

# Jump Contagion among Stock Market Indices: Evidence from Option Markets

November 25, 2025

## Abstract

We analyze the contagious propagation of jumps among international stock market indices, using a rich panel of high-frequency stock and options data (692,892 option contracts) over the period 2006–2015. We propose a bivariate option pricing model designed to allow for time and space amplification of jumps in option markets. We develop a semi-parametric estimation procedure, which employs a continuum of moment conditions in GMM with implied states and non-parametric high-frequency spot volatility estimation. A partial-information approach is introduced to reduce the computational complexity arising in the multivariate setting. We find statistical evidence of jump contagion both within and between stock market indices. Our results reveal that jump contagion from the US to the UK is more pronounced than *vice versa*, whereas the jump contagion effects between the US and Germany stand on equal footing. We illustrate the statistical and economic importance of capturing jump contagion for risk management, option pricing, and scenario analysis. We show that accounting for jump contagion, employing scenarios based on the Global Financial Crisis, leads to an increase of capital requirements in the UK from 6.3% to 8.4% for each unit invested.

**Keywords:** Jumps; Option markets; Financial crisis; Contagion; Spatio-temporal models; C-GMM.

# 1 Introduction

Intricate linkages exist between international financial markets. Shocks to financial markets tend to propagate rapidly from one market to the next, potentially amplifying the initial shock via dynamic feedback loops. Such contagious amplification over time and in space (across markets) has important implications for risk management and scenario analysis, valuation and hedging, and portfolio choice and international diversification.

Option markets provide a unique laboratory to analyze these contagion effects. A panel of option price data, observed over time for different markets, strike prices (i.e., exercise prices) and maturities (i.e., expiry dates), embeds a wealth of information on the persistence, direction, and contagious nature of shocks.<sup>1</sup> Figure 1 provides an example of the propagation of shocks among the US (S&P 500) and UK (FTSE 100) option markets at the peak of the Global Financial Crisis of 2008. Panel (a) illustrates the interplay between the cascades of declines in the two underlying stock indices, starting with the initial drop in the US; and Panel (b) shows the reflection in option-implied volatilities for the two markets. The figures visualize in particular that the implied volatility slice for short maturity options on the UK index catches up with the US implied volatility counterpart by October 8, 2008, and even outruns it in terms of its steepness by the end of what constitutes the worst week in US stock markets since 1929.

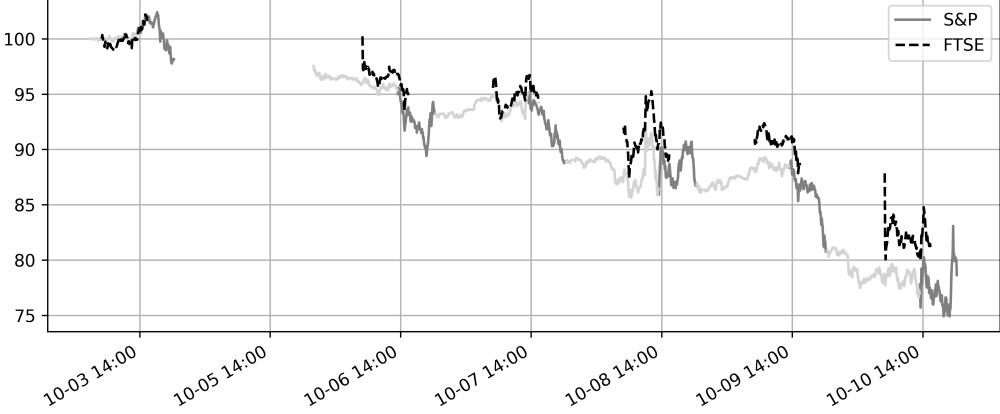
Extracting time and space amplification features from options panels, however, constitutes a challenging statistical problem. This applies in particular to jump contagion, which refers to possibly asymmetric, feedback amplification effects between large moves in asset markets, leading to jump clusters in time and space. The challenges arise from the latency of the state variables—jump intensities and stochastic volatilities—in option pricing; the multitude of dimensions—time-series, maturity, moneyness (i.e., strike-to-price ratio) and

---

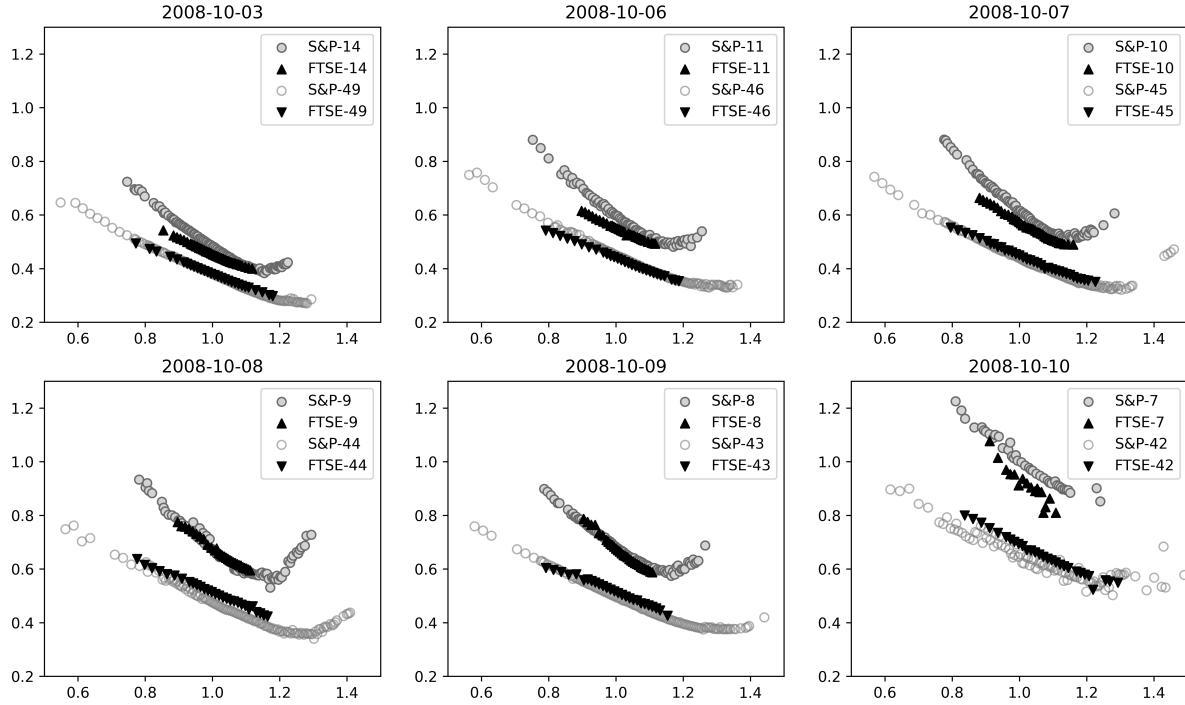
<sup>1</sup>For standard options terminology used in the sequel, we refer to the glossary in Hull (2021).

Figure 1: Contagion among the US and UK option markets, October 3–10, 2008

(a) Stock market indices



(b) Option-implied volatilities



Note: Panel (a) plots the E-Mini S&P 500 stock market futures index (S&P), and the FTSE 100 stock market index (FTSE), both scaled to 100 at the start of the sample. The observation frequency is 5 minutes, and trading times are converted to UTC (coordinated universal time), with October 4, 2008 (Saturday) omitted from the timeline. For the US futures index, the active trading periods (13:30–20:15 UTC) are highlighted in dark gray, while the remaining trading times are displayed in light gray. Panel (b) plots short-maturity Black-Scholes implied volatilities for the E-Mini S&P 500 stock market futures index options (S&P) and the FTSE 100 stock market index options (FTSE), against the moneyness level (i.e., the strike-to-price ratio). The days to maturity is indicated in the legends. The option prices are collected in the interval 14:03–14:05 UTC. More details are provided in Section 3.

cross-sectional—that play a role; and the subtlety of the features—not just occurrence of but contagion among jumps—we wish to explore. In this paper, we develop a statistical approach that exploits a rich, laboriously constructed and synchronized, panel of stock and options data to estimate a bivariate option pricing model designed to allow for, but not superimpose, time and space amplification of jumps.

We formulate our model under both the physical and the risk-neutral probability measures. The risk-neutral specification enables us to infer the unobserved state variables of our model by implying the parameter-dependent latent state variables from the panel of option prices. These are ingredients in an implied-state GMM approach with a *continuum* of moments (C-GMM) to identify the parametric components of our semi-martingale model—the drift and jump components. We treat the time-varying spot volatility components non-parametrically, by equating them to jump-robust estimates obtained from high-frequency data over short periods of time to facilitate robust identification of our rich model. This allows us to study, as is typical in classical GMM, a partially specified parametric model that only delineates a subset of statistical relationships that are of particular interest, while benefiting from the efficiency advantages of C-GMM. To reduce the computational complexity of the method, we introduce a partial-information approach, somewhat similar to the limited-information estimators considered by Singleton (2001). Monte Carlo simulations indicate that our criterion function embodies sufficient information to identify the model parameters, yielding a good finite-sample performance.

We use these statistical tools to analyze a panel of high-frequency stock and options data for the UK (FTSE 100), Germany (DAX 30) and the US (S&P 500), covering January 2006 to August 2015 and containing 692,892 option contracts. Our findings can be summarized as follows. First, we find significant evidence of both time and space amplification of jumps for all three markets. Second, our results reveal that jump contagion from the US to the UK is more pronounced than *vice versa*, in agreement with conventional wisdom

that the US plays a “leading role” in international financial markets. Third, we find that jump contagion among the US and Germany is on equal footing for both markets. This remarkable finding may be explained by the fact that the index of Germany, the largest economy of the Eurozone market (i.e., the Euro area, to which the UK does not belong), is representative for the broader Eurozone, which has played an important role in the systemic financial crises of the decade we analyze, the Eurozone Debt Crisis in particular. Indeed, in a subsample spanning 2006–2010 that excludes the Eurozone Debt Crisis of 2011–2012, the jump contagion effects among US and German financial markets are unidirectional, from the US to Germany only. Furthermore, we demonstrate the statistical and economic importance of jump contagion in risk characteristics of log-return distributions, prices of multi-index options, and implied volatility dynamics for the S&P 500 and FTSE 100. We show e.g., that, when translated into capital requirements, accounting for jump contagion, using scenarios based on the Global Financial Crisis, amounts to an increase in required capital from 6.3% to 8.4% for each unit invested in the FTSE 100 index.

Our work is related to the literature studying the international transmission of equity shocks in the form of jumps, which is relatively small compared to the vast literature on international asset return and volatility spillovers; see Aït-Sahalia, Cacho-Díaz, and Laeven (2015), and also Errais, Giesecke, and Goldberg (2010) and Aït-Sahalia, Laeven, and Pelizzon (2014) who analyze jump contagion in credit defaults. High-frequency tests for common or mutually exciting jumps are developed in Jacod and Todorov (2009) and Dungey, Erdemlioglu, Matei, and Yang (2018); see also Aït-Sahalia and Xiu (2016). An even smaller literature—most closely related to our work—analyzes jump propagation through the lens of option markets. Andersen, Fusari, and Todorov (2020) consider the pricing of index options in separate markets, and find a large coherence across markets with respect to their left tail risk. Bakshi, Carr, and Wu (2008) investigate a genuinely multi-country stochastic discount factor using currency options, allowing for global and country-specific

diffusion and jump risk factors. Kokholm (2016) considers a multivariate option pricing model with a self- and/or cross-exciting jump component, under a risk-neutral specification. He applies it to sectoral indices in one market using a calibration technique. To our best knowledge, this paper is the first to analyze jump contagion among international stock market indices using the laboratory of option panel data.

Our statistical approach is inspired by the option-implied-state GMM approach of Pan (2002), and the C-GMM approach of Carrasco and Florens (2000, 2002) and Carrasco, Chernov, Florens, and Ghysels (2007). Different from this existing literature, our approach is semi-parametric in nature, mitigates the computational complexity of multivariate C-GMM using a partial-information approach, and accounts for estimation uncertainty from state-implied moments when computing standard errors.

This paper is organized as follows. Section 2 describes the model. Section 3 describes the data. Section 4 develops the estimation procedure. In Section 5 we present our empirical analysis. Conclusions are in Section 6. Further details on the model specification, data selection and processing, the estimation procedure, and data analysis are provided in four online appendices provided as supplementary material. Computer code to implement the procedures developed in this paper is available from a GitHub repository.<sup>2</sup>

## 2 Model Specification

This section presents our multivariate continuous-time option pricing model with contagious time and space amplification. It embeds the mutually exciting jump process proposed in Ait-Sahalia et al. (2015) to characterize the stock index dynamics in  $m$  economies. Unlike contagion models of multivariate stock index returns, modeling contagion among option prices requires extension of the pricing kernel. Therefore, we propose a class of

---

<sup>2</sup><https://anonymous.4open.science/r/Jump-Contagion-CEE0/>.

country-level risk-neutral probability measures that jointly accommodate arbitrage-free international stock price dynamics with mutually exciting jumps. Next, we describe the semi-nonparametric approximation adopted in our model.

## 2.1 Index Return Dynamics and Risk-Neutral Measures

We fix a filtered probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  and consider a model of index return dynamics for  $m$  economies equipped with mutually exciting jump processes. We assume that each of the markets is characterized by a stock market index denominated in the local currency with the following dynamics:

$$\frac{dS_{i,t}}{S_{i,t}} = (r_{i,t} - q_{i,t} + \eta_i \xi_{i,t}^2 + (\mathbb{E}[J_i] - \mathbb{E}^{\mathbb{Q}_i}[J_i])\lambda_{i,t}) dt + \xi_{i,t} dW_{i,t} + J_{i,t} dN_{i,t} - \mathbb{E}[J_i]\lambda_{i,t} dt, \quad (1)$$

for  $i = 1, \dots, m$ , where  $r_{i,t}$  and  $q_{i,t}$  are deterministic risk-free rates and dividend yields;  $W_{i,t}$  are standard Brownian motions, correlated with (possibly time-varying) pairwise instantaneous correlation coefficients  $\varrho_{ij,t}$ ;  $\xi_{i,t}$  are general, adapted volatility processes; and  $J_{i,t} dN_{i,t}$  are compound Hawkes jump processes with serially and cross-sectionally independent random variables  $J_{i,t}$  governing the jump sizes, having generic law  $F_{J_i}$  and mean  $\mathbb{E}[J_i]$  (under  $\mathbb{P}$ ). By  $\mathbb{E}^{\mathbb{Q}_i}[J_i]$  we denote the expected jump size in market  $i$  under the equivalent risk-neutral probability measure  $\mathbb{Q}_i$  specific to market  $i$ , as defined below. Throughout, stochastic processes, expectation operators, and parameters without superscript are understood to be defined with respect to the physical probability measure  $\mathbb{P}$ .

The Hawkes (1971) jump process, also known as the mutually exciting jump process, is a main ingredient of our model, allowing us to capture both jump contagion across markets and clustering of jumps in time within each market.<sup>3</sup> More specifically, we define

---

<sup>3</sup>Originally proposed to model epidemics, Hawkes processes have also been used to model seismic excitation since Ogata (1988).

the multivariate Hawkes jump process through  $m$  counting processes  $N_{i,t}$ , one for each of the  $m$  markets, such that each counting process is characterized by its conditional jump intensity process  $\lambda_{i,t}$ , defined by

$$\lambda_{i,t} = \lim_{s \downarrow 0} \frac{\mathbb{E}[N_{i,t+s} - N_{i,t} | \mathcal{F}_{t-}]}{s}. \quad (2)$$

Unlike the Poisson process, the jump intensity of the Hawkes process is stochastic with dynamics (under exponential decay) given by

$$d\lambda_{i,t} = \kappa_i(\bar{\lambda}_i - \lambda_{i,t})dt + \sum_{j=1}^m \delta_{ij} dN_{j,t}, \quad i = 1, \dots, m. \quad (3)$$

In this specification, a jump event in equity index  $j$  causes the intensity  $\lambda_{i,t}$  to increase by  $\delta_{ij} \geq 0$ , followed by an exponential decay towards  $\bar{\lambda}_i > 0$  at a rate  $\kappa_i > 0$ . The parameters  $\delta_{ij}$  dictate the *self-excitation* (for  $i = j$ ) and *cross-excitation* (for  $i \neq j$ ) effects, generating two key features of the model: first, a jump event increases the probability of subsequent jump events in the same index, leading to jump clustering in time; second, a jump event in one market increases the probability of jumps in other markets, entailing jump propagation in space. Note that, following a jump event in one market, the jump intensities in other markets respond instantaneously, thereby directly making jumps much more likely. Note also that these time and space amplification features are probabilistic and not superimposed, i.e., not certain to occur. The paired vectors  $(N, \lambda)$  jointly constitute a Markov process.

Our model captures jump contagion, i.e., the propagation over time and across markets of large moves in stock market indices. Sample paths from the model may exhibit nearly concurrent jumps in the different markets. By design, the model does not allow for large common exogenous shocks (it does allow for Brownian correlation), in line with the existing

literature on “contagion” in economics and finance as “propagation” rather than common shocks. Information, also common exogenous information, needs time to spread and get reflected in the different markets. Our model captures these cascading effects. To distinguish in formal statistical tests between contagious jumps and large common exogenous shocks would require a different modeling framework and ultra-high-frequency analysis.

In addition to the risk-free interest rate  $r_{i,t}$  and dividend yield  $q_{i,t}$  in economy  $i$ , the drift term in (1) contains two risk-premium components. The diffusive risk premium  $\eta_i \xi_{i,t}^2$  is akin to the risk-return trade-off occurring in the capital asset pricing model (CAPM):  $\eta_i$  represents the additional expected return per unit of diffusive (“Brownian”) variance  $\xi_{i,t}^2$ . The jump risk premium  $(\mathbb{E}[J_i] - \mathbb{E}^{\mathbb{Q}_i}[J_i])\lambda_{i,t}$  represents the additional expected return under the physical measure (relative to the risk-neutral measure), needed to compensate for bearing jump risk. The last term in (1) is the compensator for the jump component; the compensated jump component is a local martingale. Our model specification allows for, possibly time-varying, correlations between the Brownian motions. However, in the presence of mutually exciting jumps, the contribution of the Brownian correlation to the realized correlation is swamped in crisis episodes, even if the Brownian correlation increases during such time periods, and therefore it plays only a secondary role.

As is common in the literature (e.g., Pan (2002), Broadie, Chernov, and Johannes (2007)), we assume the relative jump sizes  $J_{i,t}$ ,  $i = 1, \dots, m$  to be independent log-normal random variables. Specifically, conditional upon a jump event in market  $i$ , the equity price jumps from  $S_{i,t-}$  to  $S_{i,t} = S_{i,t-} \exp(Z_{i,t})$ , with  $Z_{i,t} \sim \mathcal{N}(\mu_i, \sigma_i^2)$ . Under this parametrization, the relative jump size in index  $i$  is  $J_{i,t} = \exp(Z_{i,t}) - 1$ , with mean  $\mathbb{E}[J_{i,t}] = \exp(\mu_i + \frac{1}{2}\sigma_i^2) - 1$ . When the mean parameters (i.e.,  $\mu_i$ ) are estimated at negative values, as is the case in our data analysis, this entails that negative jumps in our model occur more frequently, and are more contagious, than positive jumps. We also assume that the vector of stacked jump sizes  $Z_t$ , vector of Brownian motions  $W_t$ , and vector of jump processes

$N_t$  are mutually independent. Importantly, this model admits a generalized affine jump-diffusion representation as defined in Appendix B of Duffie, Pan, and Singleton (2000).

The sources of uncertainty stemming from the random jumps in our model render each market  $i$ , consisting of the equity index, a finite number of options on that index and a money market account, incomplete. Therefore, the stochastic discount factor for each of the markets is not unique. To formulate our risk-neutral pricing model, we focus on candidate pricing kernels that keep the joint dynamics of the log-equity index and the jump intensity process for each market  $i$ , under the equivalent risk-neutral probability measure  $\mathbb{Q}_i$ , within the generalized affine jump-diffusion class. We provide further details on the measure change in Appendix A.1, formally establishing in particular that the pricing kernels thus specified rule out arbitrage opportunities within each market as well as internationally.

The resulting model under  $\mathbb{Q}_i$  may be represented as

$$\frac{dS_{i,t}}{S_{i,t}} = (r_{i,t} - q_{i,t}) dt + \xi_{i,t} dW_{i,t}^{\mathbb{Q}_i} + J_{i,t} dN_{i,t} - \mathbb{E}^{\mathbb{Q}_i}[J_i] \lambda_{i,t} dt, \quad (4)$$

for  $i = 1, \dots, m$ , where  $dW_{i,t}^{\mathbb{Q}_i} = dW_{i,t} + \eta_i \xi_{i,t} dt$ , with  $W_{i,t}^{\mathbb{Q}_i}$  a standard Brownian motion under  $\mathbb{Q}_i$ . The pairwise instantaneous correlation coefficients  $\varrho_{ij,t}$  between the Brownian motions are preserved under each of the risk-neutral measures. The distribution of the jump size random variables  $J_{i,t} = e^{Z_{i,t}} - 1$  is of the same translated log-normal type under  $\mathbb{Q}_i$  as under  $\mathbb{P}$ , but with possibly different parameters. We assume that only the mean parameters are different under the physical and risk-neutral specifications, i.e.,  $\sigma_i^{\mathbb{Q}_i} \equiv \sigma_i$ , as in the univariate model of Pan (2002). The assumption of equal jump size variances under both measures is needed for the identification of the jump size parameters. As a consequence, the jump risk premium  $(\mathbb{E}[J_i] - \mathbb{E}^{\mathbb{Q}_i}[J_i])\lambda_{i,t}$  is generated by the difference between  $\mu_i$  and  $\mu_i^{\mathbb{Q}_i}$ .<sup>4</sup>

---

<sup>4</sup>Our specification of the pricing kernel sets the jump-timing risk premium to zero, so that the dynamics of the jump intensity processes in (3), and in particular the parameters  $\kappa_i$ ,  $\bar{\lambda}_i$  and  $\delta_{ij}$ , are the same under

## 2.2 Semi-Nonparametric Approximate Index Return Dynamics

In the formulation of the model, the diffusive volatility processes  $\xi_{i,t}$  have not yet been specified. In principle, they could be modelled using a parametric stochastic volatility specification. Alternatively, we adopt a semi-nonparametric approximation for the index return dynamics using nonparametric estimates of the spot volatilities  $\xi_{i,t}$ . Such an approximation leads to robust pricing of close-to-maturity options, allowing inference to be focused on the latent jump intensity dynamics and jump sizes. Moreover, a fully parametric version of the model, including a stochastic volatility specification, is prone to model misspecification and identification problems, especially in a multivariate setting; the semi-nonparametric approach considerably reduces these complications.

As in the univariate setting of Andersen, Fusari, and Todorov (2017), we use an approximate representation of the stock index process with constant spot volatility and a constant dividend yield and interest rate. Under the equivalent martingale measures  $\mathbb{Q}_i$ , we define the approximate processes  $\tilde{S}_{i,s}^{\mathbb{Q}_i}, i = 1, \dots, m$  for  $s \in [t, T]$  (the time period between pricing and expiration of the option) as follows:

$$\begin{cases} \frac{d\tilde{S}_{i,s}^{\mathbb{Q}_i}}{\tilde{S}_{i,s}^{\mathbb{Q}_i}} = (r_{i,t} - q_{i,t})ds + v_{i,s}dW_{i,s}^{\mathbb{Q}_i} + J_{i,s}dN_{i,s} - \mathbb{E}^{\mathbb{Q}_i}[J_i]\lambda_{i,s}ds, & s \in [t, T], \\ v_{i,s} = \xi_{i,t}\mathbb{1}_{\{t \leq s \leq T\}}, \quad \tilde{S}_{i,t}^{\mathbb{Q}_i} = S_{i,t}. \end{cases} \quad (5)$$

In other words, the spot volatility  $v_{i,s}$  is taken to be constant and equal to  $\xi_{i,t}$  (the true spot volatility at time  $t$ ) over the interval  $s \in [t, T]$ , and the approximate process  $\tilde{S}_{i,s}^{\mathbb{Q}_i}$  is initialized at the true index price  $S_{i,t}$  at time  $t$ . We also refer to the insightful work of Medvedev and Scaillet (2007, 2010), who consider a small time-to-maturity asymptotic approximation of the implied volatility function for stochastic volatility jump-diffusions.

This approximation is reasonable for pricing short-dated options because under the  $\mathbb{P}$  and  $\mathbb{Q}_i$ ; see Appendix A.1 for details.

risk-neutral measures, the stochastic volatility process usually exhibits slow mean reversion. For example, Pan (2002) finds the mean-reversion parameter in the volatility process to be 0.013 under  $\mathbb{Q}$ , expressed in daily terms, corresponding to a one-day autocorrelation coefficient in volatility equal to 0.987. Moreover, as close-to-maturity European-style option prices are more sensitive to the specification of the jump intensity dynamics and of the jump size distribution, pricing these options using the approximated process  $\tilde{S}_{i,t}^{\mathbb{Q}_i}$  instead of  $S_{i,t}$  leads to negligible approximation errors. We confirm this in simulations in Section 4.3 and Appendix C.4. Unlike Andersen et al. (2017) we do not “freeze” the jump intensity to its value at time  $t$ , because in our setting it can vary considerably, even over the short period, due to the self-excitation and contagion effects discussed in Section 2.1.

