European countries not also been in difficulty at the same time.

Apart from Greece, the levels of ϕ_i remain relatively small. This means that the credit risk of each European country is mostly 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 λ_∞ are statistically significant 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.

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

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, 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, times the debt to GDP ratio in country A, divided by the GDP of country B.[20] This gives a rough approximation to the strength of the transmission of risk from country A to country B.[21] 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 (1999) which enables us to calculate the importance scores. From this, 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.[22]

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, characterized by unemployment rates of over 20%[23]. 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.[24] At the same time, Spain was the 5th largest economy in Europe[25] and was perceived by markets as too big to fail. Given all of this, the high weight we estimate for Spain is unsurprising.

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 an almost perfect juxtaposition 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 implying an expected 2.31 defaults per year. Greece is followed by Portugal with a maximum intensity of default of 0.30. Italy and Spain come next reaching maximums of 0.08 and 0.07 respectively. The

[20]The debt to GDP ratio in country A proxies the risk of 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.

[21]We use consolidated banking exposures from the Bank of International Settlements. The data is for quarter 4 of 2010, the only period between 2009 and 2011 for which data for Italy was available. In addition, we collect GDP and public debt data from Eurostat.

[22]For France, Germany, Greece, Italy, Portugal, Spain and the UK respectively, the debtor-importance scores are 0.048, 0.008, 0.026, 0.063, 0.606, 0.005, 0.244, and the creditor-importance scores are 0.130, 0.136, 0.017, 0.023, 0.016, 0.580, 0.098.

[23]Source: Eurostat.

[24]See, for instance, `http://www.bbc.co.uk/news/world-europe-18384291`.

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

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 6,316, translating to an expected 121 events per week.

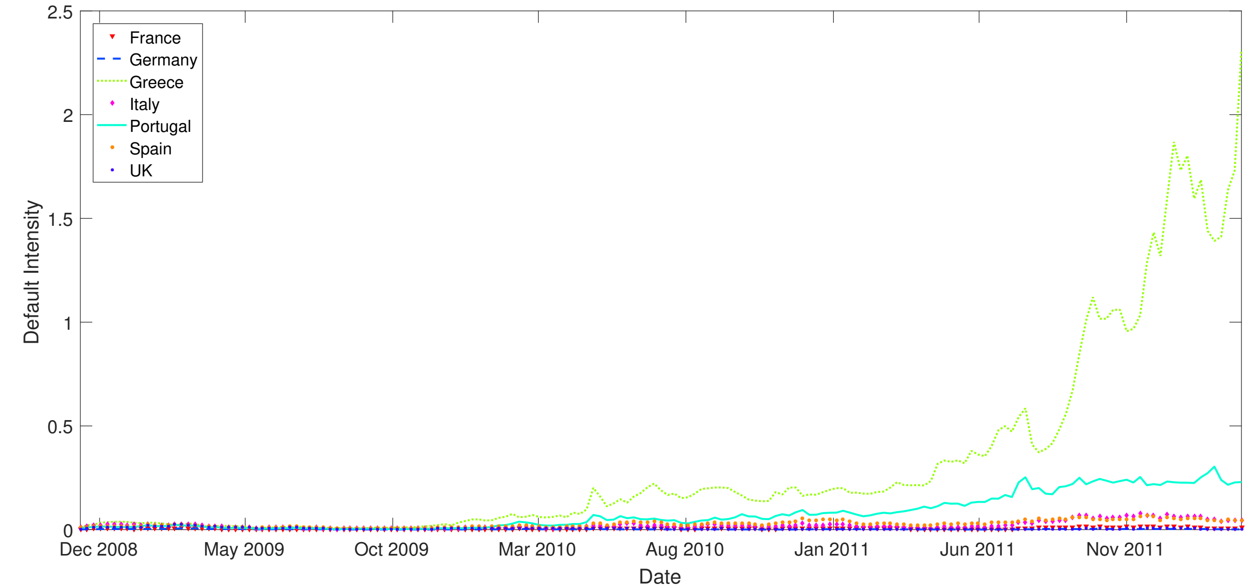
5.2. *Contagion*

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 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.

