TI 2021-086/III
Tinbergen Institute Discussion Paper

Jump Contagion among Stock Market Indices: Evidence from Option Markets

H. Peter Boswijk[1]
Roger J.A. Laeven[1]
Andrei Lalu[1]
Evgenii Vladimirov[1]

[1] University of Amsterdam

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at https://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

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

H. Peter Boswijk[1,2], Roger J. A. Laeven[1,3,†], Andrei Lalu[1,2], and Evgenii Vladimirov[1,2]

[1]Amsterdam School of Economics, University of Amsterdam, The Netherlands

[2]Tinbergen Institute, The Netherlands

[3]EURANDOM and CentER, The Netherlands

September 22, 2021

Abstract

This paper explores the contagious propagation of jumps among international stock market indices by exploiting a rich panel of stock and options data. We propose a multivariate option pricing model designed to allow for, but not superimpose, time and space amplification of jumps in option markets. We develop a semi-parametric estimation procedure employing a continuum of moments conditions in GMM with implied states. We introduce a partial-information approach to reduce the computational complexity arising in the multivariate setting, derive the asymptotic properties of our estimators, and analyze their finite-sample performance. Our empirical results reveal evidence of jump contagion in option markets, both from the US to Europe and vice versa, with the US leading the UK and standing on equal footing with Germany. We illustrate the importance of capturing jump contagion for risk management, option pricing, and scenario analysis.

Keywords: Jumps; Option markets; Crisis; Transmission; Spatio-temporal models; C-GMM.

JEL Classification: Primary: C58; G01; G15; Secondary: C14; C33; G13.

*We are very grateful to Yacine Aït-Sahalia, Christian Gouriéroux, Frank Kleibergen, Siem Jan Koopman, Olivier Scaillet and seminar and conference participants at Aarhus, Cass Business School, Lausanne, the St. Petersburg Economic Seminar, the University of Amsterdam, the Tinbergen Institute, the 12th World Congress of the Econometric Society, the 13th Annual SoFiE Conference, the 2021 International Association for Applied Econometrics (IAAE) Annual Conference, the Inaugural International Conference on Econometrics and Business Analytics, and the 2021 Econometric Society European Meeting (ESEM) for helpful comments and suggestions. `Matlab` code to implement the estimation procedure developed in this paper is available from the authors upon request. This research was funded in part by the Netherlands Organization for Scientific Research (NWO) under grants NWO-Vidi 2009 and NWO-Vici 2019/2020 (Laeven). Email addresses: `H.P.Boswijk@uva.nl`, `R.J.A.Laeven@uva.nl`, `A.Lalu@uva.nl`, and `E.Vladimirov@uva.nl`.

†Corresponding author. University of Amsterdam, Amsterdam School of Economics, PO Box 15867, 1001 NJ Amsterdam, The Netherlands. Phone: +31 6 52787501.

1 Introduction

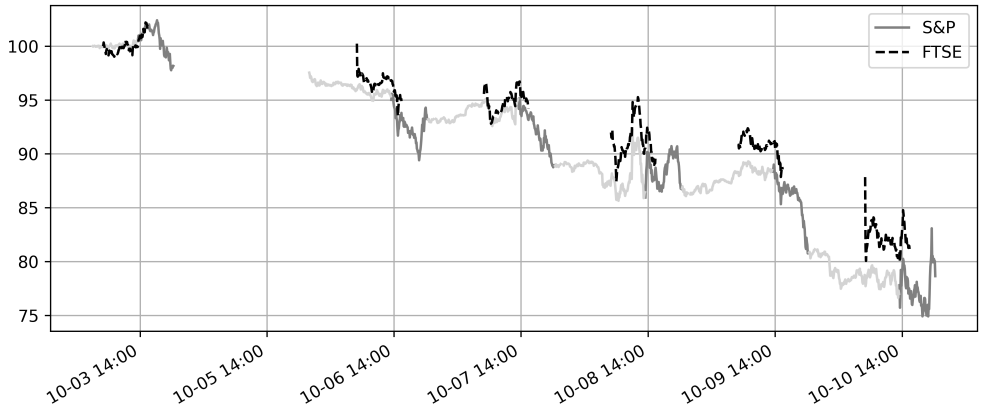
In our increasingly complex and interconnected world, intricate linkages exist between financial markets. Shocks to financial markets, e.g., due to financial or macro-economic news announcements, tend to propagate rapidly from one market to the next, potentially amplifying the initial shock via dynamic feedback loops. Such contagious amplification over time (i.e., over days or weeks) and in space (i.e., across markets), has important implications for risk management, valuation and hedging, and portfolio choice.