Finally, as the change of measure does not affect the diffusion term of the price dynamics, we can also adopt the approximate dynamics for the processes under  $\mathbb{P}$ :

$$\begin{cases} \frac{d\tilde{S}_{i,s}}{\tilde{S}_{i,s}} = (r_{i,t} - q_{i,t} + \eta_i \xi_{i,t}^2)ds + v_{i,s} dW_{i,s} + J_{i,s} dN_{i,s} - \mathbb{E}^{\mathbb{Q}_i}[J_i] \lambda_{i,s} ds, & s \in [t, T], \\ v_{i,s} = \xi_{i,t} \mathbb{1}_{\{t \leq s \leq T\}}, \quad \tilde{S}_{i,t} = S_{i,t}, \end{cases} \quad (6)$$

for  $i = 1, \dots, m$ , but with  $T = t + 1$ ; i.e., the spot volatility is assumed constant over the period of a single day. The specifications (5)–(6) will serve as a basis for the estimation procedure developed in Section 4.

### 3 The Data

Estimation is based on a rich panel of daily observations, which have, in turn, been constructed from tick-by-tick spot, futures and option price data for the FTSE 100, DAX 30 and S&P 500 stock market indices, spanning the period January 1, 2006, to August 13, 2015. That is, we exploit a very large sample of intra-day tick-by-tick observations to ob-

tain daily synchronized panels of options data for the three markets in the different time zones, as well as jump-robust spot volatility estimates constructed from intra-day returns in a time interval preceding the observation time of the options. The synchronicity of cross-market observations in our estimation sample is essential for capturing jump contagion among stock market indices. We then use the panel of daily data to estimate bivariate as well as univariate versions of the model. The sampling of tick-by-tick data underlying the option panel observations is a non-trivial task, which warrants further description. This section summarizes the data collection process; additional details are in Appendix B.

The data-set was obtained from the Thomson Reuters Tick History database, containing time-stamped tick-by-tick data from electronic exchanges for several major stock market indices and corresponding exchange-listed derivative contract prices. Data samples contain bid-ask quotes and transaction prices with time-stamps in the exchange’s local time zone that denote the time at which the price data were received by Thomson Reuters from the exchange’s servers. As the use of official exchange-determined “close” prices is not possible because the options are traded in different time zones, we use the synchronization procedure as outlined next.

We create daily option panels using tick-by-tick data subsets selected from a particular time interval during market trading hours, which we refer to as *reference interval*. We choose reference intervals for market pairs such that for each corresponding pair the trade recordings are as “synchronized” as possible to the right point of the reference interval, which we refer to as *reference point*. Throughout the sample we fix the reference interval for FTSE 100 options to 15:03–15:05 and for DAX 30 options to 16:03–16:05 (local exchange times). The reference interval for the S&P 500 options is obtained by translating the UK (and Germany) reference interval to US local exchange times using IANA Time Zone Database conventions, meaning option data for the US is usually sampled between 9:03–9:05 CST, with periodical exceptions dictated by daylight saving time adjustments used in

the US and in Europe. We impose several rules and filters in the data selection routine; see Appendix B.1 for specific details. Table 1 provides the descriptive statistics for the filtered option sample for each of the three markets. In Table B.1 in the appendix, we analyze the sensitivity to the choice of the reference intervals. Specifically, we analyze descriptive statistics for the filtered option sample using two-minutes and increased five-minutes intervals. Whereas some differences between the two reference intervals are visible, they are rather marginal. Table B.2, Panel A, shows that the bid-ask spreads are the smallest for S&P, while the spreads for DAX and FTSE are of a similar magnitude. Relative bid-ask spreads are also the lowest for S&P, especially deep out-of-the-money (Table B.2, Panel B). Our estimation procedure relies on transaction data if available and otherwise on mid bid-ask quotes. The percentage of transaction data is typically less than 1% for all three markets and for different ranges of maturity and moneyness levels, as new quotes arrive much more frequently than actual trades occur. The minimum tick sizes are 0.05 (S&P), 0.1 (DAX) and 0.5 (FTSE). The data might be noisier in less liquid markets, although this is not visible in the sample standard deviations of implied volatility in Table 1, Panel C.

In addition to the filtered option data, we use short-term interbank lending interest rates for each relevant currency, which we interpolate to match option time-to-maturity. Specifically, we use LIBOR-US, LIBOR-GBP and EURIBOR short-term rates for options on S&P 500 futures, and the FTSE 100 and DAX 30 indices, respectively. We follow Aït-Sahalia and Lo (1998) by backing out forward prices from put-call parity pairs and estimating our model using log-forward returns, circumventing the need to specify and calibrate dividend yield dynamics. The details on forward price calculations are provided in Appendix B.2. We further interpolate the sample Black-Scholes implied volatilities over a fixed set of moneyness and option maturities to construct a (homogeneous) panel of input data for the estimation procedure. Interpolating implied volatilities is a common procedure; see, e.g., Broadie et al. (2007) and Bardgett, Gourier, and Leippold (2019).

Table 1: Descriptive statistics for the option implied volatility data

FTSE 100		DAX 30		S&P 500	
$5 < \tau \leq 40$	$40 < \tau \leq 75$	$5 < \tau \leq 40$	$40 < \tau \leq 75$	$5 < \tau \leq 40$	$40 < \tau \leq 75$
Panel A: Aggregate number of option contracts					
$0.75 < k \leq 0.85$	1,375	2,968	9,505	16,000	23,413
$0.85 < k \leq 0.92$	6,646	8,914	18,139	21,831	36,986
$0.92 < k \leq 0.98$	18,743	16,670	20,637	20,843	40,197
$0.98 < k \leq 1.03$	23,127	18,285	17,300	17,332	34,234
$1.03 < k \leq 1.10$	10,911	13,052	17,838	21,861	28,560
$1.10 < k \leq 1.20$	1,258	2,721	4,764	9,875	6,782
Total	62,060	62,610	88,183	107,742	170,172
Panel B: Sample mean of implied volatility (%)					
$0.75 < k \leq 0.85$	38.7	35.3	36.6	33.1	39.3
$0.85 < k \leq 0.92$	31.2	27.2	30.4	27.2	29.0
$0.92 < k \leq 0.98$	22.0	20.8	24.1	23.1	21.8
$0.98 < k \leq 1.03$	17.2	16.9	20.0	20.1	16.7
$1.03 < k \leq 1.10$	18.8	17.1	19.2	18.2	16.7
$1.10 < k \leq 1.20$	31.0	24.6	26.0	21.1	26.7
Total	21.2	20.6	25.1	23.8	24.1
Panel C: Sample standard deviation of implied volatility (%)					
$0.75 < k \leq 0.85$	9.8	10.6	7.0	7.4	10.6
$0.85 < k \leq 0.92$	10.9	8.9	8.4	7.3	9.0
$0.92 < k \leq 0.98$	8.7	7.6	7.9	7.1	8.5
$0.98 < k \leq 1.03$	7.9	6.9	7.9	7.0	8.5
$1.03 < k \leq 1.10$	9.2	7.6	7.8	6.7	8.8
$1.10 < k \leq 1.20$	11.6	9.9	9.8	7.9	11.4
Total	10.3	9.3	9.9	8.7	11.9

This table provides descriptive statistics for filtered option data on FTSE 100, DAX 30 and S&P 500 futures. The sample contains daily option data from January 1, 2006, to August 13, 2015. The filters employed in the data selection procedure are detailed in Appendix B.1. Observations are bucketed into two categories for time-to-maturity,  $\tau$ , and into six categories with respect to the moneyness level, defined as strike-to-forward ratio  $k = K/F$ .

As the procedure involves fitting curves to the volatility smiles for specific maturities, the resulting panel is less sensitive to market microstructure noise than the raw prices.<sup>5</sup> Details about the interpolation procedure we employ are provided in Appendix B.3. Finally, for the non-parametric spot volatility estimates, we exploit high-frequency index return data. More specifically, to obtain spot volatility estimates, we use one-minute index return series<sup>6</sup>

<sup>5</sup>We therefore expect no material impact on our main results from using mid-quotes rather than alternative high-frequency fair-price estimators such as proposed by Stoikov (2018).

<sup>6</sup>Non-parametric volatility estimates from one-minute returns are not systematically higher than those based on two- and five-minutes returns. Also, one-minute return series exhibit insignificant positive auto-correlation. These features suggest that microstructure noise is not strongly present in the return series.

from the beginning of each trading day until, but not including, the reference interval for each pair. Thus, the time periods used to obtain non-parametric volatility estimates do not overlap with the option recordings’ time intervals, which allows us to use the model approximation proposed in Section 2.2.

## 4 Estimation Procedure

In this section, we develop the estimation procedure used for the data analysis. To obtain estimates of the spot volatility values  $v_{i,t}$ , we employ a jump-robust spot volatility estimator based on high-frequency returns observed before time  $t$ , with adaptive thresholding as in Bollerslev and Todorov (2011). Appendix C.1 provides details about this estimator. These estimators have been shown to be consistent under a typical in-fill asymptotic scheme, and to be robust in applications and in simulations. Given the spot volatility estimates, parameter estimation involves optimization of a GMM-type criterion function, the evaluation of which consists of two stages.

In the first stage, we back out the parameter-dependent jump intensities—the unobserved part of the state vector—using the option-pricing relation: as option prices are functions of the state variables, we can exploit this functional form to recover the latent states from the market observables, given a set of model parameter values. In the second stage, we evaluate the criterion function given this set of parameter values and the state vector consisting of observed index prices and implied jump intensities. The general method of using implied variables in a GMM-type estimation procedure was coined “implied-state GMM” by Pan (2002), who used *standard* GMM based on univariate option-implied volatilities.

In contrast to Pan (2002), we design an estimation procedure based on GMM with a *continuum* of moment conditions (C-GMM). C-GMM was initiated in the innovative work of Carrasco and Florens (2000, 2002), in a setting without latent states. The use of a

continuum of moments allows us to exploit more information than standard GMM with a finite number of moment conditions, which should result in more reliable and efficient estimates. That is, our point of departure is a novel combination of implied-state GMM and C-GMM. Beyond this combination, our methodological contribution is three-fold: (a) we develop a semi-parametric approach, which allows the parametric identification of jump contagion to be robust against misspecification of diffusive volatility; (b) we develop a partial-information version of implied-state C-GMM that mitigates the exponentially increasing computational complexity associated with larger-dimensional state vectors; and (c) we take estimation uncertainty from state-implied moments into account when computing (asymptotic or bootstrap) standard errors. These contributions demand an extension of the existing asymptotic theory for C-GMM, developed in the appendix. Our main contributions lie in the comprehensive data analysis. The semi-parametric approach requires a consistency result for non-parametric spot volatility estimation, available in the literature; our paper does not make a contribution to the specific spot volatility estimation literature. Throughout this section, we assume that for each of the markets  $i = 1, \dots, m$ , and at regular-interval (daily) observation times  $t = 0, 1, \dots, T$ , we observe a vector of (maturity- and moneyness-dependent) market-traded option prices  $p_{i,t}$ , the forward price on the index  $F_{i,t}$ , and the spot volatility estimate  $\hat{v}_{i,t}$ . Note that our empirical strategy does not require the identification of jump events from high-frequency index data; the daily option price panel is sufficient to identify jump propagation.

## 4.1 Implying the Latent States

The first stage in our estimation procedure consists of backing out the latent jump intensities from the option prices given a parameter vector  $\theta$ . Let us define the option pricing relation determining a stacked vector of option prices  $p_t = (p'_{1,t}, \dots, p'_{m,t})'$ , with

$\tau_t = (\tau'_{1,t}, \dots, \tau'_{m,t})'$  the corresponding time to maturity and  $k_t = (k'_{1,t}, \dots, k'_{m,t})'$  the moneyness level, given the vector of forward prices  $F_t$  and jump intensities  $\lambda_t$ , model parameters  $\theta$ , and volatility estimates  $\hat{v}_t$ , as follows:

$$p_t = \mathcal{P}(F_t, \lambda_t, \theta, \hat{v}_t, \tau_t, k_t), \quad (7)$$

with  $\mathcal{P} : \mathbb{R}_+^m \times \Lambda \times \Theta \times \mathbb{R}_+^m \times \mathbb{R}_+^{mn\tau} \times \mathbb{R}_+^{mnk} \rightarrow \mathbb{R}_+^{mn\tau n_k}$ . Here  $\Lambda \subseteq \mathbb{R}_+^m$  is the domain of the jump intensities and  $\Theta$  is a compact parameter space, such that the stationarity condition of the multivariate Hawkes process is satisfied. This requires the spectral radius of the matrix consisting of the entries  $(\delta_{ij}/\kappa_i)$ ,  $i, j = 1, 2$ , to be less than unity. Note that we use several option prices with different characteristics at any time  $t$  and within any market  $i$ , i.e.,  $\tau_{i,t} \in \mathbb{R}_+^{n_\tau}$  and  $k_{i,t} \in \mathbb{R}_+^{n_k}$ , where  $n_\tau$  and  $n_k$  represent the number of different maturities and moneyness levels, respectively.

We exploit the option-pricing relation (7) to imply the latent jump intensities. Formally, let us define the domain of invertibility of the option-pricing relation  $\Sigma \subset \mathbb{R}_+^{mn\tau n_k} \times \Theta \times \mathbb{R}_+^m \times \mathbb{R}_+^{mn\tau} \times \mathbb{R}_+^{mnk}$ , such that it is a maximal set for which a mapping  $f : \Sigma \rightarrow \Lambda$  is uniquely defined by

$$p_t = \mathcal{P}(F_t, f(p_t, F_t, \theta, \hat{v}_t, \tau_t, k_t), \theta, \hat{v}_t, \tau_t, k_t). \quad (8)$$

Therefore, assuming that the inversion is well-defined, the option-implied jump intensities are defined by:

$$\lambda_t^\theta = f(p_t, F_t, \theta, \hat{v}_t, \tau_t, k_t), \quad (9)$$

where we use the superscript  $\theta$  to emphasize the dependence of the implied intensity on the parameter vector  $\theta \in \Theta$ , keeping in mind its dependence on the volatility estimates  $\hat{v}_t$ .

Importantly, the vector of true intensities  $\lambda_t$  is retrieved based on the market-observables when evaluating the mapping (9) at the true model parameters  $\theta_0$  and using the true volatility process  $\xi_t$  (assuming correct model specification).

We refer to Appendix C.2.1 for further details on implying the jump intensities. As discussed in Section 2.1, the model admits a generalized affine jump-diffusion representation under the physical and risk-neutral probability measures. One of the important advantages of the class of affine jump-diffusions is that the conditional characteristic function (CCF) of the state vector  $X_T = (\log F_{1,T}, \dots, \log F_{m,T}, \lambda_{1,T}, \dots, \lambda_{m,T})'$  conditional on information available at time  $t$  is known in closed form (up to the solution of an ODE system) as an exponentially affine function of  $X_t$ ; see Appendix A.2. This property is exploited in Appendix C.2.1 to obtain model prices.

After having implied the jump intensities we can construct a series of observations for the global state vector  $X_t^\theta = (\log F_{1,t}, \dots, \log F_{m,t}, \lambda_{1,t}^\theta, \dots, \lambda_{m,t}^\theta)'$ , which we then use in the criterion function evaluation, discussed in the following sub-section.

## 4.2 Parameter Estimation in a Full-Information Setting

In addition to obtaining option prices, the CCF also allows us to obtain the model-implied conditional density function of the state vector based on Fourier inversion, and thus, in principle, to employ classical maximum likelihood, which provides asymptotically efficient estimators (see, e.g., Singleton (2001)). However, Fourier inversion requires multivariate numerical integration at every time point, which is computationally highly expensive in an optimization routine. Singleton (2001) proposed to use method-of-moment estimators directly in the “frequency domain” using the CCF of a state vector. Such an estimator based on the CCF and its empirical counterpart avoids the need for Fourier inversion, thus it is computationally more appealing. Furthermore, Carrasco and Florens (2002) show

that exploiting a continuum of moment conditions based on the CCF yields the asymptotic efficiency of maximum likelihood. We follow this route and develop, in our semi-parametric setting, a C-GMM estimator that we extend to allow for implied state variables.

Because C-GMM requires a stationary Markovian state, we consider a state process  $Y_t = (y_{1,t}, \dots, y_{m,t}, \lambda_{1,t}, \dots, \lambda_{m,t})'$ , which consists of daily returns  $y_{i,t} = \log F_{i,t} - \log F_{i,t-1}$  and latent jump intensities  $\lambda_{i,t}$  for each of the markets. The CCF of the stationary state vector  $Y_{t+1}$  given the information at time  $t$  can be obtained from the CCF of the non-stationary state vector  $X_{t+1}$ :

$$\phi(s, Y_t, \Delta; \hat{v}_t, \theta) := \mathbb{E}[e^{is \cdot Y_{t+1}} | \mathcal{F}_t] = \mathbb{E}[e^{is \cdot X_{t+1}} | \mathcal{F}_t] e^{-\sum_{j=1}^m is_j \log F_{j,t}},$$

with  $\Delta$  the sampling frequency of a single day. We consider the moment conditions based on the CCF of the state vector and its empirical counterpart. This involves combining the “raw” moment functions  $u(s, Y_t, Y_{t+1}; \hat{v}_t, \theta) := e^{is \cdot Y_{t+1}} - \phi(s, Y_t, \Delta; \hat{v}_t, \theta)$ , with an instrument function  $m(r, Y_t)$ , to obtain the moment function

$$h(r, s, Y_t, Y_{t+1}; \hat{v}_t, \theta) := m(r, Y_t) \cdot u(s, Y_t, Y_{t+1}; \hat{v}_t, \theta) = m(r, Y_t) (e^{is \cdot Y_{t+1}} - \phi(s, Y_t, \Delta; \hat{v}_t, \theta)),$$

and hence the moment conditions

$$\mathbb{E}[h(r, s, Y_t, Y_{t+1}; \hat{v}_t, \theta_0)] = 0, \quad \text{for all } r, s \in \mathbb{R}^{2m}. \quad (10)$$

The idea of GMM with a continuum of moments, developed in Carrasco and Florens (2000, 2002) and Carrasco et al. (2007), is to use not a discrete finite set of vectors  $s$  as arguments for the moment conditions (10), but rather to employ a full continuum of values of  $s$ . Furthermore, these authors show that also using a continuum of instruments of the form  $m(r, Y_t) = e^{ir \cdot Y_t}$  with  $r \in \mathbb{R}^{2m}$  leads to a considerable efficiency gain in estimation.

We will adopt both elements in our estimation approach. Unlike the regular C-GMM set-up, not all components of the state vector  $Y_t$  are observed in our model. However, we can exploit the option-pricing relation (7) and imply the jump intensities from the market observables as discussed in the previous sub-section. Under some additional assumptions, formally stated later, we can use the moment conditions (10) based on the state vector with implied intensities  $Y_t^\theta = (y_{1,t}, \dots, y_{m,t}, \lambda_{1,t}^\theta, \dots, \lambda_{m,t}^\theta)'$ . Let us denote the sample analogue of the moment conditions (10) based on the state vector with implied intensities as

$$h_T(\tau; \hat{v}, \theta) := \frac{1}{T-1} \sum_{t=1}^{T-1} h(\tau, Y_t^\theta, Y_{t+1}^\theta; \hat{v}_t, \theta), \quad (11)$$

with  $\tau = (r, s)' \in \mathbb{R}^{4m}$ .

In Appendix C.2.2, we provide further details on the criterion function of the implied-state C-GMM procedure in the present ‘full-information setting’. It turns out there that this procedure becomes computationally prohibitively expensive already for a bivariate model. We overcome this by developing a partial-information setting in the following sub-section.

### 4.3 Parameter Estimation in a Partial-Information Setting

Singleton (2001) notes that although full maximum likelihood (ML) estimation based on Fourier inversion of the CCF (ML-CCF) is computationally expensive in a multivariate setting, one could base estimation on the marginal conditional density functions  $f(y_{i,t+1}|Y_t; \theta)$  of the single state variable  $y_{i,t+1}$  conditional on the entire state vector  $Y_t$ . This limited-information (LML-CCF) approach requires at most  $N$  one-dimensional integrations for Fourier inversion instead of one  $N$ -dimensional integral evaluation. Therefore, a potential mild loss in asymptotic efficiency is traded off against the computational simplicity relative to the full ML-CCF approach.

A similar approach can be developed for the CCF-based C-GMM estimator, which

allows us to considerably decrease the computational costs when focusing on the marginal CCF of a single economy. Therefore, instead of constructing the criterion function from one  $2m$ -dimensional integral as in (C.4), we exploit a partial-information estimator based on the sum of  $m^2$  2-dimensional integrals. Although this approach can be described in the general multivariate setting, we apply it here to the bivariate model described in Appendix A.2, where we also provide the closed-form CCF for the bivariate model under  $\mathbb{P}$ .

Let us denote by  $Y_t^{(1)} = (y_{1,t}, \lambda_{1,t})$  and  $Y_t^{(2)} = (y_{2,t}, \lambda_{2,t})$  the *marginal market states* of the first and second economy, and by  $Y_t^{(3)} = (y_{1,t}, \lambda_{2,t})$  and  $Y_t^{(4)} = (y_{2,t}, \lambda_{1,t})$  the *marginal cross-market states*. Clearly, the marginal CCFs of the marginal states can be obtained from the joint CCF evaluated at the argument vectors  $s^{(1)} := (s_1, 0, s_3, 0)', s^{(2)} := (0, s_2, 0, s_4)', s^{(3)} := (s_1, 0, 0, s_4)'$  and  $s^{(4)} := (0, s_2, s_3, 0)',$  that is,

$$\phi^{(i)}(v, Y_t, \Delta; \hat{v}_t, \theta) := \phi(s^{(i)}, Y_t, \Delta; \hat{v}_t, \theta) = e^{\alpha^{(i)}(\Delta) + \beta_3^{(i)}(\Delta)\lambda_{1,t} + \beta_4^{(i)}(\Delta)\lambda_{2,t}}, \quad (12)$$

where  $\alpha^{(i)}(\Delta), \beta_3^{(i)}(\Delta), \beta_4^{(i)}(\Delta)$  are the solutions to the ODE system (A.10) in Appendix A.2 solved with the initial values  $s^{(i)}$  for  $i = 1, 2, 3, 4.$

Similar to the general setting based on the joint CCF, we exploit the marginal CCFs to obtain the moment conditions. In the bivariate case, instead of the moment condition in (10), we can consider four sets of “marginal” moment conditions stacked in a vector form:

$$\mathbb{E}[\mathbf{h}(\tau, t; \theta_0)] = \mathbb{E} \begin{bmatrix} h^{(1)}(\tau; \hat{v}_t, \theta_0) \\ h^{(2)}(\tau; \hat{v}_t, \theta_0) \\ h^{(3)}(\tau; \hat{v}_t, \theta_0) \\ h^{(4)}(\tau; \hat{v}_t, \theta_0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (13)$$

with  $h^{(i)}(\tau; \hat{v}_t, \theta) = m(r, Y_t^{(i)}) (e^{is \cdot Y_{t+1}^{(i)}} - \phi^{(i)}(s, Y_t; \hat{v}_t, \theta)),$  where  $i = 1, 2, 3, 4,$   $\tau = (r, s)'$  and  $r, s \in \mathbb{R}^2.$

In Appendix C.2.3, we provide additional details on the criterion function of this partial-information implied-state C-GMM procedure. Furthermore, we discuss in detail the asymptotic properties of this estimation procedure in Appendix C.3. More specifically, under Assumptions C.1–C.4 stated in the appendix—requiring stationarity and Markovianity of the process  $Y_t$ , regularity of the moment functions and their empirical counterparts, and consistency of the non-parametric spot volatility estimator—, we formally establish the asymptotic normality of our estimators and provide estimators for the asymptotic standard errors that account for the effect of implied-state moments. Consistency of the non-parametric spot volatility estimator is critical for obtaining these—essentially parametric—asymptotic properties of our semi-parametric estimator.

To analyze the finite-sample performance of our estimation procedure, we provide detailed Monte Carlo simulation results in Appendix C.4. The results, a summary of which is contained in Table 2 with full details provided in the appendix, show a good finite-sample performance of our partial-information implied-state C-GMM procedure, notwithstanding the challenging nature of the statistical problem. We find in particular that the model parameters capturing jump contagion (i.e.,  $\kappa_i$ ,  $\bar{\lambda}_i$ , and  $\delta_{ij}$ ,  $i, j = 1, 2$ ), which are of central interest in this paper, are identified with high precision. We note that our simulation design includes the semi-nonparametric approximation of Section 2.2 and the pricing errors induced by it. For comparison, Appendix C.4 also contains simulation results for (infeasible) moment conditions based on the fully parametric model assuming the spot volatilities and stochastic volatility parameters to be known. The results are similar, indicating that the “volatility freezing” approximation has relatively little effect on estimator behavior. Because the Monte Carlo analysis also shows that the asymptotic standard errors appear to be sensitive to the numerical implementation details of the relevant integrals, we will report, in the next section, bootstrap standard errors, using a parametric recursive-design bootstrap approach that is described in detail in Appendix C.5.