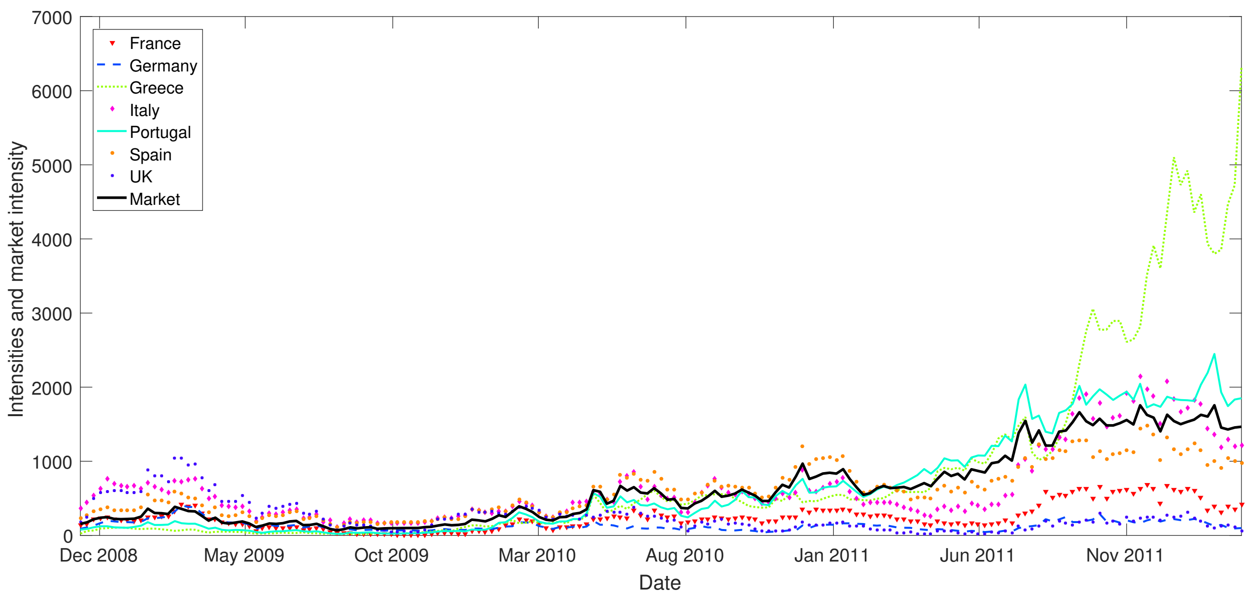
Figures 3 and 4 show the percentage quartile responses of CDS prices in all countries to events in Portugal and Spain (first figure) and Greece (second figure). Responses to events in all other countries are included in Appendix F. All figures are generated by taking the Monte Carlo quartiles of the 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.[26]

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.

[26]For 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).



(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.

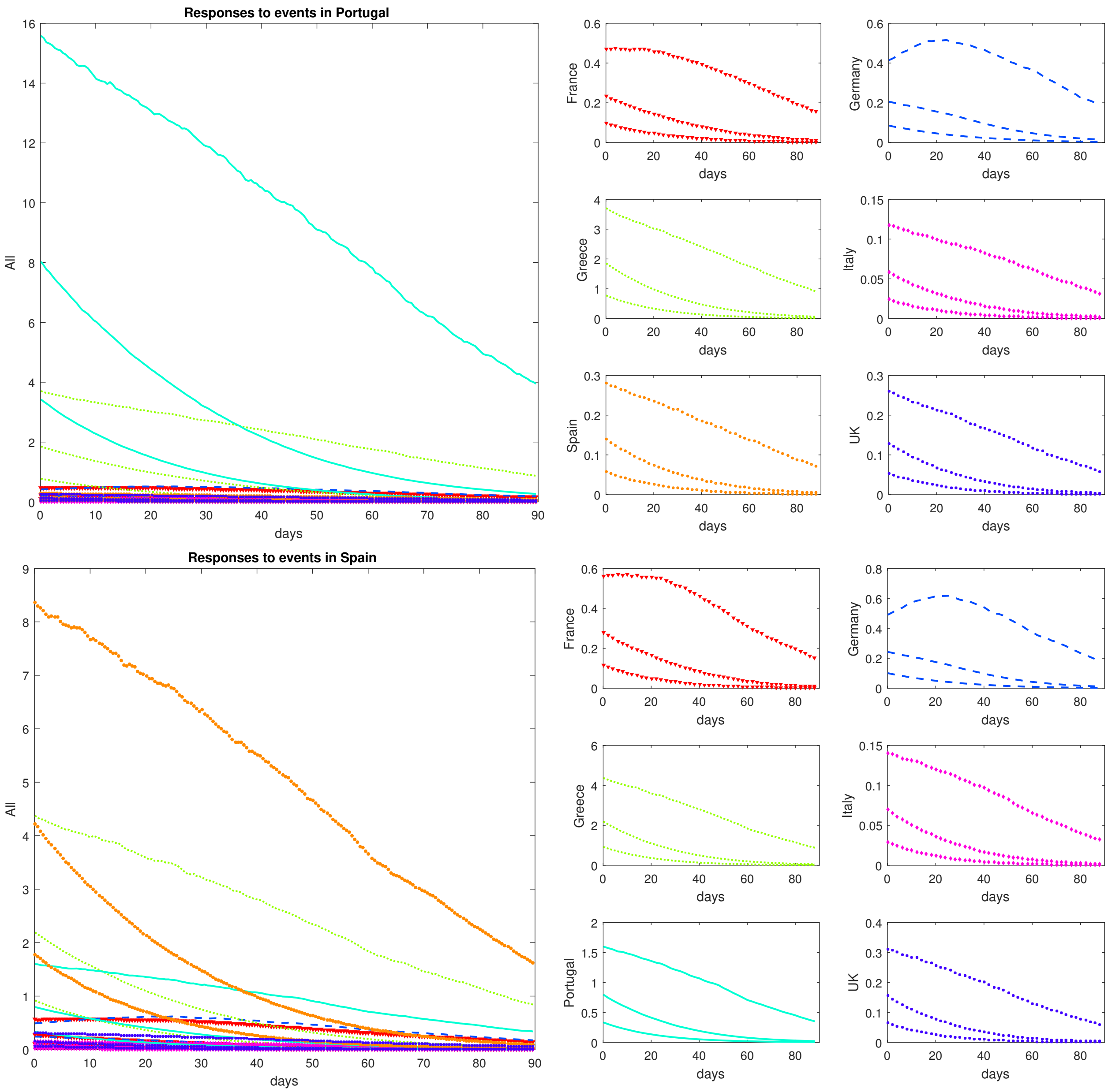


Figure 3: Percentage effect of one event at the start of the sample in Portugal (upper panel) and Spain (lower panel) on the CDS spreads of all countries (bigger figure) and the rest of countries (smaller figures)