Option markets provide a unique laboratory to analyze time and space amplification. A panel of option price data embeds this phenomenon along several dimensions: in the time-series dimension, in the maturity dimension, in the moneyness dimension, and in the cross-sectional dimension. It thus contains a wealth of information on the persistence, direction, and contagious nature of the shocks. Figure 1 provides an example of such contagious propagation of shocks among the US (S&P 500) and UK (FTSE 100) options 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 corresponding to the shortest maturity options on the UK index catches up with the US implied volatility counterpart by October 8, and even outruns it in terms of its steepness by the end of what constitutes the worst week in US stock markets since 1929.

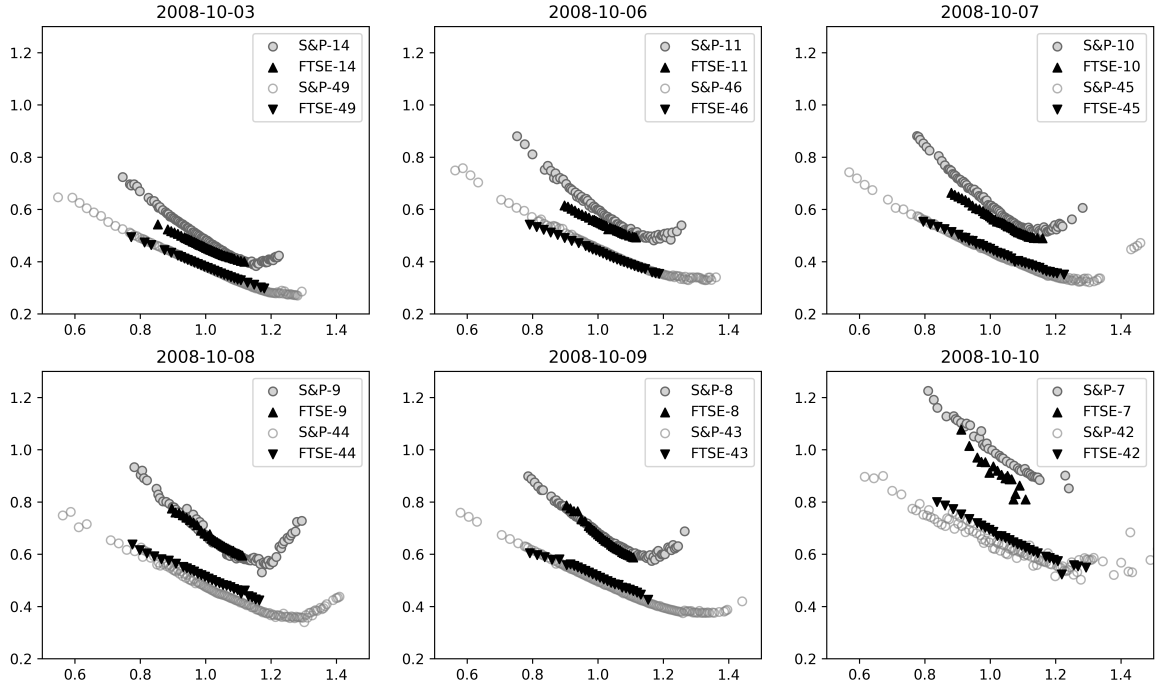
Extracting time and space amplification features from options panels, however, constitutes a challenging econometric problem. The challenges arise from the latency of the risk drivers—stochastic volatility and jump intensities—in option pricing, the multitude of dimensions—time-series, maturity, moneyness, and cross-sectional—that play a role, and the subtlety of the features—amplification of shocks in time and space, in particular, jump contagion—we wish to explore. In this paper, we develop an econometric approach that exploits a rich, carefully synchronized, panel of stock and options data to estimate a multivariate option pricing model designed to allow for, but not superimpose, time and space amplification of jumps.