Table 2: Monte Carlo simulation results for the bivariate model

	$\mu_1^{\mathbb{Q}_1}$	$\sigma_1$	$\kappa_1$	$\bar{\lambda}_1$	$\delta_{11}$	$\delta_{12}$	$\mu_1$	$\eta_1$
true	-0.130	0.030	6.000	1.000	3.000	1.000	-0.040	2.000
25%	-0.133	0.027	5.520	0.924	2.685	0.925	-0.042	1.560
50%	-0.129	0.031	5.872	1.043	2.901	1.050	-0.038	2.467
75%	-0.125	0.034	6.116	1.086	3.070	1.131	-0.035	2.957
	$\mu_2^{\mathbb{Q}_2}$	$\sigma_2$	$\kappa_2$	$\bar{\lambda}_2$	$\delta_{22}$	$\delta_{21}$	$\mu_2$	$\eta_2$
true	-0.130	0.030	5.000	1.000	2.000	3.000	-0.040	2.000
25%	-0.132	0.028	4.729	0.945	1.835	2.803	-0.043	1.333
50%	-0.127	0.030	4.925	1.083	2.002	3.074	-0.039	2.237
75%	-0.123	0.033	5.073	1.175	2.135	3.323	-0.036	2.667

This table summarizes the Monte Carlo simulation results of the partial-information implied-state C-GMM procedure. True parameters and Monte Carlo sample quantiles (at 25%, 50%, and 75%) are presented on separate rows. Further details are in Appendix C.4.

## 5 Data Analysis

In this section, we describe our estimation results for the three pairs of stock market indices we consider. We also provide three model applications to further highlight the statistical and economic importance of the jump cross-excitation effect.

### 5.1 Estimation Results

Parameter estimates for the bivariate models are displayed in Table 3. Each bivariate model is estimated using the partial-information implied-state C-GMM procedure developed in Section 4.3, using synchronized daily data for the corresponding stock market indices and their options panels following Section 3. The synchronicity between markets is crucial for the identification of jump contagion in space. We note that already in the bivariate setting, our model is a rich semi-parametric model with 16 parameters to be estimated. Whereas a trivariate analysis would be interesting and is theoretically feasible, it practically reaches the limits of what can be reliably identified in finite samples. This applies in particular to the self- and cross-excitation parameters, which increase from 4 to 9 when going from a

Table 3: Estimation results for FTSE 100, DAX 30 and S&amp;P 500

	$\mu^Q$	$\sigma$	$\kappa$	$\bar{\lambda}$	$\delta^s$	$\delta^c$	$\mu$	$\eta$
FTSE	-0.126 [0.012]	0.020 [0.002]	4.063 [0.673]	0.353 [0.042]	1.638 [0.256]	2.506 [0.328]	-0.038 [0.003]	2.186 [0.128]
	-0.131 [0.011]	0.027 [0.003]	3.482 [0.671]	0.418 [0.044]	2.190 [0.283]	1.244 [0.196]	-0.025 [0.002]	2.680 [0.168]
DAX	-0.148 [0.01]	0.036 [0.004]	3.320 [0.751]	0.296 [0.043]	2.501 [0.349]	0.517 [0.146]	-0.039 [0.004]	2.173 [0.145]
	-0.131 [0.011]	0.033 [0.003]	3.216 [0.657]	0.283 [0.035]	1.709 [0.323]	2.119 [0.439]	-0.041 [0.003]	2.051 [0.117]
S&P	-0.135 [0.016]	0.036 [0.005]	3.781 [0.749]	0.261 [0.05]	2.257 [0.273]	1.788 [0.263]	-0.037 [0.006]	1.977 [0.285]
	-0.138 [0.019]	0.039 [0.009]	4.235 [0.803]	0.394 [0.053]	2.287 [0.373]	1.658 [0.235]	-0.035 [0.007]	2.119 [0.205]

This table reports bivariate model parameter estimates for three pairs of stock market indices: FTSE 100-DAX 30, S&P 500-FTSE 100, and S&P 500-DAX 30. The  $\delta^s$  parameters capture self-excitation for each index based on pairwise estimation (i.e.,  $\delta_i^s = \delta_{ii}$ ,  $i = 1, 2$ ), while the  $\delta^c$  parameters capture cross-excitation for each pair (i.e.,  $\delta_i^c = \delta_{ij}$ ,  $i, j = 1, 2$ ,  $i \neq j$ ). Bootstrap standard errors are reported in square brackets.

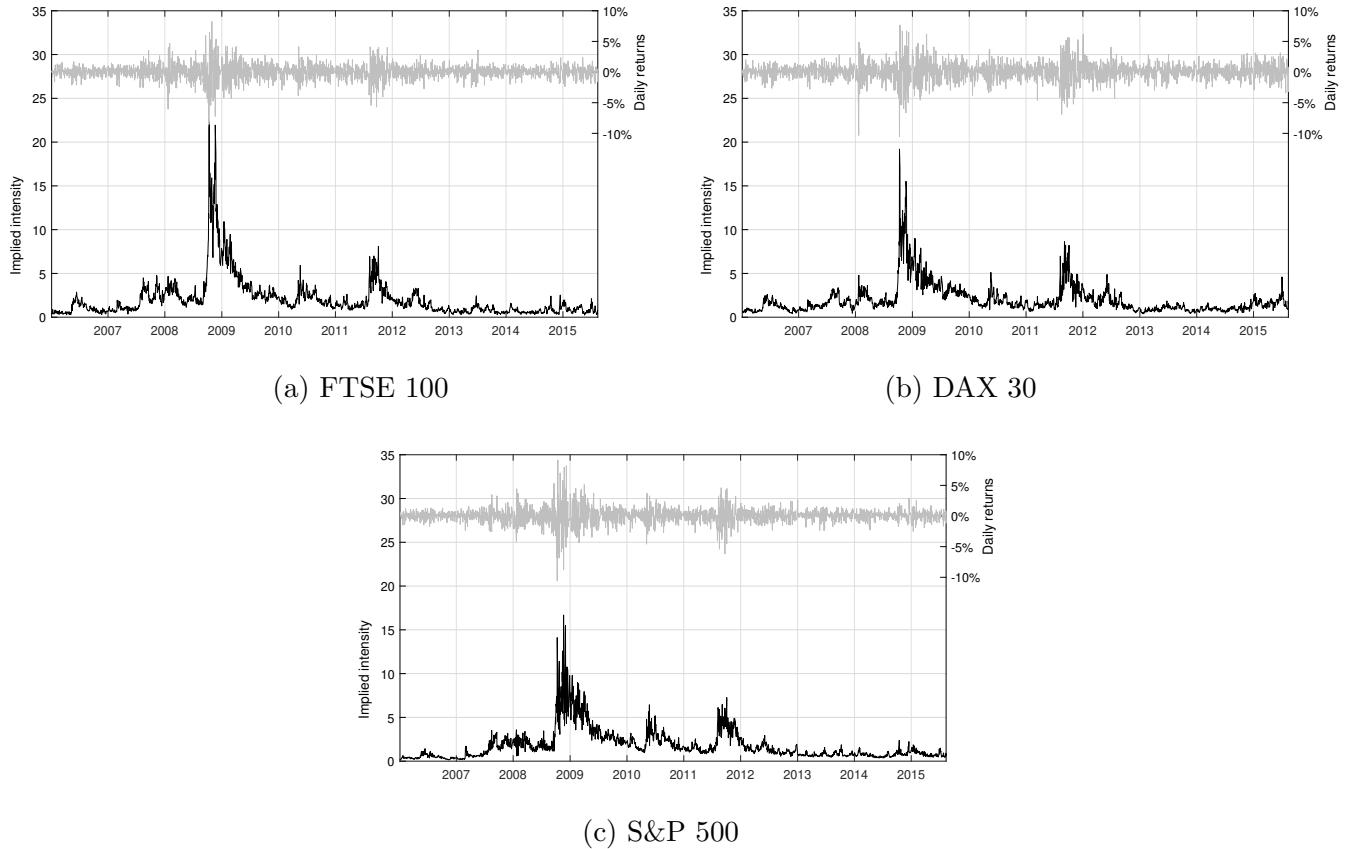
bivariate to a trivariate analysis and which are of central interest in the paper.

The estimation results provide statistically significant evidence of both self- and cross-excitation in jumps for all three markets. According to our estimates, a single jump event leads to an increase in the corresponding own jump intensity ranging from  $\delta^s = 1.6$  to 2.5 in the markets considered, given base rates  $\bar{\lambda}$  ranging from 0.3 to 0.4. This self-excitation of jumps induces jump clustering in time.<sup>7</sup> Estimates of the cross-excitation parameter  $\delta^c$  range from 0.5 to 2.5. From our cross-excitation estimates, we deduce that the UK market is about four times as much exposed to shocks in the US market than *vice versa*. In other words, we observe a large asymmetry in the jump contagion among FTSE and S&P stock market indices, in line with conventional wisdom that the US market plays a leading role in international financial markets. On the other hand, cross-excitation in jumps between the US and German stock market indices is largely symmetric; in particular, the

<sup>7</sup>The self-excitation of jumps is broadly in line with the findings in Boswijk, Laeven, and Lalu (2015) and Du and Luo (2019), who studied univariate self-excitation models with parametric volatility dynamics in the US market index using weekly data.

cross-excitation effect from Germany to the US is stronger than suggested by conventional wisdom. The cross-excitation effect from DAX to FTSE has the largest cross-excitation parameter estimate. The reverse effect from FTSE to DAX is estimated to be twice as small. We find that the Wald tests for each pair of indices reject the null hypothesis that the cross-excitation parameters are equal to zero.<sup>8</sup>

Figure 2: Time-series of the option-implied jump intensities



Note: This figure plots the time-series of option-implied jump intensities for FTSE 100, DAX 30 and S&P 500 stock market indices along with corresponding log-forward returns (secondary, right-hand axis in each subplot). The parameter estimates from the S&P-FTSE pair are used to imply the latent jump intensities for FTSE 100, while the estimates for the S&P-DAX pair are used to back out the jump intensities for DAX 30 and S&P 500.

Using the model parameter estimates in Table 3, we imply the latent jump intensities for each index from the corresponding sets of option prices. Figure 2 plots the implied

<sup>8</sup>It is conceivable that markets with the richest option information (the US, and, to a lesser degree, Germany, see Table 1, Panel A) have the most pronounced (and most precise) estimation results. This difference between the three markets could be reduced by considering a subset of option contracts for a specific strike price range. This would, however, lead to such a loss of information that parameter identification would be jeopardized. In particular, identification of jump intensities and their dynamics requires option information over the full range of strike prices.

Table 4: Option prices: Empirical fit

$k$	0.85	0.87	0.89	0.91	0.93	0.95	0.97	0.99	1.01	1.03	1.05	1.07	1.09	Total
FTSE. <i>b</i>	1.62	1.51	1.56	1.70	1.65	1.26	0.82	1.45	2.25	2.68	2.85	3.01	3.25	2.03
FTSE. <i>u</i>	1.75	1.57	1.52	1.62	1.61	1.36	1.13	1.66	2.38	2.84	3.11	3.40	3.71	2.20
S&P. <i>b</i>	1.82	1.76	1.77	1.73	1.47	0.98	0.95	1.86	2.55	2.81	2.96	3.24	3.76	2.26
S&P. <i>u</i>	1.96	1.95	1.92	1.76	1.39	1.02	1.51	2.50	3.04	3.15	3.21	3.42	3.94	2.50

This table reports the root mean square errors (RMSEs, displayed as a percentage) of the option prices written on the FTSE 100 and S&P 500 indices, expressed in terms of the market-observed and model-implied Black-Scholes implied volatility, as a function of the strike-to-forward ratio  $k = K/F$ , using the *bivariate* (.*b*) and *univariate* (.*u*) models and parameter estimates.

jump intensities, along with the index log-forward returns. To back out the jump intensity for the UK stock market index we use the parameter estimates for the pair S&P-FTSE, while for the US and German markets we use the S&P-DAX pair estimates. We note that the jump intensity time series implied using parameter estimates from other pairs exhibit very similar dynamics with only minor differences in level. The jump intensities for all three markets follow a similar pattern: in our data sample, the time series of latent jump intensities backed out from option prices start at values close to the corresponding base rate intensities, spike in the fall of 2008 during the global financial crisis, increase during the European sovereign debt crisis, gradually decay towards the base rates after each of these events, and exhibit relatively stable dynamics afterwards.

In our model set-up, the jump risk premia are driven by the difference in means between the jump sizes under the physical and risk-neutral probability measures, i.e., they are specified as  $(\mathbb{E}[J] - \mathbb{E}^{\mathbb{Q}}[J])\lambda_t$  per unit of time. The estimated jump risk premium coefficients,  $\mathbb{E}[J] - \mathbb{E}^{\mathbb{Q}}[J]$ , for the bivariate models are around 8.0%, 9.0% and 9.5% for the UK, German and US stock market indices, respectively. We note that these coefficients are commensurate with the instantaneous level of the corresponding jump intensity process. Thus, the dynamics of the jump risk premia are time-varying and are increasing during turbulent periods together with the intensity processes.

To understand the economic channel underlying our jump contagion results, Table D.1

reports subsample estimates for the first half of the sample. This includes the Global Financial Crisis of 2008, but excludes the Eurozone Debt Crisis of 2011–2012. The results show that, in this subsample, the jump contagion effects among US and German financial markets are unidirectional, from the US to Germany only. Only when we include the second half of the sample, the jump contagion effects among the US and Germany are bidirectional and on equal footing. Relative sizes of the two financial markets are similar in the full and first half of the sample; hence, size does not explain these differences. The jump contagion effects among the US and UK are already asymmetric (primarily from the US to the UK) during the full sample period, and even more so when estimated from the first half of the sample. Germany has the largest economy of the Eurozone of which it accounts for nearly a third. By contrast, the UK does not belong to the Eurozone. Hence, the German equity index can be considered representative for the broader Eurozone, a market that has played a pivotal role in the financial crises of the second half of the sample period we analyze.<sup>9</sup>

To illustrate the empirical fit of the option prices that our model achieves, we display in Table 4 the root mean square errors (RMSEs) of the option prices. In particular, we consider the bivariate parameter estimates of the pair S&P-FTSE, and display the RMSEs obtained by comparing the market-observed and model-implied option prices. The table demonstrates that the model, estimated for the full sample spanning January 1, 2006, to August 13, 2015, fits well the observed option prices. The table also anticipates the (nearly uniformly positive) gains in fit of the bivariate model compared to its univariate counterpart discussed in the next subsection for both indices. In Appendix D.1, we also

---

<sup>9</sup>These results are broadly in line with Diebold and Yilmaz (2015) who find that trans-Atlantic equity *volatility* connectedness between major financial *institutions* during 2007–2008 is unidirectional, from the US to Europe (including both Germany and the UK); this pattern changed in 2009, when the Eurozone Debt Crisis started to unfold. Furthermore, they find that directional connectedness from European to US financial institutions increased sharply in 2011 for Germany, France, Italy and Spain, but *not* for the UK. Furthermore, Guo, Sanni, and Yu (2022) find, using quarterly data over a much longer sample period of 1974Q1 to 2020Q4, that US state variables dominate the state variables in the other G7 countries in explaining the conditional equity premium of international markets, showing that the US predictor variables jointly forecast most international market returns.

Table 5: Univariate model estimation results for FTSE 100, DAX 30 and S&amp;P 500

	$\mu^Q$	$\sigma$	$\kappa$	$\bar{\lambda}$	$\delta$	$\mu$	$\eta$
FTSE	-0.127	0.030	2.132	0.318	1.798	-0.030	2.379
DAX	-0.137	0.032	3.207	0.486	2.132	-0.029	2.109
S&P	-0.161	0.043	2.445	0.305	2.176	-0.038	2.216

This table reports parameter estimates for the univariate model for FTSE 100, DAX 30 and S&P 500 stock market indices.

illustrate the empirical fit of the moment conditions within the partial-information implied-state C-GMM procedure. There, we also show that the parameter values, and the gains in fit of the bivariate model, are reasonably stable when estimated from a subsample.

## 5.2 Applications

We illustrate the statistical and economic implications of jump contagion in three applications. In the classical GMM tradition, we provide throughout this subsection comparisons between parametric models that delineate only a subset of statistical relationships that are of particular interest. Specifically, we focus on the jump contagion channel that plays a central role in this paper. To gauge the effect of cross-excitation in the jump components across markets, we first provide estimation results for the univariate model specification. The univariate model can be seen as a nested version of the bivariate specification, in which the cross-excitation parameters are turned off. We note that for the estimation of the univariate model we use the same procedure: implied-state GMM with a continuum of moments as discussed in Section 4.2. The estimation results of the univariate models for the FTSE 100, DAX 30 and S&P 500 stock market indices are provided in Table 5. Turning off the cross-excitation channel in the jump component is likely compensated for by the other parameters of the model. For this reason we observe that, while the estimates for the remaining parameters are of the same magnitude, some differences should and do appear when comparing estimates between the univariate and bivariate models.

### 5.2.1 Distribution of Index Returns

As a first application, we consider the effect of jump contagion on the (conditional) distribution of index returns, under the physical probability measure  $\mathbb{P}$  used for risk management. For this purpose, we simulate forward prices for a pair of indices using the parameter estimates of the bivariate and univariate models from Tables 3 and 5, respectively. From the set of bivariate estimates, we use the S&P 500 and FTSE 100 parameter estimates; this pair exhibits the most pronounced jump contagion asymmetry according to our estimates.

Since the simulated distribution of log-returns is conditional on the (initial) jump intensity values, we consider five different scenarios to illustrate the effect of jump contagion. Under the base scenario (a), the initial values of the intensities are given by the corresponding estimates of the base rates  $\bar{\lambda}_1$  and  $\bar{\lambda}_2$ , while in scenarios (b)–(e) we assume the initial values to be similar to levels implied from our model during the 2008 Global Financial Crisis and the Euro Debt Crisis. Table 6 displays the empirical quantiles, skewness and kurtosis statistics as well as the expected number of jumps for the simulated log-return distributions under the bivariate and univariate models. The results are based on 100,000 random paths over a 10-day horizon simulated using an Euler scheme.

It is clearly apparent from the table that the distribution of simulated log-returns is wider (i.e., more spread out) in the bivariate model than in the univariate model for the FTSE series under all scenarios, whereas the opposite holds for S&P. A natural explanation for this is that in the bivariate model the spillover of jumps from S&P to FTSE is much more pronounced than *vice versa*, while the jump size parameters imply more extreme jump sizes under the univariate specification than under the bivariate model for S&P.

Wider distributions imply larger values of standard risk measures used for risk capital calculations such as Value-at-Risk (VaR) and Expected Shortfall (ES). To illustrate, translated into 10-days Value-at-Risk capital requirements at the 99% probability level,

Table 6: Descriptive statistics for the conditional log-return distribution (simulated using model parameter estimates, horizon  $h = 10$  days)

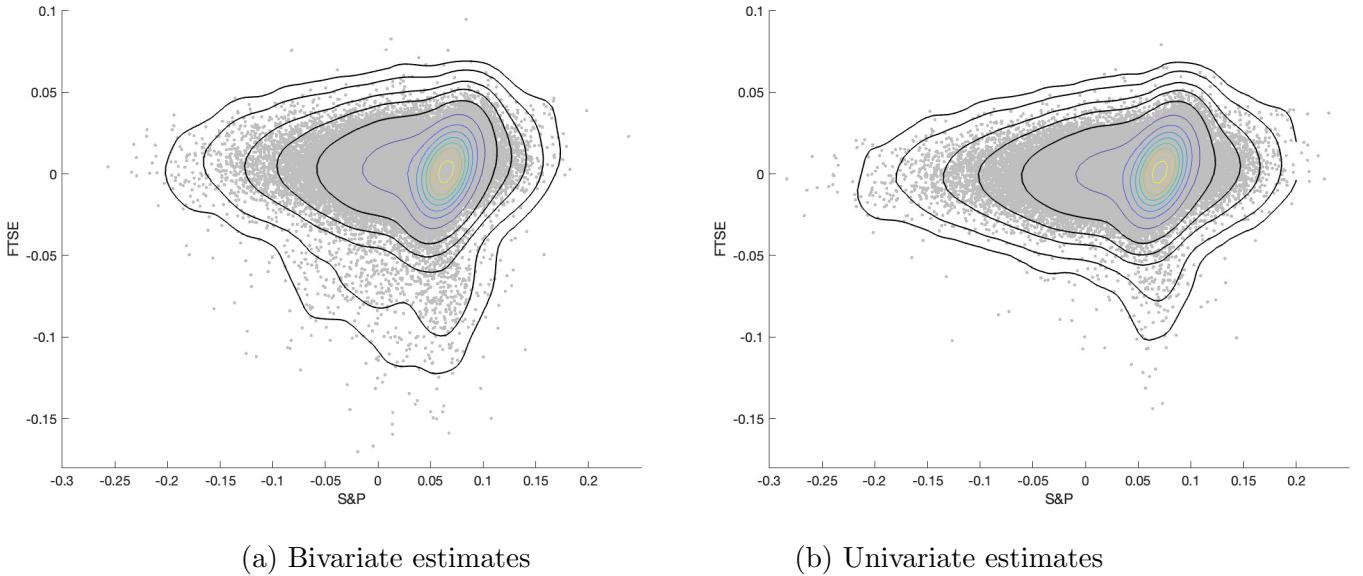
	0.1%	1%	5%	25%	50%	75%	95%	S	K	$\mathbb{E}[N_t   \lambda_0]$
(a) Base Case: $\lambda_{1,0} = \bar{\lambda}_1, \lambda_{2,0} = \bar{\lambda}_2$										
Bivariate - FTSE	-7.51	-3.17	-2.11	-0.78	0.11	1.00	2.27	-0.49	6.04	0.0071
Univariate - FTSE	-6.22	-3.11	-2.09	-0.77	0.12	1.01	2.29	-0.28	4.55	0.0077
Bivariate - S&P	-7.85	-3.21	-2.09	-0.77	0.13	1.02	2.31	-0.59	7.21	0.0075
Univariate - S&P	-8.57	-3.20	-2.08	-0.75	0.14	1.04	2.33	-0.72	8.74	0.0077
(b) Euro Debt Crisis: $\lambda_{1,0} = \lambda_{2,0} = 5$										
Bivariate - FTSE	-12.77	-7.56	-3.03	0.32	1.37	2.33	3.68	-2.15	12.07	0.1248
Univariate - FTSE	-10.63	-6.14	-2.11	0.32	1.34	2.29	3.65	-1.69	10.28	0.1234
Bivariate - S&P	-13.26	-7.57	-2.66	0.51	1.54	2.50	3.88	-2.17	12.39	0.1227
Univariate - S&P	-14.53	-8.27	-2.56	0.67	1.70	2.65	4.08	-2.29	14.06	0.1227
(c) S&P Shock: $\lambda_{1,0} = 20, \lambda_{2,0} = \bar{\lambda}_2$										
Bivariate - FTSE	-9.27	-3.73	-2.10	-0.68	0.24	1.15	2.46	-0.93	8.41	0.0194
Univariate - FTSE	-6.22	-3.11	-2.09	-0.77	0.12	1.01	2.29	-0.28	4.55	0.0077
Bivariate - S&P	-15.91	-8.41	-2.95	3.77	5.90	7.12	8.78	-1.74	7.53	0.4872
Univariate - S&P	-17.17	-8.96	-3.04	4.51	6.57	7.80	9.63	-1.74	7.89	0.4906
(d) FTSE Shock: $\lambda_{1,0} = \bar{\lambda}_1, \lambda_{2,0} = 20$										
Bivariate - FTSE	-15.80	-8.94	-3.65	2.77	5.13	6.36	7.93	-1.68	6.82	0.4838
Univariate - FTSE	-12.07	-6.38	-2.03	3.34	5.16	6.36	7.99	-1.54	6.71	0.4906
Bivariate - S&P	-8.53	-3.27	-2.07	-0.73	0.17	1.06	2.36	-0.75	8.34	0.0105
Univariate - S&P	-8.57	-3.20	-2.08	-0.75	0.14	1.04	2.33	-0.72	8.74	0.0077
(e) 2008 Global Financial Crisis: $\lambda_{1,0} = 20, \lambda_{2,0} = 15$										
Bivariate - FTSE	-15.19	-8.81	-3.88	2.28	4.00	5.16	6.69	-1.81	7.70	0.3760
Univariate - FTSE	-11.88	-6.56	-2.49	2.45	3.89	5.01	6.54	-1.65	7.62	0.3688
Bivariate - S&P	-15.93	-8.39	-2.95	3.78	5.93	7.15	8.81	-1.74	7.52	0.4895
Univariate - S&P	-17.17	-8.96	-3.04	4.51	6.57	7.80	9.63	-1.74	7.89	0.4906

This table displays the empirical quantiles (in percentages), skewness (S), kurtosis (K), and expected number of jumps implied by the conditional distribution of simulated log-returns for S&P 500 (“index 1”) and FTSE 100 (“index 2”). The stock index price paths are simulated using bivariate and univariate model parameter estimates, conditional upon different values (“scenarios”) of the latent jump intensities. The return horizon is  $h = 10$  days. Volatilities are assumed to be constant throughout the horizon and are set to  $v_{i,s} = 8.36\%$  for both indices, and the instantaneous Brownian correlation is set to be 0.6.

the effect of accounting for cross-excitation by the bivariate model, using stress scenarios similar to those in the Global Financial Crisis, implies a risk capital increase from about 6.3 to 8.4 cents for each dollar invested in the FTSE 100 index. This can be seen from panel (e), column (2) in Table 6, as  $1 - \exp(-0.0656) \approx 6.3\%$  and  $1 - \exp(-0.0881) \approx 8.4\%$ .