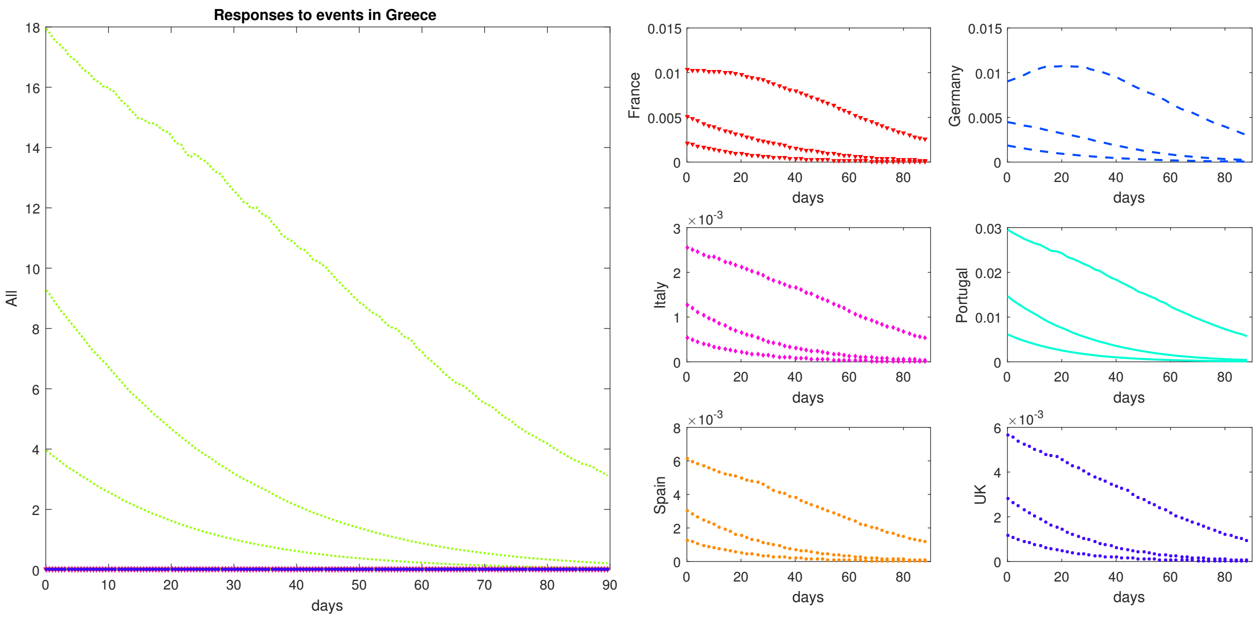


Figure 4: Percentage effect of one event at the start of the sample in Greece on the CDS spreads of all countries (bigger figure) and the rest of countries (smaller figures)

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 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.[27] Overall, Figures 3 and 4 show evidence that risk from the periphery countries is transmitted to non-peripheral countries, such as Germany, France and the UK.

6. CONCLUSION

In this paper, we model the intensities of sovereign credit related events for a group of 7 EU countries as a multi-variate selfexciting marked point process with both country-specific

[27] From the same data as was used in Footnote 21.

and market components. Each country's intensity adjusts to the market differently, with two parameters per country controlling the strength of this adjustment. This structure allows for rich feedback between risks in different countries.

Our 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.

We use the results of Duffie et al. (2000) to obtain closed form expressions for CDS prices in terms of the intensities. We estimate the model by maximum likelihood on data covering 7 maturities. In estimation, we use a novel technique to back out implied cumulated shocks to the intensities of credit-related events.

Our modelling and estimation choices, accompanied by the use of a global optimisation algorithm, enable the estimation of a rich model, with a high number of parameters and a large data set. To the best of our knowledge, this is the richest estimation problem dealt with in the credit risk literature.

Our results suggest substantial vulnerability of the system to events in the periphery. Notably, Portugal and Spain have the largest weights within the "market" factor. A much smaller, but statistically significant weight is estimated for Greece. We believe this reflects investors' views on the sources of risk in Europe during the observation window. As the Greek crisis was already being resolved during the period, investors' focus was shifted toward future dangers, of which the ones from Portugal and Spain appeared more imminent. In turn, Greece shows a higher sensitivity to the credit risk levels of all other countries, reflecting its precarious state.

REFERENCES

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.

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.

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.

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.

Eymen Errais, K. G. and L. R. Goldberg (2010). Affine point processes and portfolio credit risk. *SIAM Journal of Financial Mathematics 1*(1), 642–665.

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. (1999). 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. A., J. Pan, L. H. Pedersen, and K. J. Singleton (2011, April). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics 3*(2), 75–103.

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.