Figure 1: An example of jump contagion among the US and UK option markets

(a) Underlying stock market indices



(b) Option-implied volatility slices



Note: This figure plots an example of jump contagion among the US and UK option markets during the sample period from 3 October, 2008, to 10 October, 2008. Panel (a) plots the E-Mini S&P 500 stock market futures index (S&P) and the FTSE 100 stock market index (FTSE). Both time series are normalized, with the first observation set to 100. The plots display intra-daily data with a frequency of 5 minutes. Trading times are converted into UTC (coordinated universal time). For the US futures index, the active trading periods (between 13:30 and 20:15 UTC) are highlighted in gray color, while the remaining trading times are displayed in light gray. October 4, 2008 (Saturday) has been omitted from the timeline. Panel (b) plots the 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., strike-to-price ratio). The two shortest time-to-maturity options are displayed for each of the indices. The times to maturity are indicated in the legends. All options data are collected in the interval between 14:03 and 14:05 UTC time. For more details about the data selection procedure, we refer to Section 4.

We formulate our model both under the physical probability measure, used for risk management, and under the risk-neutral probability measure, used for pricing and hedging. The risk-neutral specification enables us to infer the unobserved states of our model by implying the parameter-dependent latent state variables from the panel of option prices. Our estimation procedure is of a semi-parametric nature. We develop an implied-state GMM approach with a *continuum* of moments (C-GMM) to identify the parametric components of our multivariate semi-martingale model—the drift and jump components. We treat the spot volatility components non-parametrically, by equating them to jump-robust spot estimates obtained from high-frequency data over short periods of time,[1] to facilitate robust identification of our rich model.

An important hurdle in estimating our multivariate model is the computational complexity, which ramifies rapidly with dimension and renders full-blown econometric estimation computationally practically infeasible already in the bivariate setting. We therefore introduce a partial-information approach to C-GMM estimation, somewhat in the spirit of the limited-information estimation in Singleton (2001) for maximum likelihood and standard GMM. This allows us to considerably reduce the computational complexity while limiting the potential loss of asymptotic efficiency. We analyze the asymptotic properties of our partial-information implied-state C-GMM procedure, and derive expressions for asymptotically valid standard errors that take into account the effect of implied-state moments on estimation uncertainty. In Monte Carlo simulations we demonstrate that our criterion function embodies sufficient information to identify the model parameters, yielding a good finite-sample performance.

We estimate our model to a synchronized panel constructed from high-frequency data on UK (FTSE 100), German (DAX 30) and US (S&P 500) stock (futures) market indices and their option contracts, during the sample period from January 2006 to August 2015. The data selection and processing has been challenging and elaborate; it is described in detail in Section 4. Conventional wisdom suggests that the US plays a "leading role" in international financial markets, often summarized by the slogan "when the US sneezes, the world catches a

[1] In a univariate setting, Andersen, Fusari, and Todorov (2017) freeze the volatility of short-dated options to identify short-term market risks and show that such an approach has negligible errors for pricing options with short time-to-maturity.

cold".[2] Our empirical results confirm that in the jump contagion among the US and the UK, the US market indeed takes a leading role. However, our results reveal that in the interaction between the US and Germany, the two markets are essentially on equal footing. That is, the jump contagion from Germany to the US appears to be stronger than conventional wisdom suggests.[3] To our best knowledge, this paper is the first to estimate jump contagion among stock market indices using multivariate spot index data and option panels.

We finally illustrate 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. The strongest effects of jump contagion are found when the initial jump intensity in the US—the leading economy in this pair—is markedly larger than in the UK, representing a stress scenario that starts in the US.

Our work is related to the literature on international stock and option markets spillovers, which can be classified into three categories. First, there is a vast literature on asset return and volatility spillovers in an international context. For instance, Hamao, Masulis, and Ng (1990), Engle, Ito, and Lin (1990), Karolyi (1995), Koutmos and Booth (1995), Jebran, Chen, Ullah, and Mirza (2017), to name a few, use multivariate GARCH-type models to study the international transmission of stock returns and volatility. Diebold and Yilmaz (2009) and Ehrmann, Fratzscher, and Rigobon (2011) use a different approach based on a VAR framework and variance error decomposition; see also Alter and Beyer (2014) and Diebold and Yilmaz (2015) and the references therein.