We also notice that the distribution of the simulated S&P 500 returns is wider than that of the FTSE 100 in all scenarios except for scenario (d), due in part to the strong

Figure 3: Contour plots, with  $\lambda_{1,0} = 20$ ,  $\lambda_{2,0} = \bar{\lambda}_2$ ,  $h = 10$



Note: Contour plots overlayed on top of scatter plots of log-return data simulated using parameter estimates for the bivariate model (panel (a)) and the two univariate models (panel (b)). Return horizon set to  $h = 10$  days. Initial jump intensities set to  $\lambda_{1,0} = 20$  for S&P 500 and  $\lambda_{2,0} = \bar{\lambda}_2$  for FTSE 100. Volatilities are assumed to be constant throughout the horizon and are set to  $v_{i,s} = 8.36\%$  for both indices, and the instantaneous correlation between Brownian increments is set to be 0.6.

self-excitation of jumps in S&P. Furthermore, the median returns on the S&P 500 are substantially larger than on the FTSE 100 in the asymmetric scenarios except for scenario (d), although the expected number of jumps in the S&P 500 is larger. This result is likely to be driven by the jump risk premia embedded in the expected returns under the physical measure. In other words, there are more jumps expected for the S&P 500, for which investors demand a larger premium to bearing this jump risk. In Appendix D.2.1, we re-compute the results of Table 6 using parameter estimates from the first half of the sample, in which we observe even more pronounced jump contagion asymmetry.

In addition to Table 6, we provide the contour plots for the model with and without cross-excitation, employing a stress scenario induced by a shock in the S&P 500 index, in Figure 3. We observe that the presence of cross-excitation in the bivariate model substantially increases the joint probability of large negative returns in both indices, compared to the situation where cross-excitation is absent (and hence the dependence is driven by the Brownian correlation only). To further analyze the statistical and economic importance

of jump contagion, we consider in Appendices D.2.2–D.2.3 two additional applications: (*i*) prices of multi-index options, and (*ii*) implied volatility dynamics, for S&P and FTSE. We find there e.g., that the strongest effects of jump contagion on multi-index option prices and implied volatility dynamics for the pair S&P-FTSE occur when mimicking a typical stress scenario that starts off in the US—the leading economy in this pair.

## 6 Conclusion

We have explored jump contagion in the laboratory of option markets. We have proposed a multivariate option pricing model to capture contagious propagation of jumps among international stock market indices. We have developed an estimation procedure exploiting the model’s conditional characteristic function. This characteristic function depends upon latent stochastic volatilities and jump intensities, and we use it both for backing out stochastic jump intensities from option prices and for the construction of a GMM criterion function based on a continuum of moments. To achieve robust identification, we have followed a semi-parametric approach, replacing spot volatilities with jump-robust realized measures obtained from high-frequency index returns. In addition, to reduce the computational complexity which increases rapidly with the dimension of the system, we have introduced a partial-information approach to implied-state continuum-of-moments GMM estimation, and established its asymptotic properties. Monte Carlo simulations have been conducted to assess the finite-sample behavior.

We have estimated the bivariate specification of our model to carefully synchronized option panels from three pairs of major international stock market indices: FTSE 100, DAX 30, and S&P 500. Our empirical results reveal the presence of significant jump contagion in these option markets. Although these contagion effects are bi-directional in all index pairs, they are partially asymmetric, with the UK being more affected by the

US and Germany than the other way around, and with the US on par with Germany. Finally, we have illustrated the importance of jump contagion for risk management, option pricing, and scenario analysis. Here we find that jump contagion is both statistically and economically relevant, with particularly strong effects in situations where the cross-excitation is asymmetric and the jump intensity in the leading economy is markedly larger than in the other economy.

## References

- Aït-Sahalia, Y., Cacho-Diaz, J., & Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3), 585–606.
- Aït-Sahalia, Y., Laeven, R. J., & Pelizzon, L. (2014). Mutual excitation in Eurozone sovereign CDS. *Journal of Econometrics*, 183(2), 151–167.
- Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance*, 53(2), 499–547.
- Aït-Sahalia, Y., & Xiu, D. (2016). Increased correlation among asset classes: Are volatility or jumps to blame, or both? *Journal of Econometrics*, 194(2), 205–219.
- Andersen, T. G., Fusari, N., & Todorov, V. (2017). Short-term market risks implied by weekly options. *The Journal of Finance*, 72(3), 1335–1386.
- Andersen, T. G., Fusari, N., & Todorov, V. (2020). The pricing of tail risk and the equity premium: Evidence from international option markets. *Journal of Business & Economic Statistics*, 38(3), 662–678.
- Bakshi, G., Carr, P., & Wu, L. (2008). Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies. *Journal of Financial Economics*, 87(1), 132–156.
- Bardgett, C., Gourier, E., & Leippold, M. (2019). Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *Journal of Financial Economics*, 131(3), 593–618.
- Bollerslev, T., & Todorov, V. (2011). Estimation of jump tails. *Econometrica*, 79(6), 1727–1783.
- Boswijk, H. P., Laeven, R. J., & Lalu, A. (2015). *Asset returns with self-exciting jumps: Option pricing and estimation with a continuum of moments* (Tech. Rep.). Amsterdam: University of Amsterdam and Tinbergen Institute.
- Broadie, M., Chernov, M., & Johannes, M. (2007). Model specification and risk premia: Evidence from futures options. *The Journal of Finance*, 62(3), 1453–1490.
- Carrasco, M., Chernov, M., Florens, J.-P., & Ghysels, E. (2007). Efficient estimation of general dynamic models with a continuum of moment conditions. *Journal of Econometrics*, 140(2), 529–573.
- Carrasco, M., & Florens, J.-P. (2000). Generalization of GMM to a continuum of moment conditions. *Econometric Theory*, 16(6), 797–834.

- Carrasco, M., & Florens, J.-P. (2002). *Efficient GMM estimation using the empirical characteristic function* (Tech. Rep.). Toulouse: Institut d'Économie Industrielle (IDEI).
- Diebold, F. X., & Yilmaz, K. (2015). Trans-Atlantic equity volatility connectedness: US and European financial institutions, 2004–2014. *Journal of Financial Econometrics*, 14(1), 81–127.
- Du, D., & Luo, D. (2019). The pricing of jump propagation: Evidence from spot and options markets. *Management Science*, 65(5), 2360–2387.
- Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6), 1343–1376.
- Dungey, M., Erdemlioglu, D., Matei, M., & Yang, X. (2018). Testing for mutually exciting jumps and financial flights in high frequency data. *Journal of Econometrics*, 202(1), 18–44.
- Errais, E., Giesecke, K., & Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1), 642–665.
- Guo, H., Sanni, S. A., & Yu, Y. (2022). *Searching for the best conditional equity premium model* (Tech. Rep.). Available at SSRN: <https://ssrn.com/abstract=4168294>.
- Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1), 83–90.
- Hull, J. C. (2021). *Options, Futures, and Other Derivatives* (11th ed.). Pearson.
- Jacod, J., & Todorov, V. (2009). Testing for common arrivals of jumps for discretely observed multidimensional processes. *The Annals of Statistics*, 37(4), 1792–1838.
- Kokholm, T. (2016). Pricing and hedging of derivatives in contagious markets. *Journal of Banking & Finance*, 66, 19–34.
- Medvedev, A., & Scaillet, O. (2007). Approximation and calibration of short-term implied volatilities under jump-diffusion. *The Review of Financial Studies*, 20(2), 427–459.
- Medvedev, A., & Scaillet, O. (2010). Pricing American options under stochastic volatility and stochastic interest rates. *Journal of Financial Economics*, 98(1), 145–159.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401), 9–27.
- 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.
- Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102(1), 111–141.
- Stoikov, S. (2018). The micro-price: A high-frequency estimator of future prices. *Quantitative Finance*, 18(12), 1959–1966.

# **Supplementary Material to**

## **“Jump Contagion among Stock Market Indices: Evidence from Option Markets”**

November 25, 2025

### **Abstract**

This supplementary material serves as an appendix to the paper “Jump Contagion among Stock Market Indices: Evidence from Option Markets”. For context, notation and definitions, see the paper. This supplement provides details concerning: *(i)* the model specification, *(ii)* the data selection and processing, *(iii)* the estimation procedure, and *(iv)* the data analysis.

## Appendix A Model Specification

### A.1 Change of Measure

This appendix provides further details on the candidate pricing kernels and the change of measure for the model specification discussed in Section 2. In particular, we show that the choice of the pricing kernel for each of the markets rules out arbitrage opportunities within each market, as well as internationally. Furthermore, we show that under the risk-neutral measures, the jump intensity dynamics are unaffected.

On our filtered probability space, we assume the existence of a stochastic discount factor process  $M_{i,t}$  that prices all assets in economy  $i$ . We consider a candidate pricing kernel  $M_{i,t}$  that has the following dynamics:

$$\frac{dM_{i,t}}{M_{i,t}} = -r_{i,t}dt - \eta_i \xi_{i,t} dW_{i,t} + \sum_{k=1}^m (U_{k,t}^i dN_{k,t} - \mathbb{E}[U_{k,t}^i] \lambda_{k,t} dt), \quad (\text{A.1})$$

where  $U_{k,t}^i$  are random jump sizes in market  $k$ , specific to pricing kernel  $i$ . That is, in order to price the jump risk in market  $i$ , we allow the pricing kernel  $M_{i,t}$  to jump simultaneously with the underlying indices of every market. We assume the relative jump sizes  $U_{k,t}^i$  in the pricing kernels to follow the same type of distribution as the index jump sizes, i.e.,  $U_{k,t}^i = e^{V_{k,t}^i} - 1$  are independently log-normally distributed with  $V_k^{i,t} \sim \mathcal{N}(a_{i,k}, b_{i,k}^2)$ . Note that  $U_{k,t}^i$  are allowed to be different from  $U_{k,t}^j$  for  $i \neq j$ , as investors in markets  $i$  and  $j \neq i$  may perceive jump events in market  $k \neq \{i,j\}$  differently, leading to different jump sizes in their corresponding pricing kernels  $M_{i,t}$  and  $M_{j,t}$ . It is assumed that  $U_{k,t}^i$  is independent of  $U_{n,t}^j$  for  $i \neq j$  and/or  $k \neq n$ , and independent of all Brownian motions, but the kernel jump log-sizes  $V_{k,t}^i$  are possibly correlated with the index jump log-sizes  $Z_{k,t}$ , with correlation coefficients  $\rho_{i,k}$ .

Similar to the univariate setting of Pan (2002), we assume the mean relative jump size in the pricing kernel  $M_{i,t}$  to be zero, i.e.,  $a_{i,k} + \frac{1}{2}b_{i,k}^2 = 0$  for  $k = 1, \dots, m$ . These constraints enable identification of the jump parameters and also set the jump-timing risk premium to zero. As we will see below, they keep the dynamics of the jump intensity processes the same under both probability measures, i.e.,  $\lambda_{k,t}^{\mathbb{Q}_i} \equiv \lambda_{k,t}$  for  $k = 1, \dots, m$ . In a more general setting,

one could allow for different intensity processes under the physical and risk-neutral measures using an additional non-trivial component in (A.1), but this would considerably increase the number of parameters to estimate and consequently weaken parameter identification.

One can show that the stochastic discount factor  $M_{i,t}$  in Eqn. (A.1) ensures that the deflated index processes  $\mathcal{S}_{i,t}^i := M_{i,t} S_{i,t} \exp(\int_0^t q_{i,s} ds)$  and the deflated money market account processes  $\mathcal{B}_{i,t} := M_{i,t} \exp(\int_0^t r_{i,s} ds)$  are local martingales. In fact, applying Itô's formula, we have:

$$d\mathcal{B}_{i,t} = \mathcal{B}_{i,t} \left( -\eta_i \xi_{i,t} dW_{i,t} + \sum_{k=1}^m U_{k,t}^i dN_{k,t} \right),$$

with  $\mathbb{E}[U_{k,t}^i] = 0$  (from the constraint  $a_{i,k} + \frac{1}{2}b_{i,k}^2 = 0$ ), and

$$d\mathcal{S}_{i,t}^i = \mathcal{S}_{i,t}^i \left[ (1 - \eta_i) \xi_{i,t} dW_{i,t} - \mathbb{E}^{\mathbb{Q}_i}[J_{i,t}] \lambda_{i,t} dt + (\exp(V_{i,t}^i + Z_{i,t}) - 1) dN_{i,t} + \sum_{k \neq i} U_{k,t}^i dN_{k,t} \right],$$

where

$$\begin{aligned} \mathbb{E}[\exp(V_i^i + Z_i) - 1] &= \exp(a_{i,i} + \frac{1}{2}b_{i,i}^2 + \mu_i + \rho_{i,i} b_{i,i} \sigma_i + \frac{1}{2}\sigma_i^2) - 1 \\ &= \exp(\mu_i^{\mathbb{Q}_i} + \frac{1}{2}\sigma_i^2) - 1 = \mathbb{E}^{\mathbb{Q}_i}[J_{i,t}], \end{aligned}$$

with  $\mu_i^{\mathbb{Q}_i} = \mu_i + \rho_{i,i} b_{i,i} \sigma_i$ . Therefore, the processes  $\mathcal{S}_{i,t}^i$  and  $\mathcal{B}_{i,t}$  are indeed local martingales.

Furthermore, in the international setting, the deflated foreign index processes and foreign money market accounts, denominated in the currency of market  $i$ , have to be local martingales as well. In other words, the processes  $\mathcal{S}_{j,t}^i := M_{i,t} E_{ij,t} S_{j,t} \exp(\int_0^t q_{j,s} ds)$  and  $\mathcal{B}_{j,t}^i := M_{i,t} E_{ij,t} \exp(\int_0^t r_{j,s} ds)$  need to be local martingales, where  $E_{ij,t}$  is the exchange rate between markets  $i$  and  $j$ , i.e., the price in currency  $i$  of one unit of currency  $j$ . This is guaranteed—and hence arbitrage opportunities across all economies are ruled out—whenever the exchange rate dynamics  $E_{ij,t}$  are such that  $M_{j,t} = M_{i,t} E_{ij,t}$  (see, for example, Brandt and Santa-Clara (2002), Backus, Foresi, and Telmer (2001)).

Therefore, arbitrage-free exchange rate dynamics can be derived from the ratio of foreign

to domestic pricing kernels:

$$\begin{aligned} dE_{ij,t} &= d\left(\frac{M_{j,t}}{M_{i,t}}\right) \\ &= E_{ij,t} [(-r_{j,t}dt - \eta_j \xi_{j,t} dW_{j,t}) - (-r_{i,t}dt - \eta_i \xi_{i,t} dW_{i,t})] \\ &\quad + E_{ij,t} \left[ (\eta_i^2 \xi_{i,t}^2 - \eta_i \xi_{i,t} \eta_j \xi_{j,t} \varrho_{ij,t}) dt + \sum_{k=1}^m \left( \frac{1+U_{k,t}^j}{1+U_{k,t}^i} - 1 \right) dN_{k,t} \right], \end{aligned}$$

where  $\varrho_{ij,t}$  is the instantaneous correlation between the Brownian motions  $W_{i,t}$  and  $W_{j,t}$ . Using the log-normal parametrization for the relative jump sizes in the pricing kernels, that is,  $U_{k,t}^i = e^{V_{k,t}^i} - 1$  with  $V_k^i \sim \mathcal{N}(a_{i,k}, b_{i,k}^2)$ , we have

$$\begin{aligned} \frac{dE_{ij,t}}{E_{ij,t}} &= (r_{i,t} - r_{j,t} + \eta_i^2 \xi_{i,t}^2 - \eta_i \xi_{i,t} \eta_j \xi_{j,t} \varrho_{ij,t}) dt + \eta_i \xi_{i,t} dW_{i,t} - \eta_j \xi_{j,t} dW_{j,t} \\ &\quad + \sum_{k=1}^m \left( e^{V_{k,t}^j - V_{k,t}^i} - 1 \right) dN_{k,t}. \end{aligned} \tag{A.2}$$

The resulting exchange rate processes feature both diffusive components with stochastic volatility and compound jump process components. In our set-up, we allow the exchange rate processes to jump simultaneously with jumps in any of the markets, and the jump sizes depend on how these jumps are perceived in the markets  $i$  and  $j$ . More specifically, due to the parametrization assumption, the exchange rate  $E_{ij,t}$  jumps simultaneously with a jump in a market  $k$  with log-jump size  $V_k^j - V_k^i \sim \mathcal{N}(a_{j,k} - a_{i,k}, b_{j,k}^2 - b_{i,k}^2)$ .

Define the equivalent martingale measure  $\mathbb{Q}_i$  in market  $i$  from the Radon-Nikodym density process  $\psi_{i,t}$ , satisfying

$$\frac{d\psi_{i,t}}{\psi_{i,t}} = -\eta_i \xi_{i,t} dW_{i,t} + \sum_{k=1}^m U_{k,t}^i dN_{k,t}. \tag{A.3}$$

Under  $\mathbb{Q}_i$ , the processes

$$W_{j,t}^{\mathbb{Q}_i} = W_{j,t} + \int_0^t \eta_i \xi_{i,s} \varrho_{ij,s} ds, \quad j = 1, \dots, m,$$

are standard Brownian motions with the same instantaneous correlations as the original Brow-

nian motions under  $\mathbb{P}$ . Note that  $\varrho_{ii,t} = 1$ , so that  $W_{i,t}^{\mathbb{Q}_i} = W_{i,t} + \int_0^t \eta_i \xi_{i,s} ds$ .

Under the defined equivalent measure  $\mathbb{Q}_i$ , the discounted foreign asset prices denominated in currency  $i$  are  $\mathbb{Q}_i$ -martingales. To see this, define  $\tilde{B}_{j,t}^i := \exp(-\int_0^t r_{i,s} ds) E_{ij,t} \exp(\int_0^t r_{j,s} ds)$  and  $\tilde{S}_{j,t}^i := \exp(-\int_0^t r_{i,s} ds) E_{ij,t} S_{j,t} \exp(\int_0^t q_{j,s} ds)$ . By applying Itô's formula, the dynamics of these processes under  $\mathbb{Q}_i$  can be characterized as follows:

$$\begin{aligned} \frac{d\tilde{B}_{j,t}^i}{\tilde{B}_{j,t}^i} &= \eta_i \xi_{i,t} dW_{i,t}^{\mathbb{Q}_i} - \eta_j \xi_{j,t} dW_{j,t}^{\mathbb{Q}_i} + \sum_{k=1}^m \left( e^{V_{k,t}^j - V_{k,t}^i} - 1 \right) dN_{k,t}, \\ \frac{d\tilde{S}_{j,t}^i}{\tilde{S}_{j,t}^i} &= (1 - \eta_j) \xi_{j,t} dW_{j,t}^{\mathbb{Q}_i} + \eta_i \xi_{i,t} dW_{i,t}^{\mathbb{Q}_i} + \left( e^{Z_{j,t} + V_{j,t}^j - V_{j,t}^i} - 1 \right) dN_{j,t} \\ &\quad - \mathbb{E}^{\mathbb{Q}_j}[J_{j,t}] \lambda_{j,t} dt + \sum_{k \neq j}^m \left( e^{V_{k,t}^j - V_{k,t}^i} - 1 \right) dN_{k,t}. \end{aligned}$$

Define  $G_t^k := \int_0^t \left( e^{V_{k,s}^j - V_{k,s}^i} - 1 \right) dN_{k,s}$  and  $H_t^j := \int_0^t \left( e^{Z_{j,s} + V_{j,s}^j - V_{j,s}^i} - 1 \right) dN_{j,s}$ . Then, given the assumptions on the zero mean relative jump sizes in the pricing kernels, i.e.,  $a_{i,k} + \frac{1}{2} b_{i,k}^2 = 0$  for  $k, i = 1, \dots, m$ , it follows that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}_i}[G_s^k] &= \mathbb{E} \left[ \psi_{i,t} G_t^k \right] \\ &= \mathbb{E} \left[ - \int_0^t \eta_i \xi_{i,s} G_s^k \psi_{i,s} dW_{i,s} + \int_0^t \psi_{i,s} \left( e^{V_{k,s}^j} - e^{V_{k,s}^i} + G_s^k (e^{V_{k,s}^i} - 1) \right) dN_{k,s} \right] \\ &= 0, \\ \mathbb{E}^{\mathbb{Q}_i}[H_t^j] &= \mathbb{E} \left[ \psi_{i,t} H_t^j \right] \\ &= \mathbb{E} \left[ - \int_0^t \eta_i \xi_{i,s} H_s^j \psi_{i,s} dW_{i,s} + \int_0^t \psi_{i,s} \left( e^{Z_{j,s} + V_{j,s}^j} - e^{V_{j,s}^i} + H_s^j (e^{V_{j,s}^i} - 1) \right) dN_{j,s} \right] \\ &= \mathbb{E} \left[ \int_0^t \mathbb{E}^{\mathbb{Q}_j}[J_{j,s}] \psi_{i,s} \lambda_{j,s} ds \right]. \end{aligned}$$

Given that

$$\mathbb{E}^{\mathbb{Q}_i} \left[ \int_0^t \mathbb{E}^{\mathbb{Q}_j}[J_{j,s}] \lambda_{j,s} ds \right] = \mathbb{E} \left[ \psi_{i,t} \int_0^t \mathbb{E}^{\mathbb{Q}_j}[J_{j,s}] \lambda_{j,s} ds \right] = \mathbb{E} \left[ \int_0^t \mathbb{E}^{\mathbb{Q}_j}[J_{j,s}] \psi_{i,s} \lambda_{j,s} ds \right],$$

it follows that the discounted processes  $\tilde{B}_{j,t}^i$  and  $\tilde{S}_{j,t}^i$  are indeed local martingales under  $\mathbb{Q}_i$ . Therefore, the pricing kernels rule out international arbitrage opportunities.

It is important to note that the jump intensity processes have the same dynamics under the defined equivalent measure  $\mathbb{Q}_i$  as under the physical probability measure. To see this, denote the compensated compound Hawkes processes by

$$\chi_{k,t} = \int_0^t J_{k,t} dN_{k,t} - \int_0^t \mathbb{E}[J_{k,s}] \lambda_{k,s} ds, \quad k = 1, \dots, m. \quad (\text{A.4})$$

The processes  $\chi_{k,t}$  are local martingales under  $\mathbb{P}$  by definition. Therefore, by the predictable version of the Girsanov-Meyer theorem (see Theorem 41 in Protter (2005)),

$$\begin{aligned} \chi_{k,t} - \int_0^t \frac{1}{\psi_{i,s}} d\langle \chi_k, \psi_i \rangle_s &= \chi_{k,t} - \int_0^t \mathbb{E}[J_{k,s} U_{k,s}^i] \lambda_{k,s} ds \\ &= \int_0^t J_{k,t} dN_{k,t} - \int_0^t (\mathbb{E}[J_{k,s}] + \mathbb{E}[J_{k,s} U_{k,s}^i]) \lambda_{k,s} ds \end{aligned}$$

is a local martingale under  $\mathbb{Q}_i$ . Using again  $J_k = e^{Z_k} - 1$  with  $Z_k \sim \mathcal{N}(\mu_k, \sigma_k^2)$  and  $U_k^i = e^{V_k^i} - 1$  with  $V_k^i \sim \mathcal{N}(a_{i,k}, b_{i,k}^2)$ , we have

$$\begin{aligned} \mathbb{E}[J_{k,s}] + \mathbb{E}[J_{k,s} U_{k,s}^i] &= \mathbb{E} \left[ e^{Z_{k,s} + V_{k,s}^i} - e^{V_{k,s}^i} \right] \\ &= \exp(a_{i,k} + \frac{1}{2}b_{i,k}^2 + \mu_i + \rho_{i,k} b_{i,k} \sigma_k + \frac{1}{2}\sigma_k^2) - \exp(a_{i,k} + \frac{1}{2}b_{i,k}^2) \\ &= \exp(\mu_k^{\mathbb{Q}_i} + \frac{1}{2}\sigma_k^2) - 1 = \mathbb{E}^{\mathbb{Q}_i}[J_{k,s}], \end{aligned}$$

with  $\mu_k^{\mathbb{Q}_i} = \mu_k + \rho_{i,k} b_{i,k} \sigma_k$ . Therefore,

$$\int_0^t J_{k,t} dN_{k,t} - \int_0^t \mathbb{E}^{\mathbb{Q}_i}[J_{k,s}] \lambda_{k,s} ds, \quad k = 1, \dots, m,$$

are  $\mathbb{Q}_i$ -local martingales, which implies, by the martingale characterization of jump intensities, that  $\lambda_{k,t}$  are intensity processes for the corresponding Hawkes processes  $N_{k,t}$  under the risk-neutral probability measure as well. In other words, the measure change in economy  $i$  does not affect the dynamics of the jump intensities  $\lambda_{k,t}$  for  $k = 1, \dots, m$ , and thus does not change jump times.