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

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

PROOF: Let $A := -\breve{\alpha}$, 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 $e^B = I + B + \frac{1}{2!}B^2 + \ldots$. 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 $\breve{\alpha}$ are non-positive. All elements of $\breve{\beta}$ 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 $D := \mathbb{N}^K \times \mathbb{R}^K_{\geq 0} \times \mathbb{R}^K_{\geq 0}$. Let $d := 3K$ be the overall dimension of X_t. Let $f\colon D \to \mathbb{R}$ be a function of X_t and let ∇f be its gradient (a row vector). The infinitesimal generator for $f(x)$, defined as $\mathrm{A}f\,(x) = \lim_{t\downarrow 0} \frac{\mathbb{E}(f(X_t)) - f(x)}{t}$, is given by:

$$\mathrm{A}f\,(x) = \begin{bmatrix} 0_{K\times 1} \\ 0_{K\times 1} \\ \breve{\alpha}\,(\lambda_\infty - \lambda_\cdot) \end{bmatrix}^\top \nabla^\top f + \sum_{i=1}^{K} \lambda_{i,\cdot} \int_{\mathbb{R}^d_{\geq 0}} \left[f\,(x + \mathrm{Z}) - f\,(x)\right] \mathrm{d}\nu^i(\mathrm{Z}), \tag{13}$$

where $x := [\lambda_\cdot^\top, J_\cdot^\top, N_\cdot^\top]^\top$, $\mathrm{Z} := [Z_1, \ldots, Z_d]^\top$ 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 $z_{i,\cdot} \sim \mathrm{Exp}(1)$ 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:

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

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

To further help understand equation (13), we note that the first term of the sum in that equation is given by:

$$\lambda_{1,\cdot}\int_{\mathbb{R}^d_{\geq 0}}\left[f\left(\begin{bmatrix}N_{1,\cdot}+1 & \cdots & N_{K,\cdot}+0 & J_{1,\cdot}+z_{1,\cdot} & \cdots & J_{K,\cdot}+0 & \lambda_{1,\cdot}+\check{\beta}_{11}z_{1,\cdot} & \cdots & \lambda_{K,\cdot}+\check{\beta}_{K1}z_1\end{bmatrix}^\top\right)- f\left(\begin{bmatrix}N_{1,\cdot} & \cdots & N_{K,\cdot} & J_{1,\cdot} & \cdots & J_{K,\cdot} & \lambda_{1,\cdot} & \cdots & \lambda_{K,\cdot}\end{bmatrix}^\top\right)\right]\mathrm{d}\nu^1(\mathrm{Z}).$$

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

$$\mathrm{d}\,\mathbb{E}\lambda_t=\check{\alpha}\left(\lambda_\infty-\mathbb{E}\lambda_t\right)\mathrm{d}t+\check{\beta}\,\mathbb{E}\lambda_t\,\mathrm{d}t,$$

with initial value λ_0 and solution:

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

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:*

$$\lim_{n\to\infty}\mathbb{E}\lambda_t=\overline{\lambda}_\infty=-\left(\check{\beta}-\check{\alpha}\right)^{-1}\check{\alpha}\lambda_\infty.$$

Proof: Apply Lemma B.1 with $t\to\infty$. *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, $\mathbb{E}\left(\lambda_t\lambda_t^\top\right)$, satisfies the following ODE:*

$$\frac{\mathbb{E}\lambda_t\lambda_t^\top}{\mathrm{d}t}=\check{\alpha}\lambda_\infty\,\mathbb{E}\lambda_t^\top+\mathbb{E}\lambda_t\lambda_\infty^\top\check{\alpha}^\top+\left(m_1\check{\beta}-\check{\alpha}\right)\mathbb{E}\left(\lambda_t\lambda_t^\top\right)+\mathbb{E}\left(\lambda_t\lambda_t^\top\right)\left(m_1\check{\beta}-\check{\alpha}\right)^\top+ m_2\left\{\check{\beta}\mathrm{diag}\left[\mathbb{E}\lambda_t\right]\check{\beta}^\top\right\}. \tag{14}$$

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

$$\mathrm{A}f(x) = \begin{bmatrix} 0_{K\times 1} \\ 0_{K\times 1} \\ \breve{\alpha}\left(\lambda_\infty - \lambda.\right) \end{bmatrix}^\top \begin{bmatrix} 0_{K\times 1} \\ 0_{K\times 1} \\ 2\lambda_{1,\cdot} \\ \vdots \\ 0 \end{bmatrix} + \lambda_{1,\cdot} \int_{\mathbb{R}^d_{\geq 0}} \left[\left(\lambda_{1,\cdot} + \breve{\beta}_{11} z_{1,\cdot}\right)^2 - \lambda_{1,\cdot}^2\right] \mathrm{d}\nu^1(Z) + \dots$$

$$+ \lambda_{K,\cdot} \int_{\mathbb{R}^d_{\geq 0}} \left[\left(\lambda_{1,\cdot} + \breve{\beta}_{1K} z_{K,\cdot}\right)^2 - \lambda_{1,\cdot}^2\right] \mathrm{d}\nu^K(Z)$$

$$= 2\lambda_{1,\cdot}\breve{\alpha}_1^\top\left(\lambda_\infty - \lambda.\right) + 2m_1\lambda_{1,\cdot}\sum_i \breve{\beta}_{1i}\lambda_{i,\cdot} + m_2\sum_i \breve{\beta}_{1i}^2\lambda_{i,\cdot}.$$

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

$$\mathbb{E}\left[f(X_t)\right] = f(X_0) + \mathbb{E}\left[\int_0^t \mathrm{A}f(X_s)\,\mathrm{d}s\right].$$