A much smaller literature studies the international transmission of shocks in the form of jumps, or, contagion of rare adverse events. Aït-Sahalia, Cacho-Diaz, and Laeven (2015) and Aït-Sahalia, Laeven, and Pelizzon (2014) use mutually exciting jump processes to model jump contagion across stock markets as well as jump clustering in time. Furthermore, Jacod and Todorov (2009) develop a high-frequency test for common jumps and Dungey, Erdemlioglu, Matei, and Yang (2018) propose a high-frequency test for mutually exciting jumps in multi-

[2]This slogan is an adaptation of a historical (1799) phrase that is attributed to the Austrian diplomat Klemens von Metternich ("When France sneezes, Europe catches a cold.").

[3]The German stock market might be viewed as a proxy for continental Europe. In a sense, this finding is globally in line with Diebold and Yilmaz (2015), who studied trans-Atlantic equity *volatility* connectedness between major financial institutions. They argue that the connectedness was directed mainly from the US to Europe during 2007-2008, but became bidirectional starting in late 2008.

dimensional asset price processes.

Yet an even smaller literature—most closely related to our work—analyzes jump propagation through the lens of option markets. Andersen, Fusari, and Todorov (2020), when studying the pricing of tail risk index options across international equity markets, consider the pricing of index options separately for each of the markets, and find a large coherence across the markets with respect to their left tail factor. Bakshi, Carr, and Wu (2008) investigate a genuinely multi-dimensional stochastic discount factor in international economies using currency options. They show that investors price differently the global risk factor and the country-specific risk factor. 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. Finally, there is also a strand in the literature that considers derivative pricing with the multivariate Wishart stochastic volatility process (see, e.g., Gourieroux and Sufana (2010), Da Fonseca, Grasselli, and Tebaldi (2008)). This literature, however, does not allow for jumps.

Our work is also related to the econometric literature on implied-state GMM; see Pan (2002) who adapts the standard GMM setting to accommodate option-implied volatility. In contrast to Pan (2002), we develop an estimation procedure in the GMM setting with implied states using a *continuum* of moment conditions (C-GMM). The use of a continuum of moments was initiated in the interesting work of Carrasco and Florens (2000, 2002) and Carrasco, Chernov, Florens, and Ghysels (2007), and analyzed by Boswijk, Laeven, and Lalu (2015) with implied states to study self-excitation features of the S&P 500 stock market index using index options. C-GMM allows us to exploit more information than standard GMM with a finite number of moment conditions, and hence obtain more efficient estimators. Different from this existing literature, our approach is semi-parametric in nature. Furthermore, as the computational complexity of C-GMM is exponentially increasing with the dimension of the state vector, we introduce a partial-information approach to handle the multivariate C-GMM setting.

The remainder of this paper is organized as follows. Section 2 describes the multivariate dynamics of the stock index returns under both the physical and risk-neutral probability measures, a semi-nonparametric approximation adopted in our model, and the multivariate option

pricing approach. Section 3 develops the estimation procedure. In Section 4, we describe the data and in Section 5 we present our empirical analysis. Conclusions are in Section 6. Some further details concerning the change of measure, asymptotic properties, jump-robust volatility estimation, and data handling are provided in four appendices.

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 Aït-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 country-level pricing kernel specifications that jointly accommodate arbitrage-free exchange rate dynamics and stock price dynamics with mutually exciting jumps.

2.1 Index Return Dynamics

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:

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

for $i = 1, \ldots, m$, 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 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 in the next sub-section.[4]

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. More specifically, we define 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

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

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

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

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 $\overline{\lambda}_i > 0$ at a rate $\kappa_i > 0$. The parameters δ_{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 country increases the probability of jumps in other countries, which entails jump propagation in space. Note that these time and space amplification features are probabilistic and not superimposed, i.e., not certain to occur. The paired vectors (N, λ) jointly constitute a Markov process.