In particular, applying Girsanov's theorem using the density process  $\psi_{i,t}$ , the index  $i$

follows, under  $\mathbb{Q}_i$ ,

$$\frac{dS_{i,t}}{S_{i,t}} = (r_{i,t} - q_{i,t})dt + \xi_{i,t}dW_{i,t}^{\mathbb{Q}_i} + J_{i,t}dN_{i,t} - \mathbb{E}^{\mathbb{Q}_i}[J_i]\lambda_{i,t}dt, \quad (\text{A.5})$$

where the random jump sizes  $J_{i,t}$  have mean  $\mathbb{E}^{\mathbb{Q}_i}[J_i]$  under  $\mathbb{Q}_i$ , and  $W_{i,t}^{\mathbb{Q}_i}$  is a standard Brownian motion under  $\mathbb{Q}_i$ , given by  $W_{i,t}^{\mathbb{Q}_i} = W_{i,t} + \int_0^t \eta_s \xi_{i,s} ds$ . The jump risk premium,  $(\mathbb{E}[J_i] - \mathbb{E}^{\mathbb{Q}_i}[J_i])\lambda_{i,t}$ , is expected to be positive if the index price jumps are more negative on average under  $\mathbb{Q}_i$  than under the physical measure. Note that the jump risk premium is proportional to the intensity  $\lambda_{i,t}$ , and hence increases following a jump event in market  $i$  as well in other markets  $j$  if  $\delta_{ij} \neq 0$ . Under the equivalent martingale measure  $\mathbb{Q}_i$  in market  $i$ , the model for log-index dynamics is given by

$$\begin{cases} d \log S_{i,t} = \left( r_{i,t} - q_{i,t} - \frac{1}{2}\xi_{i,t}^2 - \mathbb{E}^{\mathbb{Q}_i}[J_i]\lambda_{i,t} \right) dt + \xi_{i,t}dW_{i,t}^{\mathbb{Q}_i} + Z_{i,t}dN_{i,t}, \\ d\lambda_{i,t} = \kappa_i(\bar{\lambda}_i - \lambda_{i,t})dt + \sum_{j=1}^m \delta_{ij}dN_{j,t}, \quad J_{i,t} = e^{Z_{i,t}} - 1, \quad Z_{i,t} \stackrel{\mathbb{Q}_i}{\sim} \mathcal{N}(\mu_i^{\mathbb{Q}_i}, \sigma_i^2), \end{cases} \quad (\text{A.6})$$

for  $i = 1, \dots, m$ . Thus indeed, the counting processes  $N_{j,t}$  for  $j = 1, \dots, m$  are not affected by the change of measure in market  $i$ , as the jump intensity processes  $\lambda_{j,t}$  have the same dynamics under  $\mathbb{Q}_i$  as under the physical measure.

## A.2 The Bivariate Specification and Conditional Characteristic Function

In the empirical analysis, we focus on the bivariate specification, i.e.,  $m = 2$ . In this appendix, we provide its explicit form and the corresponding conditional characteristic functions needed for option pricing and parameter estimation.

We reformulate the bivariate model in terms of *log-forward* prices,  $\log \tilde{F}_{i,t}$ . Given the piece-wise constant volatility processes  $v_{i,t}$ , their dynamics under the physical measure  $\mathbb{P}$  are

given by

$$\begin{cases} d \log \tilde{F}_{1,t} = \left( (\eta_1 - \frac{1}{2}) v_{1,t}^2 - \mathbb{E}^{\mathbb{Q}_1}[J_1] \lambda_{1,t} \right) dt + v_{1,t} dW_{1,t} + Z_{1,t} dN_{1,t}, \\ d \log \tilde{F}_{2,t} = \left( (\eta_2 - \frac{1}{2}) v_{2,t}^2 - \mathbb{E}^{\mathbb{Q}_2}[J_2] \lambda_{2,t} \right) dt + v_{2,t} dW_{2,t} + Z_{2,t} dN_{2,t}, \\ d\lambda_{1,t} = \kappa_1 (\bar{\lambda}_1 - \lambda_{1,t}) dt + \delta_{11} dN_{1,t} + \delta_{12} dN_{2,t}, \\ d\lambda_{2,t} = \kappa_2 (\bar{\lambda}_2 - \lambda_{2,t}) dt + \delta_{21} dN_{1,t} + \delta_{22} dN_{2,t}. \end{cases} \quad (\text{A.7})$$

Replacing the spot volatilities by their non-parametric estimates, the state vector governing the bivariate option price dynamics is given by  $X_t = (\log \tilde{F}_{1,t}, \log \tilde{F}_{2,t}, \lambda_{1,t}, \lambda_{2,t})'$ .

Given the market-specific pricing kernels  $M_{i,t}$ , index options are priced separately under the risk-neutral measures  $\mathbb{Q}_1$  and  $\mathbb{Q}_2$  for the first and second market, respectively. The dynamics of the bivariate model under  $\mathbb{Q}_1$  or  $\mathbb{Q}_2$  can be written as a special case of the multivariate setting (A.6), following the discussion in Section 2.1 and Appendix A.1, and are semi-nonparametrically approximated following Section 2.2.

Importantly, the model specification under both risk-neutral probability measures stays within the affine jump-diffusion class in the general setting developed in Appendix B of Duffie, Pan, and Singleton (2000). The conditional characteristic function (CCF) of the state vector can therefore be obtained in closed form up to the solution of a system of ordinary differential equations. This allows to efficiently price options in each market using numerical integration methods, employing the marginal CCF of the corresponding log-forward price. In our empirical analysis, we use the COS method proposed by Fang and Oosterlee (2008) to efficiently price European options.

For example, the marginal CCF of the first log-forward price under the corresponding risk-neutral measure  $\mathbb{Q}_1$  is given in closed form by (see Proposition A.1 below for a genuinely bivariate result):

$$\begin{aligned} \phi^{\mathbb{Q}_1}(s_1, X_t, T-t; v_t, \theta) &:= \mathbb{E}^{\mathbb{Q}_1} \left[ e^{is_1 \cdot \log \tilde{F}_{1,T}} \middle| \mathcal{F}_t \right] \\ &= e^{\alpha(T-t) + \beta_1(T-t) \log F_{1,t} + \beta_3(T-t)\lambda_{1,t} + \beta_4(T-t)\lambda_{2,t}}, \end{aligned} \quad (\text{A.8})$$

where  $s_1 \in \mathbb{R}$  is the argument of the CCF,  $\theta$  is the vector of parameters, and  $\alpha(T - t)$  and  $\beta(T - t)$  are the solutions to the following system of ODEs:

$$\begin{cases} \dot{\beta}_1(u) = 0, \\ \dot{\beta}_3(u) = -(\exp(\mu_1^Q + \frac{1}{2}\sigma_1^2) - 1)\beta_1 - \kappa_1\beta_3 + \exp(\mu_1^Q\beta_1 + \frac{1}{2}\sigma_1^2\beta_1^2 + \delta_{11}\beta_3 + \delta_{21}\beta_4) - 1, \\ \dot{\beta}_4(u) = -\kappa_2\beta_4 + \exp(\delta_{12}\beta_3 + \delta_{22}\beta_4) - 1, \\ \dot{\alpha}(u) = -\frac{1}{2}v_{1,t}^2\beta_1 + \kappa_1\bar{\lambda}_1\beta_3 + \kappa_2\bar{\lambda}_2\beta_4 + \frac{1}{2}v_{1,t}^2\beta_1^2, \end{cases} \quad (\text{A.9})$$

$0 \leq u \leq T - t$ , with initial conditions  $\beta_1(0) = is_1$ ,  $\beta_3(0) = 0$ ,  $\beta_4(0) = 0$  and  $\alpha(0) = 0$ ; for notational convenience, the time dependence in  $\beta(u)$  has been omitted from the right-hand side expressions in (A.9). Note that this ODE system does not involve the Brownian price of risk coefficients, the instantaneous correlation coefficient, or the jump size parameters of the second index. An explicit analytic solution of (A.9) is not possible due to the non-linear components involved in the ODE for  $\beta_3(u)$  and  $\beta_4(u)$ . Therefore, we solve this system numerically. Recall that due to the adopted approximation,  $v_{1,t}$  is fixed to its value at time  $t$  when we price an option expiring at time  $T$ . The marginal CCF for the second index, needed to price options on the second index, can be obtained in a similar way.

The option pricing relation, while being non-linear and complex, is a key ingredient, allowing us to exploit information in option price panels about the latent jump intensity process, needed to estimate the model parameters.

Next, let  $y_{i,t} = \log F_{i,t} - \log F_{i,t-1}$ ,  $i = 1, \dots, m$ . We state the following proposition, providing the closed-form CCF for the bivariate model under  $\mathbb{P}$ , which plays a central role in the parameter estimation procedure:

**Proposition A.1** *The conditional characteristic function of the state vector  $Y_t = (y_{1,t}, y_{2,t}, \lambda_{1,t}, \lambda_{2,t})'$  under  $\mathbb{P}$  is given by*

$$\phi(s, Y_t, \Delta; \hat{v}_t, \theta) = e^{\alpha(\Delta) + \beta_3(\Delta)\lambda_{1,t} + \beta_4(\Delta)\lambda_{2,t}},$$

where  $s \in \mathbb{R}^4$  and  $\alpha(\Delta)$  and  $\beta(\Delta)$  are the solutions to the following system of ODEs:

$$\begin{cases} \dot{\beta}_1(u) = 0, \\ \dot{\beta}_2(u) = 0, \\ \dot{\beta}_3(u) = -(\exp(\mu_1^Q + \frac{1}{2}\sigma_1^2) - 1)\beta_1 - \kappa_1\beta_3 + \exp(\mu_1\beta_1 + \frac{1}{2}\sigma_1^2\beta_1^2 + \delta_{11}\beta_3 + \delta_{21}\beta_4) - 1, \\ \dot{\beta}_4(u) = -(\exp(\mu_2^Q + \frac{1}{2}\sigma_2^2) - 1)\beta_2 - \kappa_2\beta_4 + \exp(\mu_2\beta_2 + \frac{1}{2}\sigma_2^2\beta_2^2 + \delta_{12}\beta_3 + \delta_{22}\beta_4) - 1, \\ \dot{\alpha}(u) = (\eta_1 - \frac{1}{2})\hat{v}_{1,t}^2\beta_1 + \frac{1}{2}\hat{v}_{1,t}^2\beta_1^2 + \kappa_1\bar{\lambda}_1\beta_3 + (\eta_2 - \frac{1}{2})\hat{v}_{2,t}^2\beta_2 + \frac{1}{2}\hat{v}_{2,t}^2\beta_2^2 + \kappa_2\bar{\lambda}_2\beta_4 \\ \quad + \varrho_t\hat{v}_{1t}\hat{v}_{2t}\beta_1\beta_2, \end{cases} \quad (\text{A.10})$$

with initial conditions  $\beta(0) = is$  and  $\alpha(0) = 0$ .

The proof of this proposition follows from the application of the results in Appendix B of Duffie et al. (2000) to the state vector  $X_t = (\log F_{1,t}, \log F_{2,t}, \lambda_{1,t}, \lambda_{2,t})'$ , from which the CCF for  $Y_t$  can be obtained. Note that the first two ODE equations have trivial solutions  $\beta_1(u) = is_1$  and  $\beta_2(u) = is_2$ , respectively, for any  $u \in [0, \Delta]$ , while fully analytic solutions for the ODEs involving  $\dot{\beta}_3$  and  $\dot{\beta}_4$  are not available due to the non-linear terms. In the empirical analysis, we solve the system of ODEs using numerical methods, in particular, the explicit Runge-Kutta method.

## Appendix B Data Selection and Processing

This appendix provides details of the various data selection criteria and transformations applied to spot, futures and options data. First, we describe the full set of filters used to decide which option data observations were included in each reference interval. Next, we give additional details about the approach used to back out forward prices using the put-call parity. Finally, we discuss the interpolation of the Black-Scholes implied volatility surfaces.

### B.1 Option Data Selection

To select the set of options in a reference interval, we apply the following filter rule sequence:

- (i) retain recordings with message type “Trade” or “Quote”;
- (ii) retain recordings with a positive Transaction price or recordings with positive Bid and Ask prices;
- (iii) for each distinct Reuters Instrument Code (RIC) symbol retain the last Bid, Ask and Transaction price in the reference interval;
- (iv) select the Transaction price if available, otherwise calculate the mid Bid-Ask price.

The first two rules trivially filter out incomplete or erroneous recordings. The last two rules are similar to “last close” price series published by stock exchanges, which also typically prioritize trade data over submitted quotes.

To further reduce the presence of noise in the selected data (which can come from wide bid-ask spreads, or synchronicity mismatches between bid and ask quote timings), we consider a few additional filters. Complementing the aforementioned rule (iii), we have also determined for each distinct RIC the median Bid and median Ask recorded during the reference interval in order to calculate a “median spread” equal to the difference between median Ask and median Bid. We then employ the following additional filters:

- (i) drop RIC symbols only if all of the following four conditions are met (concurrently):
  - (a) the number of either Bid or Ask quotes recorded in the interval is less than or equal to 2;
  - (b) there are no trade observations available in the interval;
  - (c) the elapsed time between the last Bid and Ask is larger than 10 seconds;
  - (d) the spread between last Bid and Ask is larger than the 95% quantile of the spread.
- (ii) for each RIC symbol replace last Bid/Ask with the corresponding median Bid/Ask if all of the following three conditions are satisfied (concurrently):
  - (a) spread between last Bid and Ask is three times larger than the median spread;
  - (b) spread between last Bid and Ask is larger than 8 currency units;
  - (c) time difference between last Bid and Ask is larger than 5 seconds.

We note that, since we are utilizing tick-by-tick data, we identify high-spread options by comparing the spread between the last Ask and Bid to the median spread within the reference interval for the same option contract. We then either remove these high-spread options (under some additional conditions; see (i) above); or we replace them with the median quotes (if some additional conditions are met; see (ii) above). More generally, the first filter above removes illiquid, infrequently traded instruments with potentially unreliable quotes. The second filter aims to strike a balance between quote (un)reliability and data synchronicity (mis)matches. For FTSE 100 options, the first filter removes around 3-4% of the quote data, depending on the exact year. For the other two, more liquid index options, this percentage is lower. The second filter is active for only less than 0.5% of the quote data. We also note that in theory it is possible that there are days with no data for a given reference interval, but in practice this happens in only a very few instances, which are then excluded.

We also analyze the sensitivity of the selected option data to the choice of the reference intervals. We recall that throughout the sample we have fixed the reference interval for FTSE 100 options to 15:03-15:05 (local exchange times). Specifically, we analyze descriptive statistics for FTSE 100 options using both two-minutes and increased five-minutes reference intervals. We consider two different years within the sample, i.e., 2006 and 2008, representing different ‘volatility regimes’. We choose to focus on FTSE 100 options since they are the least liquid among the three option markets we consider and this is especially the case at the beginning of the sample. The results are reported in Table B.1. As we can see from the table, there are some differences between the two reference intervals, but they are rather marginal. In particular, there is no consistent de- or increase in the standard deviations of implied volatility. We also note that the differences between the two reference intervals for other years are even smaller. Since we aim to have synchronized data, i.e., information as close to 15:05 as possible for FTSE 100, we conclude it is sensible to consider the two-minutes reference intervals.

Table B.2, Panel A, shows that the bid-ask spreads are the smallest for S&P, while the spreads for DAX and FTSE are of a similar magnitude. Relative bid-ask spreads are also the lowest for S&P, especially deep out-of-the-money (Table B.2, Panel B).

Table B.1: Descriptive statistics for the option implied volatility data: Sensitivity to reference intervals

	2006				2008			
	15:03 - 15:05		15:00 - 15:05		15:03 - 15:05		15:00 - 15:05	
	5 < $\tau \leq 40$	40 < $\tau \leq 75$	5 < $\tau \leq 40$	40 < $\tau \leq 75$	5 < $\tau \leq 40$	40 < $\tau \leq 75$	5 < $\tau \leq 40$	40 < $\tau \leq 75$
Panel A: Aggregate number of option contracts								
0.75 < $k \leq 0.85$	1	3	6	9	157	474	236	696
0.85 < $k \leq 0.92$	30	26	73	74	873	1,108	1,123	1,323
0.92 < $k \leq 0.98$	603	703	877	977	1,868	1,572	1,949	1,597
0.98 < $k \leq 1.03$	1,303	1,207	1,480	1,457	2,033	1,374	2,035	1,376
1.03 < $k \leq 1.10$	188	263	301	413	1,530	1,670	1,703	1,773
1.10 < $k \leq 1.20$	1	0	1	1	458	734	583	930
Total	2,126	2,300	2,738	2,942	6,987	7,077	7,729	7,977
Panel B: Sample mean of implied volatility (%)								
0.75 < $k \leq 0.85$	-	25.3	30.7	21.7	53.2	48.7	48.7	44.9
0.85 < $k \leq 0.92$	24.4	18.8	22.9	19.1	45.7	37.6	42.5	36.1
0.92 < $k \leq 0.98$	16.3	15.6	16.3	15.3	34.2	31.6	33.9	31.5
0.98 < $k \leq 1.03$	13.0	13.1	12.9	12.9	29.1	28.2	29.1	28.2
1.03 < $k \leq 1.10$	13.0	12.5	12.8	12.0	30.0	26.2	29.0	25.7
1.10 < $k \leq 1.20$	-	-	-	-	41.0	33.0	38.7	31.0
Total	14.1	13.9	14.3	13.8	34.2	32.1	33.8	31.9
Panel C: Sample standard deviation of implied volatility (%)								
0.75 < $k \leq 0.85$	-	2.9	4.4	3.0	15.1	14.5	14.7	14.4
0.85 < $k \leq 0.92$	4.7	3.1	4.4	3.2	16.4	13.5	15.9	13.0
0.92 < $k \leq 0.98$	3.7	3.0	3.6	2.8	14.1	11.6	13.9	11.6
0.98 < $k \leq 1.03$	3.0	2.6	2.9	2.5	13.0	11.1	13.0	11.1
1.03 < $k \leq 1.10$	2.8	2.3	2.8	2.2	14.2	11.0	13.9	10.9
1.10 < $k \leq 1.20$	-	-	-	-	11.7	11.7	12.3	11.8
Total	3.8	3.0	3.9	3.0	15.3	13.3	15.0	13.1

This table provides descriptive statistics for filtered option data on FTSE 100 futures. The table restricts attention to two subsamples containing daily option data from January 1, 2006, to December 31, 2006 and from January 1, 2008, to December 31, 2008. Two reference intervals are considered: 15:03-15:05 and 15:00-15:05 (local exchange times). The filters employed in the data selection procedure are detailed in Appendix B.1. Observations are bucketed into two categories for time-to-maturity,  $\tau$ , and into six categories with respect to the moneyness level, defined as strike-to-forward ratio  $k = K/F$ .

Table B.2: Descriptive statistics on bid-ask spreads for the option implied volatility data

		FTSE 100		DAX 30		S&P 500	
		$5 < \tau \leq 40$	$40 < \tau \leq 75$	$5 < \tau \leq 40$	$40 < \tau \leq 75$	$5 < \tau \leq 40$	$40 < \tau \leq 75$
Panel A: Sample mean of bid-ask spread (local currency)							
$0.75 < k \leq 0.85$	2.52	3.44	3.12	3.18	0.25	0.33	
$0.85 < k \leq 0.92$	2.32	2.83	2.60	2.62	0.26	0.44	
$0.92 < k \leq 0.98$	2.22	2.94	2.03	2.70	0.38	0.60	
$0.98 < k \leq 1.03$	2.62	3.35	2.84	3.52	0.50	0.68	
$1.03 < k \leq 1.10$	2.41	2.76	2.75	2.60	0.30	0.48	
$1.10 < k \leq 1.20$	3.50	3.74	4.09	3.37	0.41	0.52	
Total	2.45	3.09	2.70	2.97	0.35	0.50	
Panel B: Sample mean of relative bid-ask spread (%)							
$0.75 < k \leq 0.85$	62.3	26.8	191.5	35.0	32.5	13.1	
$0.85 < k \leq 0.92$	29.0	9.9	46.4	8.0	15.9	6.7	
$0.92 < k \leq 0.98$	11.2	4.6	8.4	2.6	6.6	3.9	
$0.98 < k \leq 1.03$	5.7	3.2	3.3	2.0	4.1	2.6	
$1.03 < k \leq 1.10$	49.9	13.8	124.7	8.5	23.9	12.1	
$1.10 < k \leq 1.20$	129.2	56.9	574.8	275.6	56.7	43.7	
Total	22.0	11.0	150.0	40.0	18.0	14.0	

This table provides descriptive statistics on bid-ask spreads for filtered option data on FTSE 100, DAX 30 and S&P 500 futures. The sample contains daily option data from January 1, 2006, to August 13, 2015. The filters employed in the data selection procedure are detailed in Appendix B.1. Observations are bucketed into two categories for time-to-maturity,  $\tau$ , and into six categories with respect to the moneyness level, defined as strike-to-forward ratio  $k = K/F$ .

## B.2 Implying Forward Prices from Put-Call Parity Pairings

To circumvent potential issues which would arise if we were to make explicit modeling choices for future dividend yields, we follow the route described in Aït-Sahalia and Lo (1998) and back out forward prices using the put-call parity relationship and estimate our model based on log-forward returns instead of log-index returns.

More specifically, to imply forward prices, we collect for each day all the put-call pairs with the same strike price and maturity, subject to an additional constraint that there are at least two Bid and two Ask quotes for each option during the reference interval. The additional constraint on the number of quotes filters out illiquid options and ensures we obtain reliable forwards. After implying forward prices from all the available put-call pairs, we take the average of the forward prices implied from pairs with the same option maturity and use the resulting term structure of forward prices to calculate Black-Scholes implied volatilities. For this last step, we require risk-free interest rates for each market. In principle, these could also

be backed out from box spreads built from the option sets available in each interval, but this would have required an overly complicated option pairing algorithm. We therefore opted to use publicly available data-sets with daily LIBOR-US, LIBOR-GBP and EURIBOR interest rate fixings. We have used linear interpolation for these fixings where needed to match the considered option's maturity.

We also need to interpolate the forward prices implied from put-call parity pairs of observed options for each maturity. We do that by exploiting a raw interpolation of discount factors, i.e., a linear interpolation between the log of discount factors yields that  $\log D_\tau = \alpha \log D_{\tau_1} + (1 - \alpha) \log D_{\tau_2}$ , where  $D_\tau = e^{(r-q)\tau}$  and  $\alpha = \frac{\tau_2 - \tau}{\tau_2 - \tau_1}$ . Therefore, an interpolated forward price for maturity  $\tau = 40$  can be obtained as

$$F_t(\tau) = D_\tau S_t = (D_{\tau_1} S_t)^\alpha (D_{\tau_2} S_t)^{1-\alpha} = F_t(\tau_1)^\alpha F_t(\tau_2)^{1-\alpha}.$$

Given that E-Mini S&P 500 future options are American style options, we extract forward prices for these by matching put and call volatilities calculated using a binomial tree pricer which, up to a modest degree of residual pricing noise, can account for early exercise pricing premiums. We note that although our estimation procedure uses option pricing methods designed for European options, the inputs are Black-Scholes implied volatilities. Therefore, having implied volatilities from a binomial tree for American style E-Mini options, the estimation can make use of these volatilities.

### B.3 Volatility Surface Interpolation

This sub-section provides details of the standard interpolation technique we use to construct the implied volatility data panel, used as input in the estimation procedure. Using an implied volatility option panel as input for the estimation procedure has two advantages. First, it ensures a homogeneous information set is used at each sample observation time-point to imply latent jump intensities from option prices as the grid of (relative) moneyness levels and option maturities is fixed.<sup>1</sup> Second, it reduces computational costs as obtaining model-implied option

---

<sup>1</sup>The number of near-ATM price quotes is typically larger than the number of OTM option price quotes. Absent any standardization, the set of quotes used to imply latent jump intensities would over-weigh information

prices for a fixed set of maturities is computationally less-demanding. We first provide details of the filters employed to select the option price quotes from which implied volatilities are calculated. Next, we provide more information about the interpolation procedure and summary statistics for the resulting implied volatility surfaces.

Defining the moneyness level,  $k$ , as the strike-to-forward ratio, i.e.,  $k = K/F$ , we designate an option as an out-of-the-money (OTM) option if it has moneyness level  $k > 1.02$  for call options and  $k < 0.98$  for put options. We consider options to be close to at-the-money (ATM) if  $0.98 \leq k \leq 1.02$ . We designate an option as in-the-money (ITM) if it is not OTM or close to ATM. We use call options to imply volatilities when  $k > 1$ , unless a particular call option has a spread which is more than twice as large as its put counterpart, or the put counterpart was quoted closer to the temporal reference point. A mirrored condition is applied for  $k \leq 1$ . These conditions trade off the liquidity of relevant options against the synchronicity of the data points used as inputs for building volatility smiles. When building implied volatility smiles, we make sure that for each volatility smile the call (put) prices (calculated for all options using put-call parity) are monotonically decreasing (increasing) functions of  $k$ .

To construct the homogeneous option panel, the sample implied volatility points for each index option were interpolated over a fixed set of moneyness and option maturities. Having experimented with different techniques, we have decided to use an industry-standard SVI parametrization to interpolate in the moneyness dimension and then proportionally interpolated volatility slices in the maturity dimension. The SVI parametrization, proposed by Gatheral (2011), has several appealing features, which are important in our application. Popular among practitioners, the SVI model typically produces close fits for volatility quotes and, thus, can be reliably used for interpolation. Furthermore, it can also be used in cases when volatility quotes are sparse, as opposed to, for instance, kernel smoothing which we found can perform poorly in such cases. We note that our application only relies on SVI as an interpolation method akin to polynomial fit used in, for instance, Broadie, Chernov, and Johannes (2007). Its dynamics and parametrization are not in any way related to our model specification.