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

$$\mathbb{E}\left[\lambda_{1,t}^2\right] = \lambda_{1,0}^2 + \mathbb{E}\left[\int_0^t \mathrm{A}f(X_s)\,\mathrm{d}s\right] \Longrightarrow$$

$$\frac{\mathrm{d}\,\mathbb{E}\left[\lambda_{1,t}^2\right]}{\mathrm{d}t} = 2\breve{\alpha}_{1\cdot}^\top\lambda_\infty\,\mathbb{E}\left(\lambda_{1,t}\right) - 2\breve{\alpha}_{1\cdot}^\top\,\mathbb{E}\left(\lambda_{1,t}\lambda_t\right) + 2m_1\breve{\beta}_{1\cdot}^\top\,\mathbb{E}\left(\lambda_{1,t}\lambda_t\right) + m_2\breve{\beta}_{1\cdot}^\top\,\mathrm{diag}\left[\mathbb{E}\lambda_t\right]\breve{\beta}_{1\cdot}$$

$$= 2\breve{\alpha}_{1\cdot}^\top\lambda_\infty\,\mathbb{E}\left(\lambda_{1,t}\right) + 2\left(\breve{\beta}_{1\cdot}m_1 - \breve{\alpha}_{1\cdot}\right)^\top\,\mathbb{E}\left(\lambda_{1,t}\lambda_t\right) + m_2\breve{\beta}_{1\cdot}^\top\,\mathrm{diag}\left[\mathbb{E}\lambda_t\right]\breve{\beta}_{1\cdot}.$$

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

$$\mathrm{A}f(x) = \begin{bmatrix} 0_{K\times 1} \\ 0_{K\times 1} \\ \check{\alpha}\left(\lambda_{\infty} - \lambda_{\cdot}\right) \end{bmatrix}^{\top} \begin{bmatrix} 0_{K\times 1} \\ 0_{K\times 1} \\ \lambda_{2,\cdot} \\ \lambda_{1,\cdot} \\ \vdots \\ 0 \end{bmatrix} + \lambda_{1,\cdot} \int_{\mathbb{R}^d_{\geq 0}} \left[\left(\lambda_{1,\cdot} + \check{\beta}_{11} z_{1,\cdot}\right)\left(\lambda_{2,\cdot} + \check{\beta}_{21} z_{1,\cdot}\right) - \lambda_{1,\cdot}\lambda_{2,\cdot}\right] \mathrm{d}\nu^1(Z)$$

$$+ \ldots + \lambda_{K,\cdot} \int_{\mathbb{R}^d_{\geq 0}} \left[\left(\lambda_{1,\cdot} + \check{\beta}_{1K} z_{K,\cdot}\right)\left(\lambda_{2,\cdot} + \check{\beta}_{2K} z_{K,\cdot}\right) - \lambda_{1,\cdot}\lambda_{2,\cdot}\right] \mathrm{d}\nu^K(Z)$$
$$= \lambda_{2,\cdot}\check{\alpha}_{1\cdot}^{\top}\left(\lambda_{\infty} - \lambda_{\cdot}\right) + \lambda_{1,\cdot}\check{\alpha}_{2\cdot}^{\top}\left(\lambda_{\infty} - \lambda_{\cdot}\right) + m_1\lambda_{1,\cdot}\sum_i \check{\beta}_{2i}\lambda_{i,\cdot} + m_1\lambda_{2,\cdot}\sum_i \check{\beta}_{1i}\lambda_{i,\cdot} + m_2\sum_i \check{\beta}_{1i}\check{\beta}_{2i}\lambda_{i,\cdot}$$

In terms of expectations, we have:

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

Given the above ODEs for $\mathbb{E}\left(\lambda_{1,t}\right)$ and $\mathbb{E}\left[\lambda_{1,t}\lambda_{2,t}\right]$, 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 $\overline{\Lambda}_{\infty} = \lim_{n\to\infty}\mathbb{E}\left(\lambda_t\lambda_t^{\top}\right) - \overline{\lambda}_{\infty}\overline{\lambda}_{\infty}^{\top}$, solves the following algebraic equation:*

$$\left(m_1\check{\beta} - \check{\alpha}\right)\overline{\Lambda}_{\infty} + \overline{\Lambda}_{\infty}\left(m_1\check{\beta} - \check{\alpha}\right)^{\top} + m_2\check{\beta}\mathrm{diag}\left(\overline{\lambda}_{\infty}\right)\check{\beta}^{\top} = 0.$$

PROOF: Apply Lemma B.3 with $t \to \infty$. *Q.E.D.*

APPENDIX C: TRANSFORMS

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

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

$$\mathbb{E}\left[\left(1-\gamma_i\right)^{N_{i,T}}\middle|\mathscr{F}_t\right] = \exp\left(a^i(t) + b^i_i(t)N_{i,t} + b^i_{2K+1}(t)\lambda_{1,t} + \ldots + b^i_{3K}(t)\lambda_{K,t}\right) \quad (15)$$

where for $i \in \{1, \dots, K\}$, the coefficients $a^i(t)$ and $\boldsymbol{b}^i(t) = \left(b^i_1(t), \cdots, b^i_{3K}(t)\right)^\top$ are solutions to the ODEs given below.

a.