The standard SVI parametrization of implied total variance,  $w(x, \tau)$ , with time-to-expiry from ATM options. Using a fixed moneyness and maturity grid therefore improves the likelihood that information about tail events is extracted from options.

$\tau$  is given as a function of log-moneyness  $x = \log(k) = \log(K/F)$  and a parameter set  $\chi = \{a, b, \rho, m, \sigma\}$ :

$$w(x, \tau) = \sigma_{BS}^2(x)\tau = a + b \left( \rho(x - m) + \sqrt{(x - m)^2 + \sigma^2} \right), \quad (\text{B.1})$$

where  $a \in \mathbb{R}, b \geq 0, |\rho| < 1, m \in \mathbb{R}, \sigma > 0$  and  $a + b\sigma\sqrt{1 - \rho^2} \geq 0$ . In fact, when testing different approaches, we also considered a quadratic function to fit volatility smiles. However, the SVI parametrization most of the times displayed a better fit compared to the quadratic function. We do not treat SVI as an option pricing model per se in the sense that we do not calibrate it to all option data using a single set of parameter values. Instead we fit the functional form (B.1) independently for every reference interval and for every option maturity. This allows us to compromise between interpolating with fully flexible non-parametric approaches such as kernel smoothing and calibrating a parametric option pricing model.

To build the input for our estimation procedure, we calibrate the SVI model at every time point for two volatility slices using a quasi-explicit calibration approach as per De Marco and Martini (2009). For each day we choose two volatility slices such that times-to-maturity for the first slice  $\tau_1 \leq \tau$  and for the second  $\tau_2 > \tau$ , and  $\tau_1, \tau_2$  are the closest available maturities to  $\tau$ . After having calibrated an SVI fit for these two volatility smiles, we interpolate between these slices linearly in total variance to  $\tau$ , which we set equal to 40 days.

Table B.3 reports the root mean square errors (RMSEs) for implied volatility data based on SVI interpolations for each of the markets we consider and for different data buckets. The results show that the SVI interpolation generally has very small approximation errors, with RMSEs less than 0.5% for options with moneyness levels between 0.85 and 1.1.

The moneyness range we use for our standardized option panel at each time point is determined by the following interval rule:

$$\max\{\min\{k_1, k_2\} - 0.05, 0.85\} \leq k \leq \min\{\max\{k_1, k_2\} + 0.01, 1.1\}.$$

Although it would be better to have a fully homogeneous option panel with fixed moneyness range at every time point, there are days when the observed range is considerably narrower

Table B.3: SVI interpolation RMSEs

	FTSE 100		DAX 30		S&P 500	
	$5 < \tau \leq 40$	$40 < \tau \leq 75$	$5 < \tau \leq 40$	$40 < \tau \leq 75$	$5 < \tau \leq 40$	$40 < \tau \leq 75$
$0.75 < k \leq 0.85$	0.68	0.37	0.81	0.35	0.56	0.30
$0.85 < k \leq 0.92$	0.17	0.09	0.20	0.10	0.41	0.14
$0.92 < k \leq 0.98$	0.13	0.07	0.20	0.09	0.29	0.17
$0.98 < k \leq 1.03$	0.15	0.07	0.24	0.10	0.33	0.11
$1.03 < k \leq 1.10$	0.22	0.11	0.34	0.14	0.44	0.16
$1.10 < k \leq 1.20$	0.29	0.15	0.44	0.23	0.41	0.22
Total	0.19	0.12	0.37	0.18	0.40	0.19

This table reports the SVI interpolation RMSEs, reported as a percentage, for the filtered samples of options written on the FTSE 100, DAX 30 and S&P 500 indices. The sample consists of the daily options data covering the period January 1, 2006, to August 13, 2015. The data are interpolated for each market, each day, and each maturity slice separately.

than it is on other days. Extrapolating these narrow ranges to obtain a wider fixed moneyness range would generate unreliable information. Therefore, we limit extrapolations to a maximum of up to 5% on the left wing (relative to the ATM point) and only 1% on the right wing of each implied volatility smile. For the estimation procedure we sample from the resulting interpolated volatility fit up to 13 option implied volatilities evenly spaced between 0.85 and 1.09 moneyness levels.

In addition to the resulting homogeneity of the option price panel, our volatility surface interpolation procedure has the advantage that market microstructure noise in the original prices is averaged out: as the number of option price observations used in fitting the curves increases, the effect of observation errors will tend to zero (in probability).

## Appendix C Estimation Procedure

### C.1 Jump-Robust Volatility Estimation

This appendix provides the details of the jump-robust volatility estimation procedure. We assume that for each day  $t = 1, \dots, T$ , we observe  $n + 1$  intra-day equity prices at equidistant time points:  $S_{t-1+j/n}$ ,  $j = 0, \dots, n$  (implying that the opening price of day  $t$  equals the closing price of day  $t - 1$ ). Omitting the market-specific subscripts for notational convenience, we

denote the intra-day log-returns by

$$\Delta_j^{t,n} S = \log(S_{t-1+j/n}) - \log(S_{t-1+(j-1)/n}).$$

We use the so-called threshold estimator for realized variance, originally proposed by Mancini (2001):

$$\hat{v}_t^2 := \sum_{j=1}^n \left( \Delta_j^{t,n} S \right)^2 \mathbb{1}\{|\Delta_j^{t,n} S| \leq r_n\}, \quad (\text{C.1})$$

where  $r_n$  is some deterministic sequence, converging to 0 as  $n \rightarrow \infty$ , used as a threshold to disentangle continuous variation from the jump contribution.

This threshold estimator has been shown to be consistent for the piece-wise constant variance  $v_t^2$ ; its efficiency depends on the choice of the threshold  $r_n$ . Following Bollerslev and Todorov (2011), we consider an adaptive thresholding with  $r_n = \alpha n^{-\bar{\omega}}$  and set  $\bar{\omega} = 0.49$  and  $\alpha = 3\sqrt{\frac{1}{5} \sum_{i=1}^5 RV_{t-i}}$ , where  $RV_t$  is the realized variance estimator imposing no threshold. We base the parameter  $\alpha$  on the average of the previous five days' estimates for better option pricing performance.<sup>2</sup>

When the true spot variance is not piece-wise constant, the quadratic variation estimator (C.1) can be turned into a spot variance estimator, replacing  $n$  by a sequence  $\ell_n = \mathcal{O}(n^{1/2})$ , such that the time interval  $(t - \ell_n/n, t]$  over which the quadratic variation is estimated shrinks to the time point  $t$  as  $n \rightarrow \infty$ .

The non-parametric jump-robust volatility estimator (C.1) allows us to forego a parametric representation of the volatility processes, and focus on the estimation of the jump parameters in our multivariate option pricing model. Hence, in the estimation procedure described in Section 4, we consider a semi-nonparametrically approximated representation of the model with “frozen” spot volatilities. In our empirical analysis, we obtain the spot volatility estimates based on high-frequency data of the equity indices just prior to the observation time of the option panel.

---

<sup>2</sup>For the first day in the sample, we use  $\alpha = 3\sqrt{\min(BV_t, RV_t)}$ , where  $BV_t$  is the bipower variation estimator proposed by Barndorff-Nielsen and Shephard (2004).

## C.2 Semi-Parametric Estimation

In this appendix, we provide technical details omitted from the exposition in Sections 4.1–4.3 for brevity.

### C.2.1 Implying the Latent States

In practice, the latent jump intensities are backed out by minimizing the difference between the market-observed and model-implied option prices due to (8) at every time point. Because the Black-Scholes implied volatility function is a monotonic transformation, it is common in the literature to minimize the difference between market-observed and model-implied option prices expressed in Black-Scholes volatility terms as a form of standardization of option prices.

Formally:

$$\lambda_t^\theta = \arg \min_{\lambda} \sum_{i=1}^m \sum_{j=1}^{n_{i,t}} \left( \mathcal{IV}(F_{i,t}, \lambda_t, \theta, \hat{v}_{i,t}, \tau_{i,t}^j, k_{i,t}^j) - \mathcal{BSIV}(F_{i,t}, \tau_{i,t}^j, k_{i,t}^j) \right)^2, \quad (\text{C.2})$$

where  $n_{i,t}$  is the number of cross-sectional option prices observed at every time point in one market and where we use a superscript  $j$  for option characteristics  $(\tau_{i,t}^j, k_{i,t}^j)$  to index different options within a single market. Here, we minimize squared differences of the market-observed and model-implied Black-Scholes implied volatilities,  $\mathcal{BSIV}(\cdot)$  and  $\mathcal{IV}(\cdot)$ , respectively, for a given set of parameters  $\theta$ , jointly for all  $m$  economies.

The model prices, required for (C.2), are obtained using the numerical option pricing method of Fang and Oosterlee (2008), given the marginal characteristic function of the log index price. Note that we solve the minimization problem (C.2) at every time point, independently of previous or later points, i.e., we do not superimpose the dynamic relationship in the implied intensities  $\{\lambda_t^\theta\}_{t=1}^T$  as generated by mutually exciting jump processes. As part of the estimation procedure, we back out the latent intensities for every update in the set of parameters until a suitably chosen criterion function is minimized.

### C.2.2 Parameter Estimation in a Full-Information Setting

In order to employ a continuum of moment conditions, we define a Hilbert space with the following inner product for two complex-valued functions  $f$  and  $g$ :

$$\langle f, g \rangle = \int f(\tau) \overline{g(\tau)} \pi(\tau) d\tau, \quad (\text{C.3})$$

implying the norm  $\|f\| = \langle f, f \rangle^{\frac{1}{2}}$ , where  $\overline{g(\tau)}$  indicates the complex conjugate of  $g(\tau)$ , and  $\pi(\tau)$  is a continuous probability density function typically selected to be Gaussian. The objective function of the (first-step) GMM estimator based on a continuum of values for  $\tau$  is given by

$$Q_T(\theta) = \|h_T(\tau; \hat{v}, \theta)\|^2 = \int h_T(\tau; \hat{v}, \theta) \overline{h_T(\tau; \hat{v}, \theta)} \pi(\tau) d\tau. \quad (\text{C.4})$$

This is similar to minimizing the Euclidean norm in GMM with a finite number of moments.

Note that minimizing (C.4) defines a first-step C-GMM estimator; a second step, leading to full efficiency, was further developed in Carrasco, Chernov, Florens, and Ghysels (2007). However, the use of the implied state variable  $Y_t^\theta$  would lead the covariance operator in the second step (akin to the optimal weighting matrix in regular GMM settings) to implicitly depend upon the parameter vector, which limits the efficiency gains from a second step. Pan (2002) ignores the dependence of the implied volatility on the parameter vector when constructing an optimal weighting matrix, thereby sacrificing part of the efficiency gains. Given that our use of spot volatility estimates in the option-pricing relation already implies approximations that will make it hard to establish additional efficiency gains from a second step, we focus on first-step implied-state C-GMM estimation. We will formally establish the corresponding asymptotic distributional properties, which account for the effect of implied-state moments on estimation uncertainty.

In practice, the criterion function (C.4) has to be evaluated numerically using quadrature methods. Carrasco et al. (2007) show that introducing optimal instruments of the form  $m(r, Y_t) = e^{ir \cdot Y_t}$  does not increase the computational complexity, because all elements associated with the index  $r$  have an analytical form. Therefore, the numerical integration of (C.4) is

of the same dimension as the state vector. In other words, the instruments can be integrated out from the criterion function using a property of Fourier transforms. For a univariate version of the model, the estimation procedure would require successive numerical integration over the two-dimensional space, which can be done at reasonable computational costs. However, with the increase in dimensionality of the state vector, the computational burden increases exponentially. Even four-dimensional efficient integration, required in the bivariate setting, becomes prohibitively costly. To overcome this practical issue we consider a partial-information version of the criterion function (C.4) for the bivariate model, which we describe in sub-section 4.3.

### C.2.3 Parameter Estimation in a Partial-Information Setting

We can define the objective function of the C-GMM estimator for the bivariate set-up as follows:

$$Q_T(\theta) = \int \mathbf{h}_T(\tau; \hat{v}, \theta)' \overline{\mathbf{h}_T(\tau; \hat{v}, \theta)} \pi(\tau) d\tau = \sum_{i=1}^4 \int h_T^{(i)}(\tau; \hat{v}, \theta) \overline{h_T^{(i)}(\tau; \hat{v}, \theta)} \pi(\tau) d\tau, \quad (\text{C.5})$$

where by  $h_T^{(i)}(\tau; \hat{v}, \theta)$  we denote the sample counterpart of the  $i^{th}$  component of the moment condition (13).

Intuitively, we decompose the joint criterion function (C.4) into the sum of the four criterion functions based on the marginals.<sup>3</sup> Thus, we significantly gain in terms of computational power as the estimation problem in (C.5) requires four numerical evaluations of 2-dimensional integrals, rather than the numerical integration over 4-dimensional space as in (C.4).

Importantly, there is sufficient information in the decomposed criterion function (C.5) that allows us to identify all parameters in the bivariate model. If the cross-excitation parameters  $\delta_{12}$  and  $\delta_{21}$  were zero, and the Brownian motions  $(W_{1,t}, W_{2,t})$  were uncorrelated, then we could split the estimation procedure into two separate problems without any loss of efficiency, as the first two marginal processes  $Y_t^{(1)}$  and  $Y_t^{(2)}$ , each associated with one economy, would be independent. Having the cross-market marginals  $Y_t^{(3)}$  and  $Y_t^{(4)}$  provides additional identification if the cross-

---

<sup>3</sup> Alternatively, one could see the criterion function (C.5) as a specific choice of the probability density function  $\pi(\tau)$  in (C.4). More specifically, given  $\tau = (r, s)$  with  $r, s \in \mathbb{R}^4$ , we can define the probability density function  $\pi(s) = f(s_1, s_3)\mathbb{1}_{\{s_2=0, s_4=0\}} + f(s_2, s_4)\mathbb{1}_{\{s_1=0, s_3=0\}} + f(s_1, s_4)\mathbb{1}_{\{s_2=0, s_3=0\}} + f(s_2, s_3)\mathbb{1}_{\{s_1=0, s_4=0\}}$ , where  $f(\cdot)$  is the bivariate Gaussian density. In other words, the density function  $\pi(s)$  puts all mass on points  $(s_1, 0, s_3, 0)$ ,  $(0, s_2, 0, s_4)$ ,  $(s_1, 0, 0, s_4)$  or  $(0, s_2, s_3, 0)$ . Defining in a similar way  $\pi(r)$ , one can show that the criterion function (C.4) with this choice of density functions results in the decomposed criterion function (C.5).

excitation parameters are non-zero. We emphasize that all marginal states are conditioned on the full state vector  $Y_t$ . Therefore, the marginal CCFs exploit the information about the feedback within and between markets through the conditional moments of each  $Y_{t+1}^{(i)}$ , i.e., through  $\mathbb{E}[(Y_{t+1}^{(i)})^k | Y_t]$ , which allows us to identify the excitation parameters based on (C.5).

### C.3 Asymptotic Properties of the Estimation Procedure

In this appendix, we derive in detail the asymptotic properties of our estimators. This ultimately leads to expressions for asymptotic standard errors of the parameter estimates in our partial-information implied-state C-GMM procedure.

We start by introducing the required Hilbert space. Let  $\pi$  be a probability density function on  $\mathbb{R}^d$ . We denote by  $L^2(\pi)$  the Hilbert space of complex-valued functions such that

$$L^2(\pi) := \left\{ f : \mathbb{R}^d \rightarrow \mathbb{C} : \int |f(\tau)|^2 \pi(\tau) d\tau < \infty \right\}.$$

The inner product  $\langle \cdot, \cdot \rangle$  and the norm  $\|\cdot\|$  on  $L^2(\pi)$  are defined as

$$\langle f, g \rangle := \int f(\tau) \overline{g(\tau)} \pi(\tau) d\tau, \quad \text{and} \quad \|f\| := \langle f, f \rangle^{\frac{1}{2}},$$

where  $\overline{g(\tau)}$  denotes the complex conjugate of  $g(\tau)$ .

Let us further extend the notion of inner product for vectors of functions in  $L^2(\pi)$ . For this purpose, we first define the  $L^2(\pi)^k$  space of vector functions as

$$L^2(\pi)^k := \left\{ \mathbf{f} = (f_1, \dots, f_k)' : f_i \in L^2(\pi) \right\}.$$

Then the inner product of two (column) vector functions  $\mathbf{f} = (f_1, \dots, f_k)'$  and  $\mathbf{g} = (g_1, \dots, g_k)'$  is defined as

$$\langle \mathbf{f}, \mathbf{g} \rangle := \int \mathbf{f}(\tau)' \overline{\mathbf{g}(\tau)} \pi(\tau) d\tau = \sum_{i=1}^k \int f_i(\tau) \overline{g_i(\tau)} \pi(\tau) d\tau.$$

Similarly, for matrices  $\mathbf{F}$  and  $\mathbf{G}$  of  $L^2(\pi)$  functions, with dimensions  $k \times p$  and  $k \times d$ , respectively,

$$\langle \mathbf{F}, \mathbf{G} \rangle := \int \mathbf{F}(\tau)' \overline{\mathbf{G}(\tau)} \pi(\tau) d\tau, \text{ a } p \times d \text{ matrix.}$$

Recall that, in the full-information setting, we consider the moment function based on the CCF of the state vector  $Y_t$  and its empirical counterpart:

$$h_t(\tau; \hat{v}_t, \theta) := h(\tau, Y_t^\theta, Y_{t+1}^\theta; \hat{v}_t, \theta) = m(r, Y_t) (e^{is \cdot g Y_{t+1}} - \phi(s, Y_t, \Delta_t; \hat{v}_t, \theta)),$$

where  $\tau = (r, s)'$  with  $r, s \in \mathbb{R}^{2m}$ , and  $m(r, Y_t) = e^{ir \cdot Y_t}$  is an ‘‘instrument’’ function. However, in the partial-information setting, we have  $k$  sets of ‘‘marginal’’ moment conditions stacked in the vector

$$\mathbf{h}_t(\tau; \hat{v}_t, \theta) = \begin{pmatrix} h_t^{(1)}(\tau; \hat{v}_t, \theta) \\ \vdots \\ h_t^{(k)}(\tau; \hat{v}_t, \theta) \end{pmatrix},$$

with

$$h^{(i)}(\tau; \hat{v}_t, \theta) = m(r, Y_t^{(i)}) (e^{is \cdot Y_{t+1}^{(i)}} - \phi^{(i)}(s, Y_t, \Delta_t; \hat{v}_t, \theta)), \quad \text{for } i = 1, \dots, k,$$

where  $r, s \in \mathbb{R}^2$ , and where  $Y_t^{(i)}$  and  $\phi^{(i)}(\cdot)$  are the marginal states and marginal CCFs, respectively.

Before we state our formal convergence result, we first introduce some assumptions. We start by imposing the following assumptions on our stochastic process and moment functions:

**Assumption C.1** *The stochastic process  $Y_t$  is a stationary Markov process.*

**Assumption C.2** *The moment functions  $\mathbf{h}_t(\tau; \hat{v}_t, \theta)$  satisfy the following conditions:*

- (i)  $\mathbf{h}_t(\tau; v, \theta)$  is continuously differentiable w.r.t.  $\theta$  and  $v$ ;
- (ii)  $\mathbf{h}_t(\tau; v, \theta) \in L^2(\pi)^k, \forall \theta \in \Theta$  and  $\forall v \in \mathbb{R}_+^m$ ;
- (iii) The equation  $\mathbb{E}^{\theta_0}[\mathbf{h}_t(\tau; v_t, \theta_0)] = \mathbf{0}, \forall \tau \in \mathbb{R}^{2 \times 2m}$   $\pi$ -almost everywhere, has a unique solution  $\theta_0$  in the interior of  $\Theta$ .

For the next assumption, recall that the sample analogue of the moment conditions, given  $T + 1$  observations, is given by

$$\mathbf{h}_T(\tau; \hat{v}, \theta) := \frac{1}{T} \sum_{t=1}^T \mathbf{h}(\tau, Y_t^\theta, Y_{t+1}^\theta; \hat{v}_t, \theta).$$

**Assumption C.3** *The sample moment conditions satisfy, as  $T \rightarrow \infty$ :*

- (i)  $\sup_{\theta \in \Theta} \|\mathbf{h}_T(\cdot, v, \theta) - \mathbb{E}^{\theta_0}[\mathbf{h}_t(\cdot, v_t, \theta)]\| \xrightarrow{P} 0$ ;
- (ii)  $\sqrt{T}\mathbf{h}_T(\tau; v, \theta_0) \xrightarrow{d} \mathcal{N}(0, \mathbf{K})$  on  $L^2(\pi)^k$ , where  $\mathcal{N}(0, \mathbf{K})$  is the distribution of an  $n$ -dimensional Gaussian random element of  $L^2(\pi)^k$  with mean zero and covariance operator  $\mathbf{K}$ , the Hilbert-Schmidt operator, defined by

$$\mathbf{K} : L^2(\pi)^k \rightarrow L^2(\pi)^k, \quad \mathbf{K}\mathbf{f}(\tau_1) := \int \mathbf{k}(\tau_1, \tau_2)\mathbf{f}(\tau_2)\pi(\tau_2)d\tau_2, \quad (\text{C.6})$$

with kernel  $\mathbf{k}(\tau_1, \tau_2) := \mathbb{E}^{\theta_0} \left[ \mathbf{h}_t(\tau_1; v_t, \theta_0) \overline{\mathbf{h}_t(\tau_2; v_t, \theta_0)} \right]$ .

Note that in the partial-information setting, the kernel  $\mathbf{k}(\tau_1, \tau_2)$  is a  $k \times k$  matrix function with  $(i, j)^{\text{th}}$  element  $\mathbb{E}^{\theta_0} \left[ h_t^{(i)}(\tau_1; v_t, \theta_0) \overline{h_t^{(j)}(\tau_2; v_t, \theta_0)} \right]$ .

Finally, we impose the following condition on the non-parametric spot volatility estimator:

**Assumption C.4** *The non-parametric spot volatility estimator  $\hat{v}_t$ , defined from  $\ell_n = \mathcal{O}(n^{1/2})$  high-frequency returns prior to time  $t$  (with  $n$  the number of intra-day observations), satisfies the conditions of Theorem 8.7 of Aït-Sahalia and Jacod (2014), with  $\tau = \frac{1}{2}$ . Furthermore,  $T/n \rightarrow 0$  as  $(T, n) \rightarrow \infty$ .*

This assumption is required for the estimation error in  $\hat{v}_t$  to be negligible in the (large  $(T, n)$ ) asymptotic properties of the estimator.