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

and $\dot{\boldsymbol{b}}^i_{1:2K}(t) = 0$, where for $j_1 < j_2$, $\boldsymbol{b}^i_{j_1:j_2}(t) = \left(b^i_{j_1}(t), \dots, b^i_{j_2}(t)\right)^\top$, and where $\theta : \mathbb{R}^d \to \mathbb{R}^K_{\geq 0}$ is defined by:

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

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

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

b.

$$\dot{a}^i\left(t\right) = -\left(\check{\alpha}\lambda_\infty\right) \cdot \boldsymbol{b}^i_{2K+1:3K}(t),$$

with terminal condition $\alpha^i(T) = 0$, 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). 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:

$$\begin{aligned}
\mathrm{d}X_t &= \begin{bmatrix} \mathrm{d}N_{1,t} & \cdots & \mathrm{d}N_{K,t} & \mathrm{d}J_{1,t} & \cdots & \mathrm{d}J_{K,t} & \mathrm{d}\lambda_{1,t} & \cdots & \mathrm{d}\lambda_{K,t} \end{bmatrix} \\
&= \begin{bmatrix} 0_{2K\times 1} \\ \breve{\alpha}\left(\lambda_\infty - \lambda_t\right) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 1 & \cdots & 0 & Z_{K+1} & \cdots & 0 & Z_{K+1}\breve{\beta}_{11} & \cdots & Z_{K+1}\breve{\beta}_{K1} \end{bmatrix} \mathrm{d}N_{1,t} + \ldots \\
&+ \begin{bmatrix} 0 & \cdots & 1 & 0 & \cdots & Z_{2K} & Z_{2K}\breve{\beta}_{1K} & \cdots & Z_{2K}\breve{\beta}_{KK} \end{bmatrix} \mathrm{d}N_{K,t}
\end{aligned}$$

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

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

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

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

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

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

with $\boldsymbol{b}^i(T) = \left(0, \ldots, \log\left(1-\gamma_i\right), \ldots, 0, 0, \ldots, 0, 0, \ldots, 0\right)^\top$. So:

$$\dot{\boldsymbol{b}}^i_{2K+1:3K}(t) = \breve{\alpha}^\top \boldsymbol{b}^i_{2K+1:3K}(t) + 1_{K\times 1} - \theta\left(\boldsymbol{b}^i(t)\right)$$

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

To show that this ODE is well-defined despite the pole in θ, first define $\psi : \mathbb{R}^K \to \mathbb{R}^k$ by $\psi(c) = -\log\left(1_{K\times 1} - \check{\beta}^\top c\right)$, for all $c \in \mathbb{R}^K$, where the logarithm is elementwise. Now, note that $\theta\left(\boldsymbol{b}^i(t)\right) = \left(1_{K\times 1} - \gamma_i e_i\right) \circ \exp\psi\left(\boldsymbol{b}^i_{2K+1:3K}(t)\right)$, and that $\psi^{-1}(d) = \check{\beta}^{-\top}\left(1_{K\times 1} - \exp\left(-d\right)\right)$ 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:

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

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

Moreover, we have:

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

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

Proposition 2 *For each $i \in \{1, \cdots, K\}$ and $T \geq t$,*

$$
\begin{aligned}
\mathbb{E}\left[\gamma_i\lambda_{i,T}\left(1-\gamma_i\right)^{N_{i,T}}\middle|\mathscr{F}_t\right] = \exp&\left(a^i(t) + b^i_i(t)N_{i,t} + b^i_{2K+1}(t)\lambda_{1,t} + \ldots + b^i_{3K}(t)\lambda_{K,t}\right) \\
&\left(A^i(t) + B^i_{2K+1}(t)\lambda_{1,t} + \ldots + B^i_{3K}(t)\lambda_{K,t}\right)
\end{aligned}
$$

where for $i \in \{1, \dots, K\}$, the coefficients $A^i(t)$ and $\boldsymbol{B}^i(t) = \left(B^i_1(t), \cdots, B^i_{3K}(t)\right)^\top$ are solutions to the ODEs given below.

a.

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

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

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

b.

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

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). In addition to all the notation introduced in the proof of Proposition 1, we define $\varphi : \mathbb{R}^d \to \mathbb{R}^K_{\geq 0}$ by:

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

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

$$\nabla\theta(c) = \begin{bmatrix} \operatorname{diag}\theta(c) & \operatorname{diag}\left(\theta(c) \circ \varphi(c)\right) & \operatorname{diag}\left(\theta(c) \circ \varphi(c)\right)\breve{\beta}^\top \end{bmatrix}.$$

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

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

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

$$-\dot{\boldsymbol{B}}^i_{2K+1:3K}(t) = -\breve{\alpha}^\top \boldsymbol{B}^i_{2K+1:3K}(t) + \nabla\theta\left(\boldsymbol{b}^i(t)\right)\boldsymbol{B}^i(t),$$

and $\dot{\boldsymbol{B}}^i_{1:2K}(t) = 0$. Moreover, we have: $-\dot{A}^i(t) = K_0 \cdot \boldsymbol{B}^i(t) = (\breve{\alpha}\lambda_\infty) \cdot \boldsymbol{B}^i_{2K+1:3K}(t)$, 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* $\mathrm{d}\tilde{J}_t = \tilde{z}\,\mathrm{d}\tilde{N}_t$ *where* $\tilde{N}_t$ *is a univariate Poisson process with constant intensity* $\tilde{\lambda}$, *and where* $\tilde{z} \sim \operatorname{Exp}(1)$. *Assume* $\tilde{J}_0 = 0$. *Then the p.d.f. of* $\tilde{J}_\Delta$ *is given by:*

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

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

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

Q.E.D.