Recall that the criterion function for the C-GMM estimator  $\hat{\theta}$  is given by

$$Q_T(\hat{v}, \theta) = \|\mathbf{h}_T(\cdot, \hat{v}, \theta)\|^2 = \int \mathbf{h}_T(\tau, \hat{v}, \theta)' \overline{\mathbf{h}_T(\tau, \hat{v}, \theta)} \pi(\tau) d\tau.$$

We are now equipped to state the following proposition:

**Proposition C.1** *Under Assumptions C.1–C.4, as  $T \rightarrow \infty$ ,*

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}),$$

where

$$\begin{aligned}\mathbf{A} &:= \left\langle \mathbb{E}^{\theta_0}[\nabla_{\theta}\mathbf{h}_t(\cdot, v, \theta_0)], \mathbb{E}^{\theta_0}[\nabla_{\theta}\mathbf{h}_t(\cdot, v, \theta_0)] \right\rangle, \\ \mathbf{B} &:= \left\langle \mathbb{E}^{\theta_0}[\nabla_{\theta}\mathbf{h}_t(\cdot, v, \theta_0)], \mathbf{K}\mathbb{E}^{\theta_0}[\nabla_{\theta}\mathbf{h}_t(\cdot, v, \theta_0)] \right\rangle,\end{aligned}$$

with  $\mathbf{K}$  the Hilbert-Schmidt operator as defined in (C.6). The asymptotic covariance matrix of  $\hat{\theta}$  is estimated consistently by  $T^{-1}\hat{\mathbf{A}}_T^{-1}\hat{\mathbf{B}}_T\hat{\mathbf{A}}_T^{-1}$ , where

$$\begin{aligned}\hat{\mathbf{A}}_T &:= \int \nabla_{\theta}\mathbf{h}_T(\tau, \hat{v}, \hat{\theta})' \overline{\nabla_{\theta}\mathbf{h}_T(\tau, \hat{v}, \hat{\theta})} \pi(\tau) d\tau, \\ \hat{\mathbf{B}}_T &:= \int \nabla_{\theta}\mathbf{h}_T(\tau_1, \hat{v}, \hat{\theta})' \overline{\int \mathbf{k}_T(\tau_1, \tau_2) \nabla_{\theta}\mathbf{h}_T(\tau_2, \hat{v}, \hat{\theta}) \pi(\tau_2) d\tau_2} \pi(\tau_1) d\tau_1,\end{aligned}$$

$$\text{with } \mathbf{k}_T(\tau_1, \tau_2) := \frac{1}{T} \sum_{t=1}^T \mathbf{h}_t(\tau_1; \hat{v}_t, \hat{\theta}) \overline{\mathbf{h}_t(\tau_2; \hat{v}_t, \hat{\theta})}.$$

*Proof:* The consistency of the C-GMM procedure follows from Carrasco and Florens (2000) and Boswijk, Laeven, and Lalu (2015). To establish the asymptotic distribution of our estimators, we start from a mean value expansion of  $\mathbf{h}_T(\tau, \hat{v}, \hat{\theta})$ , which yields

$$\mathbf{h}_T(\tau, \hat{v}, \hat{\theta}) = \mathbf{h}_T(\tau, v, \theta_0) + \nabla_{\theta}\mathbf{h}_T(\tau, \bar{v}, \bar{\theta})(\hat{\theta} - \theta_0) + \nabla_v\mathbf{h}_T(\tau, \bar{v}, \bar{\theta})(\hat{v} - v),$$

where  $\bar{\theta}$  and  $\bar{v}$  are mean values. Note that in our implied-state GMM setting we have to take

into account both ‘‘direct’’ and ‘‘indirect’’ effects in the moment functions, i.e.,

$$\begin{aligned}
\nabla_{\theta} \mathbf{h}_T(\tau, v, \theta) &= \frac{1}{T} \sum_{t=1}^T \nabla_{\theta} \mathbf{h}(\tau, Y_t^{\theta}, Y_{t+1}^{\theta}, v_t, \theta) \\
&= \frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{h}(\tau, Y_t, Y_{t+1}, v_t, \theta)}{\partial \theta'} + \frac{\partial \mathbf{h}(\tau, Y_t, Y_{t+1}, v_t, \theta)}{\partial Y_t} \frac{\partial Y_t(\theta)}{\partial \theta'} \\
&\quad + \frac{\partial \mathbf{h}(\tau, Y_t, Y_{t+1}, \xi, \theta)}{\partial Y_{t+1}} \frac{\partial Y_{t+1}(\theta)}{\partial \theta'}, \\
\nabla_v \mathbf{h}_T(\tau, v, \theta) &= \frac{1}{T} \sum_{t=1}^T \nabla_v \mathbf{h}(\tau, Y_t^{\theta}, Y_{t+1}^{\theta}, v_t, \theta) \\
&= \frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{h}(\tau, Y_t, Y_{t+1}, v_t, \theta)}{\partial v'} + \frac{\partial \mathbf{h}(\tau, Y_t, Y_{t+1}, v_t, \theta)}{\partial Y_t} \frac{\partial Y_t(v_t)}{\partial v'} \\
&\quad + \frac{\partial \mathbf{h}(\tau, Y_t, Y_{t+1}, v_t, \theta)}{\partial Y_{t+1}} \frac{\partial Y_{t+1}(v_t)}{\partial v'},
\end{aligned}$$

where the first elements on the right-hand sides of both equations capture only the direct dependence of the moment function on  $\theta$  and  $v$ , while the remaining terms are due to the implied-state procedure.

Employing the mean value expansion in the first-order condition for optimality, we obtain

$$\begin{aligned}
0 &= \left\langle \nabla_{\theta} \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}) \right\rangle \\
&= \left\langle \nabla_{\theta} \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \mathbf{h}_T(\tau, v, \theta_0) + \nabla_{\theta} \mathbf{h}_T(\tau, \bar{v}, \bar{\theta})(\hat{\theta} - \theta_0) + \nabla_v \mathbf{h}_T(\tau, \bar{v}, \bar{\theta})(\hat{v} - v) \right\rangle,
\end{aligned}$$

so that

$$\begin{aligned}
&\sqrt{T}(\hat{\theta} - \theta_0) \\
&= - \left\langle \nabla_{\theta} \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \nabla_{\theta} \mathbf{h}_T(\tau, \bar{v}, \bar{\theta}) \right\rangle^{-1} \left\langle \nabla_{\theta} \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \sqrt{T} \mathbf{h}_T(\tau, v, \theta_0) \right\rangle \\
&\quad - \left\langle \nabla_{\theta} \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \nabla_{\theta} \mathbf{h}_T(\tau, \bar{v}, \bar{\theta}) \right\rangle^{-1} \left\langle \nabla_{\theta} \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \sqrt{T} \nabla_v \mathbf{h}_T(\tau, \bar{v}, \bar{\theta})(\hat{v} - v) \right\rangle. \quad (\text{C.7})
\end{aligned}$$

The second term on the right-hand side of (C.7) vanishes asymptotically by Assumption

C.4, which is seen as follows. For fixed  $\tau$ , we have

$$\sqrt{T} \nabla_v \mathbf{h}_T(\tau, \bar{v}, \bar{\theta})(\hat{v} - v) = \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial \mathbf{h}_t(\tau, \bar{v}_t, \bar{\theta})}{\partial \bar{v}_t} (\hat{v}_t - v_t). \quad (\text{C.8})$$

Theorem 8.7 of Aït-Sahalia and Jacod (2014) implies that, as  $n \rightarrow \infty$ ,  $\{\ell_n^{1/2}(\hat{v}_t - v_t)\}_{t \geq 1}$  converges stably to a sequence  $\{\zeta_t\}_{t \geq 1}$  that is, conditionally on  $\mathcal{F}_T$ , independent Gaussian with mean zero and finite  $\mathcal{F}_T$ -measurable variances. Rewriting the corresponding Edgeworth expansion (see Yoshida (2013)) as  $\hat{v}_t - v_t = \ell_n^{-1/2} \zeta_t + \ell_n^{-1} r_{nt}$ , with  $r_{nt} = \mathcal{O}_p(1)$ , we find that the right-hand side of (C.8) can be written as

$$\ell_n^{-1/2} \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial \mathbf{h}_t(\tau, \bar{v}_t, \bar{\theta})}{\partial \bar{v}_t} \zeta_t + T^{1/2} \ell_n^{-1} \frac{1}{T} \sum_{t=1}^T \frac{\partial \mathbf{h}_t(\tau, \bar{v}_t, \bar{\theta})}{\partial \bar{v}_t} r_{nt}.$$

As  $(T, n) \rightarrow \infty$ , the first term is  $\mathcal{O}_p(\ell_n^{-1/2})$  due to a (large  $T$ ) stable central limit theorem. Because the mean of  $\{r_{nt}\}_{t \geq 1}$  may be non-zero, the second term is  $\mathcal{O}_p(T^{1/2} \ell_n^{-1}) = \mathcal{O}_p(\sqrt{T/n})$ . Therefore, the condition  $T/n \rightarrow 0$  as  $(T, n) \rightarrow \infty$  guarantees that (C.8) converges in probability to 0, so that the second right-hand side term in (C.7) is asymptotically negligible.

For the first term on the right-hand side of (C.7), Assumption C.3 implies that

$$\left\langle \mathbb{E}^{\theta_0} [\nabla_\theta \mathbf{h}_t(\cdot, v_t, \theta_0)], \sqrt{T} \mathbf{h}_T(\tau, v, \theta_0) \right\rangle \xrightarrow{d} \mathcal{N}(0, \mathbf{B}).$$

Together with consistency and Slutsky's lemma, this yields the first part of the stated result.

We finally discuss the estimation of the standard errors. First, given the consistent estimators  $\hat{\theta}$  and  $\hat{v}$ , we obtain a consistent estimator of the matrix  $\mathbf{A}$ :

$$\begin{aligned} \hat{\mathbf{A}}_T &= \left\langle \nabla_\theta \mathbf{h}_T(\cdot, \hat{v}, \hat{\theta}), \nabla_\theta \mathbf{h}_T(\cdot, \hat{v}, \hat{\theta}) \right\rangle \\ &= \int \nabla_\theta \mathbf{h}_T(\tau, \hat{v}, \hat{\theta})' \overline{\nabla_\theta \mathbf{h}_T(\tau, \hat{v}, \hat{\theta})} \pi(\tau) d\tau \\ &= \sum_{i=1}^k \int \nabla_\theta h_T^{(i)}(\tau, \hat{v}, \hat{\theta})' \overline{\nabla_\theta h_T^{(i)}(\tau, \hat{v}, \hat{\theta})} \pi(\tau) d\tau. \end{aligned} \quad (\text{C.9})$$

Next, let us denote the estimator of the covariance operator by

$$\mathbf{K}_T \mathbf{f}(\tau_1) = \int \mathbf{k}_T(\tau_1, \tau_2) \mathbf{f}(\tau_2) \pi(\tau_2) d\tau_2, \quad (\text{C.10})$$

with kernel

$$\mathbf{k}_T(\tau_1, \tau_2) = \frac{1}{T} \sum_{t=1}^T \mathbf{h}_t(\tau_1; \hat{v}_t, \hat{\theta}) \overline{\mathbf{h}_t(\tau_2; \hat{v}_t, \hat{\theta})}.$$

Then, asymptotic standard errors of our parameter estimates are obtained as the square root of the diagonal elements of

$$T^{-1} \hat{\mathbf{A}}_T^{-1} \hat{\mathbf{B}}_T \hat{\mathbf{A}}_T^{-1}, \quad (\text{C.11})$$

where

$$\begin{aligned} \hat{\mathbf{B}}_T &= \left\langle \nabla_{\theta} \mathbf{h}_T(\cdot, \hat{v}, \hat{\theta}), \mathbf{K}_T \nabla_{\theta} \mathbf{h}_T(\cdot, \hat{v}, \hat{\theta}) \right\rangle \\ &= \int \nabla_{\theta} \mathbf{h}_T(\tau_1, \hat{v}, \hat{\theta})' \overline{\mathbf{K}_T \nabla_{\theta} \mathbf{h}_T(\tau_1, \hat{v}, \hat{\theta})} \pi(\tau_1) d\tau_1 \\ &= \int \nabla_{\theta} \mathbf{h}_T(\tau_1, \hat{v}, \hat{\theta})' \overline{\int \mathbf{k}_T(\tau_1, \tau_2) \nabla_{\theta} \mathbf{h}_T(\tau_2, \hat{v}, \hat{\theta}) \pi(\tau_2) d\tau_2} \pi(\tau_1) d\tau_1 \\ &= \sum_{i=1}^k \sum_{j=1}^k \int \nabla_{\theta} h_T^{(i)}(\tau_1, \hat{v}, \hat{\theta})' \overline{\int k_T^{(ij)}(\tau_1, \tau_2) \nabla_{\theta} h_T^{(j)}(\tau_2, \hat{v}, \hat{\theta}) \pi(\tau_2) d\tau_2} \pi(\tau_1) d\tau_1. \end{aligned} \quad (\text{C.12})$$

This proves the second part of the proposition.  $\square$

## C.4 Simulation Results

In this appendix, we analyze the finite-sample performance of the partial-information estimation procedure described in Section 4.3 in a Monte Carlo simulation study for the bivariate model described explicitly in Appendix A.2.

Our estimation procedure is designed for the semi-nonparametric specification, in which spot volatilities  $\xi_{i,s}$  are “freezed” to their values at time  $t$  for some short time interval. In other words, we approximate the stochastic volatilities by the processes  $v_{i,s} = \xi_{i,t}$  for  $s \in [t, T]$ . As has been discussed in Section 2.2, this approximation has negligible errors when pricing options

with short expiration time. However, in order to take this approximation into account in our Monte Carlo analysis, we simulate state vector series jointly with the stochastic volatility processes  $\xi_{i,t}$  from a fully parametric specification. In particular, we use the Heston (1993) volatility process:

$$d\xi_{i,t}^2 = \nu_i(\bar{\xi}_i^2 - \xi_{i,t}^2)dt + \sigma_{\xi,i}\xi_{i,t} \left( \rho_{\xi,i} dW_{i,t} + \sqrt{1 - \rho_{\xi,i}^2} dW_{i,t}^\xi \right), \quad (\text{C.13})$$

where the drift term allows for mean-reversion in the volatility process and  $W_{i,t}^\xi$  is a standard Brownian motion uncorrelated with the Brownian motions  $W_{j,t}^\xi$ , for  $j \neq i$ , and  $W_{i,t}$  in the corresponding index dynamics. Therefore, the Brownian component in (C.13) and in the corresponding index are correlated with constant coefficient  $\rho_{\xi,i}$ , which captures the leverage effect. Note that although  $W_{1,t}^\xi$  and  $W_{2,t}^\xi$  are independent, the Brownian part in one stochastic volatility process is not independent of the Brownian component in the other volatility due to the contemporaneous correlation between  $W_{1,t}$  and  $W_{2,t}$  in the index dynamics, which in turn we fix to  $\varrho = 0.6$  in our simulation study. Finally, when estimating the semi-nonparametric model, we use the true process  $v_{i,s} = \xi_{i,t}$  for  $s \in [t, T]$ .

We simulate the state vector series from the bivariate model specification coupled with the stochastic volatility processes (C.13) for each market using the Euler discretization technique with an additional truncation scheme for stochastic volatility. Then we price options using the characteristic function of the state vector including the stochastic volatility processes based on the COS method of Fang and Oosterlee (2008). For each sample, we simulate dynamics of 8 options per index, covering the most traded levels of moneyness (with strike-to-price ratios from 0.8 to 1.15) with a time to maturity of 0.1. Given the time discretization  $\Delta = 1/365$  between two time points, we simulate 1500 time observations. The stock indices and synchronized option panels are used as inputs for the estimation routine.

We note that the marginal characteristic functions of the log-prices and jump intensities have different oscillatory frequencies due to their different levels. In particular, the frequency of the marginal characteristic function for log-prices is much lower, which leads to only small changes in the CCF around the origin given the standard Gaussian choice of the probability

density function  $\pi(\tau)$ . This, in turn, leads to a potential loss of probabilistic information, which could deteriorate the parameter estimation. To overcome this issue, we re-scale the log-prices in the criterion function evaluation. That is, we use the CCF of  $c \cdot y_t$  with  $c > 0$  to construct the moment conditions. The parameters of the log-price dynamics are then also re-scaled accordingly. Based on preliminary simulation exercises, and aiming for a comparable magnitude in the levels and oscillatory frequencies of the states, we choose the scaling parameter to be  $c = 50$ .

Although the computational burden is significantly reduced when we employ the partial-information setting, the estimation routine is still computationally demanding: at every iteration, first, we have to back out the implied intensity by solving at every time point the non-linear least-squares problem (C.2) (which, in turn, involves numerical option pricing, and hence solving an ODE system), and next numerically evaluate four 2-dimensional integrals for the criterion function (C.5). Therefore, we run the Monte Carlo simulation with 100 replications, to obtain an (admittedly somewhat crude) indication of the finite-sample performance of the estimators.

The bivariate model specification involves 16 parameters we wish to identify, i.e., 8 parameters for each market. We recall that our multivariate option pricing model allows for, possibly time-varying, correlations between the stock indices (see Eqn. (1)). These correlations, however, do not appear in the pricing formulae of vanilla options, which depend only on the marginal CCF of a single index price (see e.g., Eqn. (A.8)). Hence, while these correlations are present in the model and the Monte Carlo simulations, they do not enter the moment conditions in our partial-information C-GMM estimation approach.

The simulation results are provided in Table C.1. We report the true parameter values used in the simulations and the corresponding Monte Carlo means, standard deviations and quantiles of the estimates. Overall, notwithstanding the challenging nature of the econometric problem, the results indicate a good finite-sample performance of our partial-information estimation procedure for the bivariate model. In particular, the self- and cross-excitation parameters, which are of central interest, are estimated with good precision. As is usual, estimates of the Brownian prices of risk,  $\eta_1$  and  $\eta_2$ , are less precise, due to the fact that their identification is

Table C.1: Simulation results for the bivariate model, semi-nonparametric approximation

	$\mu_1^{\mathbb{Q}_1}$	$\sigma_1$	$\kappa_1$	$\bar{\lambda}_1$	$\delta_{11}$	$\delta_{12}$	$\mu_1$	$\eta_1$
true	-0.130	0.030	6.000	1.000	3.000	1.000	-0.040	2.000
mean	-0.129	0.032	5.816	1.005	2.909	1.035	-0.038	1.966
std	0.010	0.008	0.464	0.193	0.335	0.186	0.007	1.972
25%	-0.133	0.027	5.520	0.924	2.685	0.925	-0.042	1.560
50%	-0.129	0.031	5.872	1.043	2.901	1.050	-0.038	2.467
75%	-0.125	0.034	6.116	1.086	3.070	1.131	-0.035	2.957
	$\mu_2^{\mathbb{Q}_2}$	$\sigma_2$	$\kappa_2$	$\bar{\lambda}_2$	$\delta_{22}$	$\delta_{21}$	$\mu_2$	$\eta_2$
true	-0.130	0.030	5.000	1.000	2.000	3.000	-0.040	2.000
mean	-0.128	0.030	4.895	1.071	2.010	3.052	-0.039	1.676
std	0.008	0.006	0.281	0.243	0.244	0.410	0.008	2.240
25%	-0.132	0.028	4.729	0.945	1.835	2.803	-0.043	1.333
50%	-0.127	0.030	4.925	1.083	2.002	3.074	-0.039	2.237
75%	-0.123	0.033	5.073	1.175	2.135	3.323	-0.036	2.667

This table provides Monte Carlo results for the bivariate model using the partial-information criterion function, and the semi-nonparametric approximation of Section 2.2. Each iteration consists of 1500 time points including simulated stock prices and 8 option prices for each time observation. True parameters and Monte Carlo sample means, standard deviations and 25%, 50%, 75% quantiles are presented on separate rows. The following parameters are used to simulate the stochastic volatility processes:  $\nu_1 = \nu_2 = 4.8$ ,  $\xi_1^2 = \xi_2^2 = 0.015$ ,  $\sigma_{\xi,1} = \sigma_{\xi,2} = 0.22$ ,  $\rho_{\xi,1} = \rho_{\xi,2} = -0.6$ .

based solely on the return dynamics.

To investigate the effect of “freezing” the spot volatilities on the behavior of the estimators, we have also conducted Monte Carlo simulations for the situation in which the true conditional characteristic function (based on the non-freezed stochastic volatility model) is used to back out the state variables and to obtain the moment conditions. For this comparison, we consider the ideal, but infeasible, situation in which we use the true values of the parameters characterizing the stochastic volatility process and the true spot volatilities, for estimation of the remaining parameters. The results are provided in Table C.2. We generally observe that the parameters are estimated with only a modest bias, comparable to the freezed volatility case in Table C.1. The estimation uncertainty is generally somewhat larger in Table C.2, suggesting that the net effect of freezing the volatility on the estimators’ behavior is actually beneficial.

The Monte Carlo simulations were also used to investigate the reliability of the asymptotic standard errors, as derived in Appendix C.3. In practice, these standard errors appear to be sensitive to the step size used in the calculation of numerical gradients. Therefore, we report, in

Table C.2: Simulation results for the bivariate model, fully parametric model

	$\mu_1^{\mathbb{Q}_1}$	$\sigma_1$	$\kappa_1$	$\bar{\lambda}_1$	$\delta_{11}$	$\delta_{12}$	$\mu_1$	$\eta_1$
true	-0.130	0.030	6.000	1.000	3.000	1.000	-0.040	2.000
mean	-0.131	0.028	5.722	0.996	2.968	1.078	-0.040	2.522
std	0.015	0.013	0.546	0.278	0.478	0.207	0.013	3.325
25%	-0.138	0.019	5.430	0.812	2.691	0.963	-0.048	0.343
50%	-0.129	0.028	5.714	0.984	3.027	1.088	-0.041	2.599
75%	-0.122	0.032	6.040	1.150	3.235	1.204	-0.034	5.065

	$\mu_2^{\mathbb{Q}_2}$	$\sigma_2$	$\kappa_2$	$\bar{\lambda}_2$	$\delta_{22}$	$\delta_{21}$	$\mu_2$	$\eta_2$
true	-0.130	0.030	5.000	1.000	2.000	3.000	-0.040	2.000
mean	-0.130	0.029	4.838	1.063	2.080	3.068	-0.039	1.759
std	0.014	0.009	0.513	0.443	0.312	0.544	0.011	3.364
25%	-0.135	0.024	4.578	0.782	1.853	2.837	-0.048	-0.868
50%	-0.127	0.028	4.907	1.070	2.103	3.129	-0.040	2.396
75%	-0.122	0.035	5.121	1.271	2.285	3.362	-0.033	4.324

This table provides Monte Carlo results for the bivariate model using the partial-information criterion function, and a fully parametric model with known stochastic volatility parameters and spot volatilities. Each iteration consists of 1500 time points including simulated stock prices and 8 option prices for each time observation. True parameters and Monte Carlo sample means, standard deviations and 25%, 50%, 75% quantiles are presented on separate rows. The following parameters are used to simulate the stochastic volatility processes:  $\nu_1 = \nu_2 = 4.8$ ,  $\xi_1^2 = \xi_2^2 = 0.015$ ,  $\sigma_{\xi,1} = \sigma_{\xi,2} = 0.22$ ,  $\rho_{\xi,1} = \rho_{\xi,2} = -0.6$ .

the empirical results in Section 5, bootstrap standard errors, the details of which are described in the next sub-section.

## C.5 Bootstrap Standard Errors

In this appendix, we provide technical details on the construction of bootstrap standard errors for our estimation method. We consider a parametric, recursive-design bootstrap procedure, see Bose (1988). This implies that bootstrap samples are generated largely analogously to the Monte Carlo simulations as described in Appendix C.4, with parameters  $\theta$  evaluated at their estimated values  $\hat{\theta}$ , and initializing the state vector at the empirically observed  $X_0 = (\log F_{1,0}, \dots, \log F_{m,0}, \lambda_{1,0}^{\hat{\theta}}, \dots, \lambda_{m,0}^{\hat{\theta}})'$ . The main difference with the Monte Carlo data generating process (DGP) is that the bootstrap DGP does not involve a parametric model for the volatilities  $\xi_{i,t}$ . Instead, we use the realized spot volatilities  $\hat{v}_{i,t}$ , which are kept fixed in repeated samples.

The algorithm therefore involves obtaining bootstrap replications  $\hat{\theta}^*$  of  $\hat{\theta}$  as follows:

1. Initialize the state vector at  $X_0^* = X_0$ ;
2. Obtain  $\{X_t^*\}_{t=1}^T$  and hence  $\{Y_t^*\}_{t=1}^T$  from the Euler discretization of the model, evaluating the parameters at  $\hat{\theta}$ , and replacing the stochastic volatilities by the realized spot volatilities  $\hat{v}_{i,t}$ ;
3. Obtain, for each  $t$ , a panel of option prices  $p_t^*$  using the COS method of Fang and Oosterlee (2008), based on the current state vector  $X_t^*$ , the parameter estimates  $\hat{\theta}$ , and assuming that the stochastic volatilities are freezed at their current spot volatility level  $\hat{v}_{i,t}$ ; use the same set of moneyness levels and maturities as in the original empirical data-set;
4. Use these option prices to imply, for each  $t$ , the jump intensities  $\lambda_{i,t}^{\theta*}$ , and apply the full-information or partial-information estimation method to yield  $\hat{\theta}^*$ .

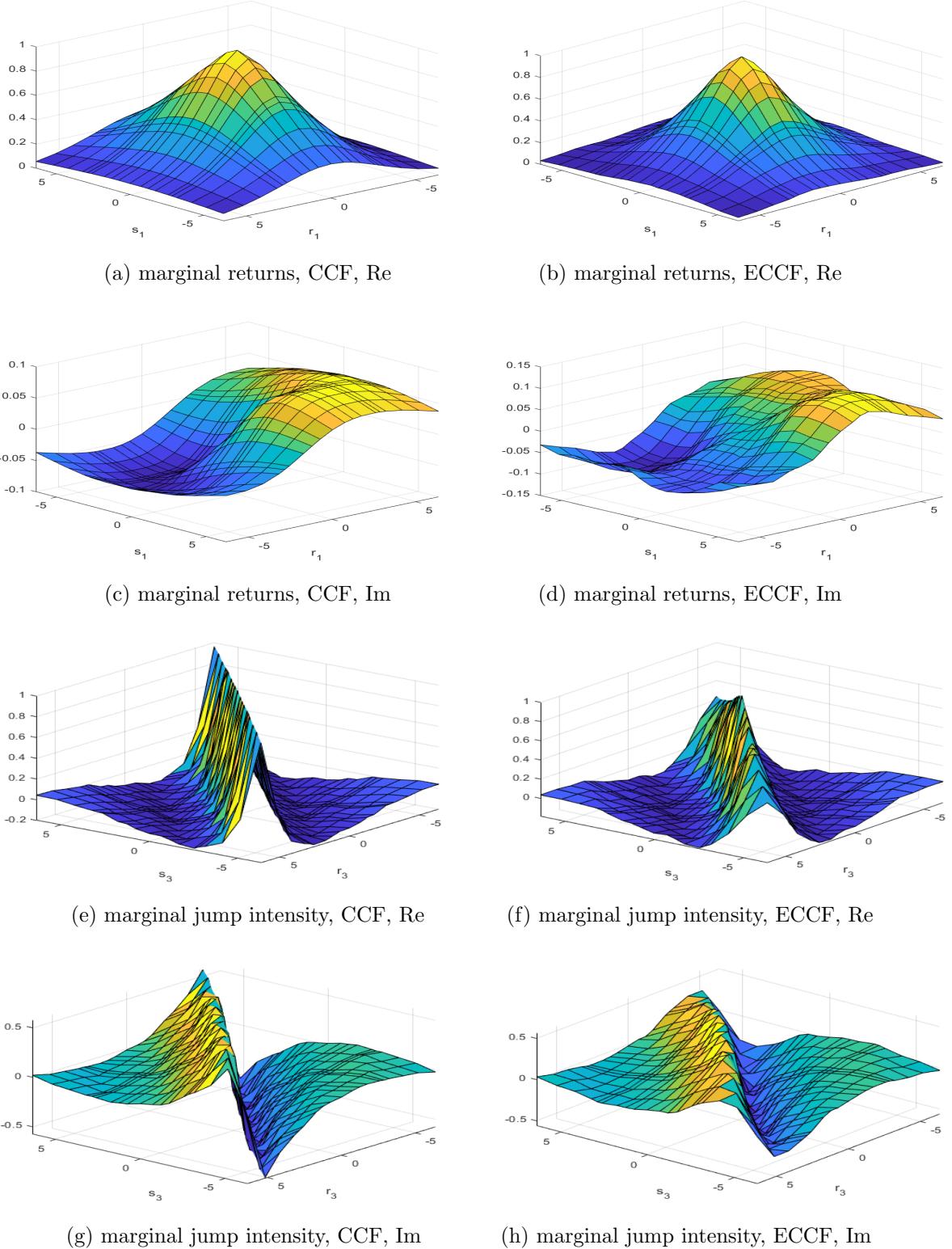
Repeating this procedure  $B$  times yields bootstrap replicates  $\hat{\theta}_b^*, b = 1, \dots, B$ . The bootstrap standard error of  $\hat{\theta}_i$  is then given by the sample standard deviation of  $\{(\hat{\theta}_{i,b}^* - \hat{\theta}_i)\}_{b=1}^B$ .