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

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

PROOF:

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

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:

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

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 $t < \tau_1^* \leq \cdots \leq \tau_{L_{t,t+\Delta}}^* \leq t+\Delta$ is given by:[28] $f(\text{jump times in } (t, t+\Delta]) = \int_t^{t+\Delta} \left[-1_{K\times 1}^T \lambda_s + \dot{N}_s^{\,T} (\log\lambda_s - z_s) \right] \mathrm{d}s$. Conditioning on the paths of λ_s and ζ_s and taking expectations leads to:

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

We define the following maximization problem:

Problem

$$\max_{\lambda_s, \zeta_s, \omega_s,\ s \in (t, t+\Delta]} \left\{ \int_t^{t+\Delta} \left[-1_{K\times 1}^T \lambda_s + \lambda_s^T \log \lambda_s - 1_{K\times 1}^T \zeta_s + \omega_s^T \zeta_s \right] \mathrm{d}s \right\} \tag{16}$$

subject to:

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

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

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

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 (16), we replace $-1_{K\times 1}^T \lambda_s + \lambda_s^T \log \lambda_s$ 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 (16), pre-multiplied by a smoothing constant set to $-\frac{\kappa}{2}$.

This results in the following approximated maximisation problem:

Problem E.1

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

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

subject to:

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

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

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

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:

$$\left\{\begin{array}{l} \dot{\mu_\tau} = \breve{\alpha}^T\mu_\tau - \log\lambda_t \\ \dot{\lambda_\tau} = \breve{\alpha}\left(\lambda_\infty - \lambda_\tau\right) + \breve{\beta}\zeta_\tau \\ \omega_s + \breve{\beta}^T\mu_s - 1_{K\times 1} = \kappa\zeta_s \end{array}\right. \Rightarrow \left\{\begin{array}{l} \mu_s = e^{\breve{\alpha}^T(s-t)}\upsilon + \left(\breve{\alpha}^T\right)^{-1}\log\lambda_t \\ \lambda_s = \lambda_\infty + e^{-\breve{\alpha}(s-t)}\left(\lambda_t - \lambda_\infty\right) + \int_t^s e^{-\breve{\alpha}(s-\tau)}\breve{\beta}\zeta_\tau\,\mathrm{d}\tau \\ \zeta_s = \max\left\{0, \frac{1}{\kappa}[\breve{\beta}^T\mu_s - 1_{K\times 1}]\right\}, \end{array}\right.$$

with $\tau \in (t, s]$ and υ 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