Validity of this bootstrap implementation, and hence consistency of the bootstrap standard errors, requires consistency of  $\hat{\theta}$ , together with the condition that the bootstrap DGP coincides with the actual process that has generated the original data (except for the values of the parameters). In practice, the latter condition will not be satisfied if the Euler discretization errors are non-negligible, and if the stochastic volatilities are not exogenous because of leverage. However, even in such cases we expect the bootstrap to provide a better characterization of estimation uncertainty than the asymptotic standard errors.

## Appendix D Data Analysis

In this appendix, we provide additional estimation results and discuss the details of three applications, to supplement the results in Section 5.

Figure D.1: C-GMM marginal moment conditions: Empirical fit



Note: This figure plots the real (Re) and imaginary (Im) parts of the marginal components of the moment conditions for the S&P returns and jump intensity, based on the parameter estimates of the pair S&P-FTSE. The marginal components of the moment conditions for the returns are the model-implied marginal CCF,  $\frac{1}{T} \sum_{t=1}^T m(r_1, Y_t) \phi(s_1, Y_t, \hat{\theta})$ , and the empirical CCF (ECCF),  $\frac{1}{T} \sum_{t=1}^T m(r_1, Y_t) e^{is_1 y_{1,t+1}}$ .

## D.1 Supplementary Estimation Results

To illustrate the empirical fit of the moment conditions that our model achieves within the partial-information implied-state C-GMM procedure, we plot in Figure D.1 the marginal components of the moments conditions. The marginal components of the moments functions are given by the model-implied marginal CCF and the empirical marginal CCF, both multiplied by the corresponding marginal instrument function. In particular, we consider the bivariate parameter estimates of the pair S&P-FTSE, and plot the (real and imaginary parts of the) marginal components of the moments conditions for the S&P index. We vary the arguments  $s_1$  and  $r_1$  of the moment functions, and fix the other arguments to zero, to obtain the marginal moments for the returns; and vary  $s_3$  and  $r_3$ , while fixing the other arguments to zero, to obtain the marginal moments for the jump intensity. The figure reveals that the model picks up well the shapes of the empirical CCFs.

To analyze the stability of the parameter estimates to changes in the sample, we have also estimated the model using only the first half of the sample period (January 1, 2006, to December 31, 2010). This part of the sample covers the Global Financial Crisis of 2008. The estimation results are displayed in Table D.1. We first note that identification of jump contagion from subsamples is expected to be weaker than from the full sample, as a consequence of the classical peso problem. Nevertheless, we observe from the table that the parameter estimates remain of a similar order of magnitude and various parameter estimates, such as those representing self-excitation, change only very little. There are also some noticeable differences. In particular, the jump contagion among the FTSE and S&P stock market indices is even more asymmetric than when estimated from the full sample: the parameter estimate for the cross-excitation from the US to the UK further increases, while the parameter estimate for the reverse effect (which was already limited) further decreases. Furthermore, the jump contagion effects among the S&P and DAX are unidirectional, from the US to Germany only, when estimated from the first half of the sample that excludes the Eurozone Debt Crisis of 2011–2012. This stands in contrast to the jump contagion effects among these two markets as estimated from the full sample, which are bidirectional and on equal footing.

To quantify the benefits of the bivariate model in an out-of-sample experiment, we have

Table D.1: Estimation results for S&P 500, FTSE 100 and DAX 30 based on the first half of the sample

	$\mu^Q$	$\sigma$	$\kappa$	$\bar{\lambda}$	$\delta^s$	$\delta^c$	$\mu$	$\eta$
S&P	-0.146	0.0792	2.702	0.244	2.697	0.000	-0.018	2.023
FTSE	-0.120	0.0178	1.725	0.251	1.724	4.865	-0.056	1.606
S&P	-0.118	0.1197	2.898	0.071	2.893	0.000	-0.042	1.776
DAX	-0.098	0.1026	2.672	0.067	2.663	2.278	-0.020	1.608

This table reports bivariate model parameter estimates for the S&P 500-FTSE 100 and S&P 500-DAX 30 pairs of stock market indices. The  $\delta^s$  parameters capture self-excitation for each index based on pairwise estimation (i.e.,  $\delta_i^s = \delta_{ii}$ ,  $i = 1, 2$ ), while the  $\delta^c$  parameters capture cross-excitation for each pair (i.e.,  $\delta_i^c = \delta_{ij}$ ,  $i, j = 1, 2$ ,  $i \neq j$ ). The subsample considered is January 1, 2006, to December 31, 2010.

Table D.2: Option prices: Out-of-sample empirical fit

$k$	0.85	0.87	0.89	0.91	0.93	0.95	0.97	0.99	1.01	1.03	1.05	1.07	1.09	Total
FTSE. <i>b</i>	1.41	1.16	0.97	1.01	1.21	1.12	0.65	1.11	1.91	2.22	2.17	2.10	2.20	1.54
FTSE. <i>u</i>	1.36	1.12	1.04	1.46	1.76	1.49	0.83	1.32	2.25	2.76	3.09	3.43	3.73	2.09
S&P. <i>b</i>	2.51	2.40	2.16	1.72	1.06	0.63	1.69	2.96	3.51	3.44	3.13	2.79	2.69	2.51
S&P. <i>u</i>	2.85	2.58	2.15	1.53	0.77	0.94	2.26	3.48	3.88	3.68	3.27	2.83	2.63	2.70

This table reports the root mean square errors (RMSEs, displayed as a percentage) of the option prices written on the FTSE 100 and S&P 500 indices, expressed in terms of the market-observed and model-implied Black-Scholes implied volatility, as a function of the strike-to-forward ratio  $k = K/F$ , using the bivariate (.*b*) and univariate (.*u*) models and parameter estimates. Option pricing errors are computed over the second half of the sample (January 1, 2011, to August 13, 2015), using parameter values estimated from the first half of the sample (January 1, 2006, to December 31, 2010).

computed option pricing errors for the univariate and bivariate models over the second half of the sample (January 1, 2011, to August 13, 2015), using parameter values estimated from only the first half of the sample (January 1, 2006, to December 31, 2010); the univariate parameter estimates for this subsample are in Table D.3. The results are reported in Table D.2; these results supplement the results in Table 4 in the main text. We observe from Table D.2 that the bivariate model continues to have (nearly uniformly) positive gains in fit compared to the univariate models.

Table D.3: Univariate model estimation results for FTSE 100, DAX 30 and S&P 500 based on the first half of the sample

	$\mu^Q$	$\sigma$	$\kappa$	$\bar{\lambda}$	$\delta$	$\mu$	$\eta$
FTSE	-0.137	0.0103	1.466	0.564	1.414	-0.038	1.945
DAX	-0.071	0.1202	1.632	0.084	1.607	-0.066	2.480
S&P	-0.163	0.0898	2.675	0.038	2.650	-0.066	1.610

This table reports parameter estimates for the univariate model for FTSE 100, DAX 30 and S&P 500 stock market indices. The subsample considered is January 1, 2006, to December 31, 2010.

## D.2 Supplementary Applications

### D.2.1 Distribution of Index Returns under a Sample Split

To supplement the results of the first application discussed in Section 5.2, we now provide results under a sample split. Specifically, we have re-computed the conditional log-return distributions and corresponding Value-at-Risk measures using parameter values estimated from the first half of the sample (January 1, 2006, to December 31, 2010). The results are reported in Table D.4; the required parameter values for the bivariate and univariate models, estimated from the first half of the sample, are provided in Tables D.1 and D.3. We observe moderate and reasonable changes when compared to the results in Table 6 in the main text, in line with the stability of the parameter estimates between the two (sub)samples (see also the previous subsection). Interestingly, as the jump contagion among the FTSE and S&P stock market indices is even more asymmetric when estimated from the first half of the sample only instead of from the full sample, this gets reflected in the 10-days Value-at-Risk capital requirements at the 99% probability level. Indeed, the effect of accounting for cross-excitation by the bivariate model, using stress scenarios similar to those in the Global Financial Crisis, would now imply a risk capital increase from about 4.9 (instead of 6.3) to 10.1 (instead of 8.4) cents for each dollar invested in the FTSE 100 index. This can be seen from panel (e), column (2) in Table D.4, as  $1 - \exp(-0.0503) \approx 4.9\%$  and  $1 - \exp(-0.1067) \approx 10.1\%$ .

Table D.4: Descriptive statistics for the conditional log-return distribution (simulated using model parameter estimates from the first half of the sample, horizon  $h = 10$  days)

	0.1%	1%	5%	25%	50%	75%	95%	S	K	$\mathbb{E}[N_t \lambda_0]$
(a) Base Case: $\lambda_{1,0} = \bar{\lambda}_1, \lambda_{2,0} = \bar{\lambda}_2$										
Bivariate - FTSE	-7.70	-3.37	-2.16	-0.81	0.09	0.97	2.25	-0.57	6.10	0.0069
Univariate - FTSE	-6.01	-3.39	-2.08	-0.72	0.18	1.08	2.37	-0.31	4.11	0.0142
Bivariate - S&P	-8.38	-3.09	-2.08	-0.79	0.11	0.99	2.29	-0.18	20.89	0.0061
Univariate - S&P	-4.34	-3.02	-2.13	-0.85	0.03	0.91	2.19	-0.77	15.77	0.0010
(b) Euro Debt Crisis: $\lambda_{1,0} = \lambda_{2,0} = 5$										
Bivariate - FTSE	-13.79	-8.09	-4.62	0.08	1.22	2.21	3.57	-1.97	9.23	0.1309
Univariate - FTSE	-8.13	-4.84	-2.40	0.29	1.37	2.34	3.65	-1.14	5.85	0.1205
Bivariate - S&P	-20.22	-10.97	-1.72	0.59	1.59	2.58	4.39	-1.16	19.91	0.1223
Univariate - S&P	-30.47	-17.73	-6.11	0.62	1.67	2.64	4.15	-3.48	24.13	0.1222
(c) S&P Shock: $\lambda_{1,0} = 20, \lambda_{2,0} = \bar{\lambda}_2$										
Bivariate - FTSE	-10.65	-5.87	-2.40	-0.69	0.30	1.28	2.73	-1.38	10.33	0.0371
Univariate - FTSE	-6.06	-3.43	-2.10	-0.73	0.17	1.08	2.36	-0.34	4.18	0.0147
Bivariate - S&P	-24.95	-13.82	-5.46	4.68	6.23	7.60	14.57	-0.72	8.80	0.4958
Univariate - S&P	-40.73	-24.25	-12.66	3.62	6.52	7.81	10.82	-1.91	8.82	0.4935
(d) FTSE Shock: $\lambda_{1,0} = \bar{\lambda}_1, \lambda_{2,0} = 20$										
Bivariate - FTSE	-17.85	-11.06	-5.68	0.36	4.38	5.80	7.31	-1.44	5.52	0.4923
Univariate - FTSE	-9.56	-4.79	-1.27	2.82	5.18	6.54	8.06	-1.19	4.94	0.4956
Bivariate - S&P	-10.22	-3.13	-2.09	-0.79	0.11	0.99	2.27	-0.84	25.04	0.0068
Univariate - S&P	-4.35	-3.05	-2.12	-0.85	0.03	0.91	2.20	-0.63	11.89	0.0011
(e) 2008 Global Financial Crisis: $\lambda_{1,0} = 20, \lambda_{2,0} = 15$										
Bivariate - FTSE	-17.32	-10.67	-5.40	0.39	3.55	4.87	6.50	-1.54	6.09	0.3979
Univariate - FTSE	-9.35	-5.03	-1.73	2.05	3.96	5.14	6.58	-1.25	5.30	0.3745
Bivariate - S&P	-24.67	-13.70	-5.24	4.69	6.25	7.61	14.48	-0.72	8.76	0.4926
Univariate - S&P	-39.94	-24.76	-12.56	3.46	6.52	7.82	10.88	-1.91	8.78	0.4961

This table displays the empirical quantiles (in percentages), skewness (S), kurtosis (K), and expected number of jumps implied by the conditional distribution of simulated log-returns for S&P 500 (“index 1”) and FTSE 100 (“index 2”). The stock index price paths are simulated using bivariate and univariate model parameter estimates from the first half of the sample (January 1, 2006, to December 31, 2010), conditional upon different values (“scenarios”) of the latent jump intensities. The return horizon is  $h = 10$  days. Volatilities are assumed to be constant throughout the horizon and are set to  $v_{i,s} = 8.36\%$  for both indices, and the instantaneous Brownian correlation is set to be 0.6.

### D.2.2 Two-Index Options

As a second application, we investigate the economic value of cross-excitation by pricing different types of multi-index options, the prices of which are typically sensitive to assumptions about dependence between the indices. As before, we restrict attention to the bivariate and univariate model estimates for the S&P 500 and FTSE 100 pair.

The following two-index option payoff types are considered:

- Correlation option:  $(K_1 - F_{1,T})^+ \cdot (K_2 - F_{2,T})^+$ ;
- Put option on the maximum between two indices:  $(K - \max\{F_{1,T}, F_{2,T}\})^+$ ;
- Basket option with fixed weights  $w_1$  and  $w_2$ :  $(K - (w_1 F_{2,T} + w_2 F_{1,T}))^+$ .

We focus on these put-type options with OTM strikes because they are sensitive to the joint occurrence of left tail events, i.e., to both indices substantially decreasing in value (this holds in particular for the first two payoff types). Among the various available option pricing approaches that have been proposed for pricing these types of multi-asset options, we opt for a Monte Carlo pricing approach using 100 000 simulations based on an Euler scheme, and we consider several initial jump intensity levels for illustration purposes. As we want to focus on the impact of cross-excitation, we make the additional simplifying assumption that these options are priced under a single arbitrage-free risk-neutral measure, disregarding any pricing contributions coming from foreign-exchange rate dynamics.

Two-index option price data points are provided in Table D.5, together with single-index vanilla European put prices for reference purposes. We first note that, given different parameter estimates for the bivariate and univariate models, we cannot isolate a “pure” cross-excitation effect. To this point, a larger (in absolute terms) jump size mean and standard deviation under the risk-neutral measure for the S&P series in the univariate model relative to the bivariate counterpart, results in more expensive European puts on the S&P index under the univariate specification than under the bivariate model for all scenarios, except scenario (d).

Nevertheless, we clearly observe the effect that jump cross-excitation has on the pricing of, in particular, correlation and put on max options. The prices of these options are markedly higher under the bivariate model with non-zero cross-excitation than under the univariate model. The strongest effects are found in scenario (c), where the initial jump intensity in the US, which is the leading economy in this pair, is substantially larger than in the UK. We also observe that under the bivariate model, the prices of single puts in asymmetric scenarios are larger due to exposure to shocks in the other market. The results for basket options, which are relatively less sensitive to joint left tail events, depend upon the chosen weights. For the weights  $w_1 = 0.3$  and  $w_2 = 0.7$  (as in Table D.5), we can see an effect due to the presence of

Table D.5: Two-index options

	Single Puts		Correlation		Put on max		Basket	
	S&P	FTSE	h=10	h=30	h=10	h=30	h=10	h=30
(a) $\lambda_{1,0} = \bar{\lambda}_1, \lambda_{2,0} = \bar{\lambda}_2$								
Bivariate	0.214	0.177	0.021	0.188	0.0018	0.0152	0.0042	0.0230
Univariate	0.235	0.182	0.001	0.044	0.0001	0.0050	0.0031	0.0171
(b) $\lambda_{1,0} = \lambda_{2,0} = 5$								
Bivariate	2.143	1.828	0.835	4.836	0.073	0.313	0.085	0.433
Univariate	2.406	1.699	0.727	4.330	0.061	0.289	0.074	0.382
(c) $\lambda_{1,0} = 20, \lambda_{2,0} = \bar{\lambda}_2$								
Bivariate	5.115	0.809	0.841	7.983	0.059	0.386	0.047	0.407
Univariate	5.716	0.179	0.168	1.045	0.012	0.056	0.026	0.207
(d) $\lambda_{1,0} = \bar{\lambda}_1, \lambda_{2,0} = 20$								
Bivariate	0.432	4.326	0.296	3.064	0.022	0.163	0.371	1.333
Univariate	0.237	4.237	0.160	1.109	0.011	0.063	0.350	1.286
(e) $\lambda_{1,0} = 20, \lambda_{2,0} = 15$								
Bivariate	5.129	3.828	4.521	21.709	0.319	0.945	0.362	1.435
Univariate	5.759	3.503	4.526	20.223	0.316	0.881	0.332	1.305

This table provides option prices for correlation, put on max, and basket options under five scenarios. For reference, single European put options are also priced for each index. Option prices are obtained using Monte Carlo simulations of the bivariate model for the pair S&P-FTSE and univariate models for the same indices. Initial prices are set to 100 for both indices. Correlation option strikes are set to  $K_1 = K_2 = 95$ ; put on max two-index strike is set to  $K = 95$ ; basket option weights used are  $w_1 = 0.3, w_2 = 0.7$  with strike set to  $K = 90$ . Two different maturities are priced:  $h = \{10, 30\}$  days. The single-index put option strike is set to  $K = 95$ . Volatilities are assumed to be constant throughout the horizon and are set to  $v_{i,s} = 8.36\%$  for both indices. The contemporaneous correlation between Brownian increments is set to 0.6.

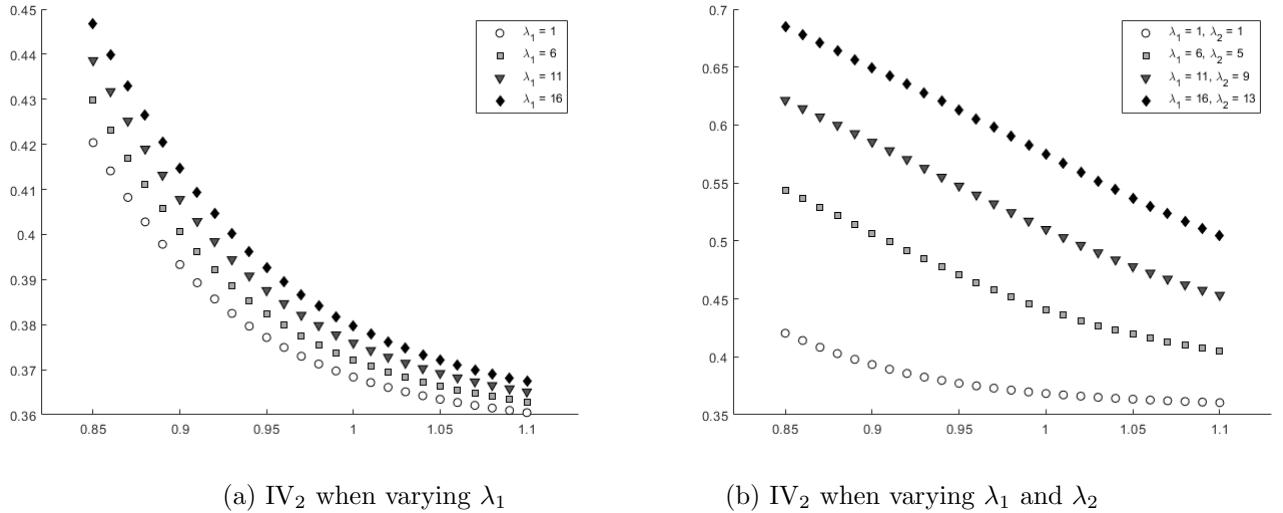
cross-excitation in the bivariate model, although it is less pronounced than for the other two two-index option payoff types.

### D.2.3 Comparative Statics of Implied Volatilities

Finally, with Figure 1 in mind, we are interested in the effect of cross-excitation on the dynamics of implied volatilities. To illustrate implied volatility dynamics, we conduct a comparative statics analysis and investigate how the implied volatility changes after the assumed occurrence of jumps in our multivariate option pricing model. This approach helps to exclude any other effects that potentially impact the implied volatility surface.

In particular, we consider again the parameter estimates from the bivariate model for the pair S&P-FTSE and mimic a scenario in which jumps occur in the US market. We fix the volatility levels in both markets to 35% and consider short-dated options with an expiration period of 15 days. Figure D.2(a) shows changes to implied volatility smiles coming from the

Figure D.2: Cross-excitation effects of jumps on implied volatilities



Note: This figure plots option implied volatilities (IV<sub>2</sub>) for the second index (i.e., FTSE) for different initial jump intensity levels. In Panel (a), the jump intensity level of the second index is fixed to  $\lambda_2 = 1$ , while Panel (b) shows the effect when both  $\lambda_1$  and  $\lambda_2$  vary. The spot volatilities are fixed to  $v_{i,s} = 35\%$  in both markets and the time-to-maturity is set to  $\tau = 15$  days.

different assumed initial levels of the intensity process  $\lambda_1$  (for fixed  $\lambda_2$ ). Although this only captures a marginal effect of jumps occurring in the S&P 500 index (since the jump intensity for the FTSE 100 index process is fixed), it illustrates that prices of options written on the second index are sensitive to the intensity of shocks in the first market. In particular, deep OTM options are more sensitive to the changes in  $\lambda_1$  than ITM counterparts. Furthermore, we observe changes in the slopes of the implied volatility curve.

Figure D.2(b) plots the implied volatilities when both intensity processes ( $\lambda_1$  and  $\lambda_2$ ) vary. This scenario mimics the occurrence of a jump in the US market: after a shock, the jump intensity  $\lambda_1$  increases with the value of the self-excitation parameter, and  $\lambda_2$  increases with the value of the cross-excitation parameter. For this analysis, we assume the self- and cross-excitation parameters to be 2.5 and 2, respectively, rounding the estimates for the S&P-FTSE pair in Table 3. Due to the simultaneous increase in  $\lambda_2$ , we observe more pronounced shifts in the implied volatility smile than in Figure D.2(a), corroborating once again the importance of jump contagion.

## References

- Aït-Sahalia, Y., & Jacod, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press.
- Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance*, 53(2), 499–547.
- Backus, D. K., Foresi, S., & Telmer, C. I. (2001). Affine term structure models and the forward premium anomaly. *The Journal of Finance*, 56(1), 279–304.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1), 1–37.
- Bollerslev, T., & Todorov, V. (2011). Estimation of jump tails. *Econometrica*, 79(6), 1727–1783.
- Bose, A. (1988). Edgeworth correction by bootstrap in autoregressions. *The Annals of Statistics*, 16(4), 1709–1722.
- Boswijk, H. P., Laeven, R. J., & Lalu, A. (2015). *Asset returns with self-exciting jumps: Option pricing and estimation with a continuum of moments* (Tech. Rep.). Amsterdam: University of Amsterdam and Tinbergen Institute.
- Brandt, M. W., & Santa-Clara, P. (2002). Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. *Journal of Financial Economics*, 63(2), 161–210.
- Broadie, M., Chernov, M., & Johannes, M. (2007). Model specification and risk premia: Evidence from futures options. *The Journal of Finance*, 62(3), 1453–1490.
- Carrasco, M., Chernov, M., Florens, J.-P., & Ghysels, E. (2007). Efficient estimation of general dynamic models with a continuum of moment conditions. *Journal of Econometrics*, 140(2), 529–573.
- Carrasco, M., & Florens, J.-P. (2000). Generalization of GMM to a continuum of moment conditions. *Econometric Theory*, 16(6), 797–834.
- De Marco, S., & Martini, C. (2009). Quasi-explicit calibration of Gatheral's SVI model. *Zeliade White Paper*, 1–15.
- Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6), 1343–1376.
- Fang, F., & Oosterlee, C. W. (2008). A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, 31(2), 826–848.
- Gatheral, J. (2011). *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2), 327–343.
- Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *Giornale dell'Istituto Italiano degli Attuari*, 64, 19–47.

- 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.
- Protter, P. E. (2005). *Stochastic Integration and Differential Equations* (2nd ed.). Springer.
- Yoshida, N. (2013). Martingale expansion in mixed normal limit. *Stochastic Processes and their Applications*, 123(3), 887–933.