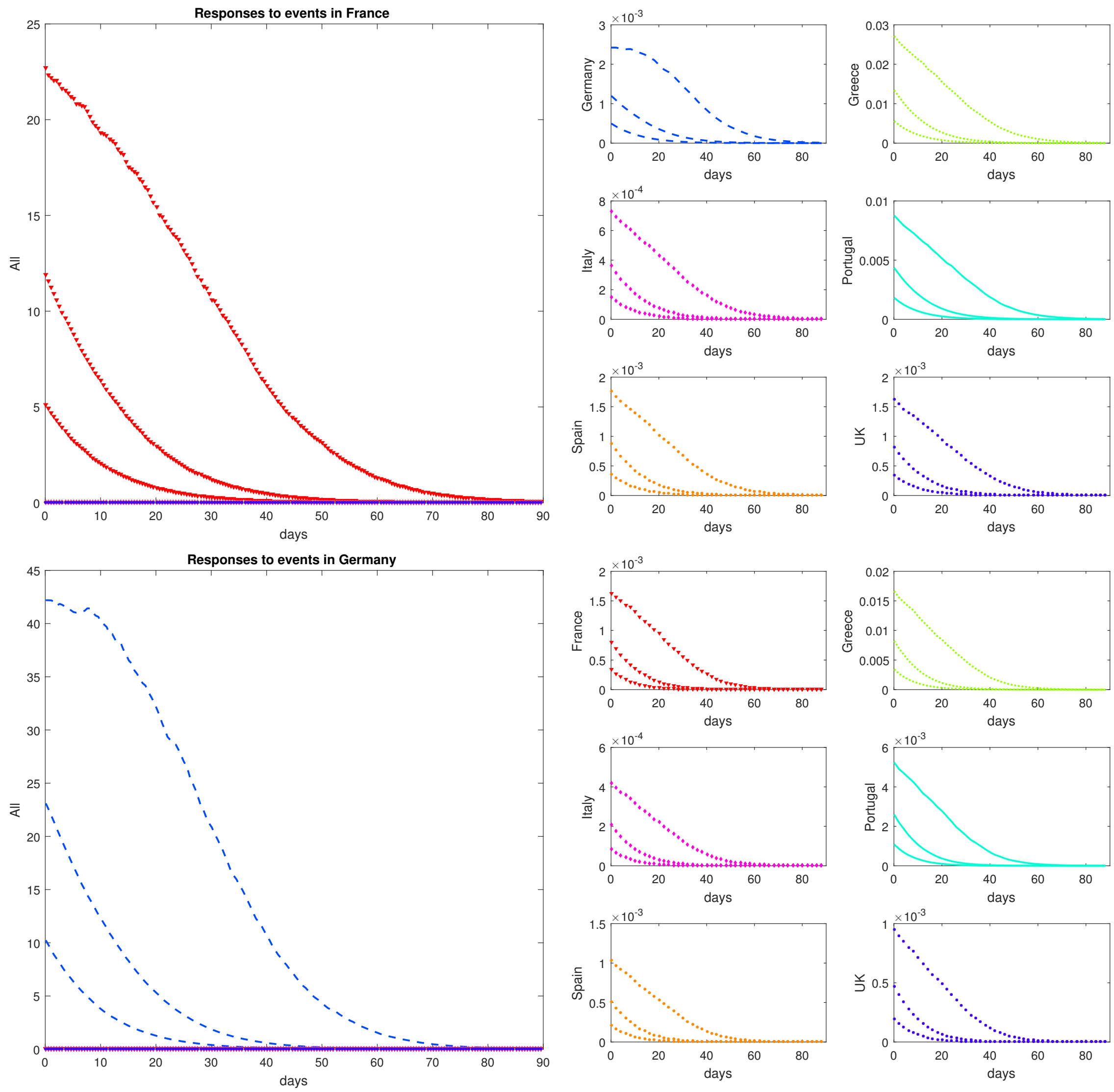


Figure 5: Percentage effect of one event at the start of the sample in France (upper panel) and Germany (lower panel) on the CDS spreads of all countries (bigger figure) and the rest of countries (smaller figures)

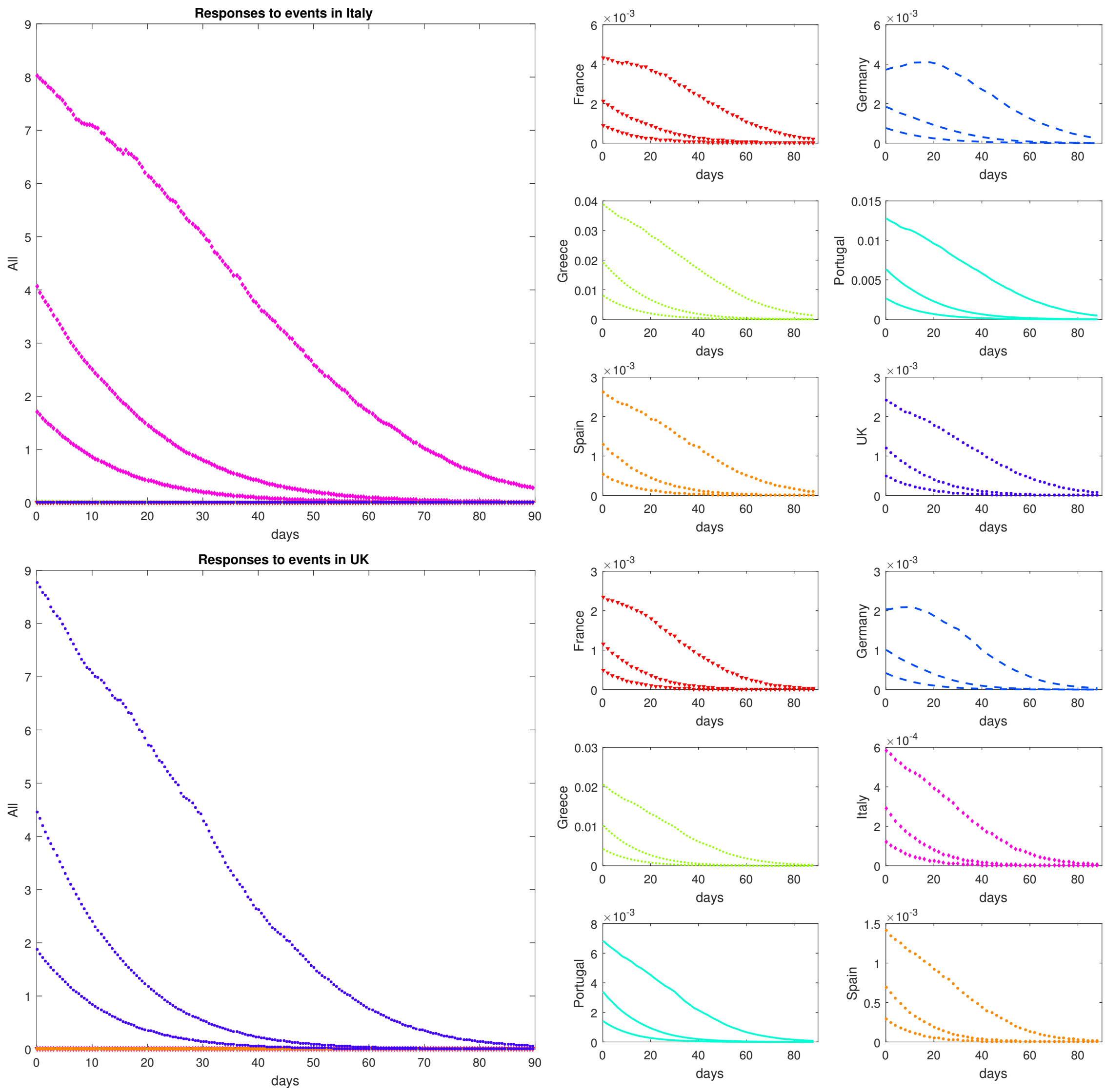


Figure 6: Percentage effect of one event at the start of the sample in Italy (upper panel) and the UK (lower panel) on the CDS spreads of all countries (bigger figure) and the rest of countries (smaller figures)