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 in the CAPM: η_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. We discuss the jump risk premium in more detail in the next sub-section, where we introduce risk-neutral dynamics. The last term in (1) is the

[4] Throughout, stochastic processes, expectation operators, and parameters without superscript are understood to be defined with respect to the physical probability measure $\mathbb{P}$.

compensator for the jump component; the compensated jump component is a local martingale.

As is common in the literature, we assume the relative jump sizes $J_{i,t},\ i = 1, \ldots, m$ to be independent log-normal random variables. More 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 $\mathbb{E}[J_{i,t}] = \exp(\mu_i + \frac{1}{2}\sigma_i^2) - 1$. 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. The model for the log-equity dynamics together with the jump intensity processes for all markets can be summarized as:

$$\begin{cases} d\log S_{i,t} = \left(r_{i,t} - q_{i,t} + (\eta_i - \frac{1}{2})\xi_{i,t}^2 - \mathbb{E}^{\mathbb{Q}_i}[J_i]\lambda_{i,t}\right)dt + \xi_{i,t}dW_{i,t} + Z_{i,t}dN_{i,t}, \\ \qquad d\lambda_{i,t} = \kappa_i(\overline{\lambda}_i - \lambda_{i,t})dt + \sum_{j=1}^{m} \delta_{ij} dN_{j,t}, \quad J_{i,t} = e^{Z_{i,t}} - 1, \quad Z_{i,t} \sim \mathcal{N}(\mu_i, \sigma_i^2), \end{cases} \tag{4}$$

for $i = 1, \ldots, m$. Importantly, this model admits a generalized affine jump-diffusion representation as defined in Appendix B of Duffie, Pan, and Singleton (2000) under both the physical probability measure $\mathbb{P}$ and the market-specific equivalent martingale measures $\mathbb{Q}_i$, as introduced in the following sub-section.

2.2 Pricing Kernels and Dynamics under $\mathbb{Q}_i$

The sources of uncertainty stemming from the random jumps in the model in Eqn. (4) 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.

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. Suitably extending the univariate specification used in Pan (2002), we assume that the candidate pricing kernel $M_{i,t}$ has the following

dynamics:

$$\frac{dM_{i,t}}{M_{i,t}} = -r_{i,t}dt - \eta_i \xi_{i,t} dW_{i,t} + \sum_{k=1}^{m} \left(U^i_{k,t} dN_{k,t} - \mathbb{E}[U^i_{k,t}] \lambda_{k,t} dt \right), \tag{5}$$

where $U^i_{k,t}$ are random jump sizes in market k, specific to pricing kernel i. Hence, 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 $U^i_{k,t}$ in the pricing kernels to follow the same type of distribution as the index jump sizes, i.e., $U^i_{k,t} = e^{V^i_{k,t}} - 1$ are independently log-normally distributed with $V^{i,t}_k \sim \mathcal{N}(a_{i,k}, b^2_{i,k})$. Note that $U^i_{k,t}$ are allowed to be different from $U^j_{k,t}$ 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 $U^i_{k,t}$ is independent of $U^j_{n,t}$ for $i \neq j$ and/or $k \neq n$, and independent of all Brownian motions, but the kernel jump log-sizes $V^i_{k,t}$ are possibly correlated with the index jump log-sizes $Z_{k,t}$, with correlation coefficients $\rho_{i,k}$.

Similar to Pan (2002), we assume the mean relative jump size in the pricing kernel $M_{i,t}$ to be zero, i.e., $a_{i,k} + \frac{1}{2} b^2_{i,k} = 0$ for $k = 1, \ldots, m$. These constraints enable identification of the jump parameters and also set the jump-timing risk premium to zero. As a result, they keep the dynamics of the jump intensity processes the same under both probability measures, i.e., $\lambda^{\mathbb{Q}_i}_{k,t} \equiv \lambda_{k,t}$ for $k = 1, \ldots, m$. In a more general setting, one could allow for different intensity processes under the physical and risk-neutral measures using an additional non-zero component in (5), but this would considerably increase the number of parameters to estimate and consequently weaken parameter identification. We provide further details on the measure change in Appendix A, formally establishing in particular that the pricing kernels thus specified rule out arbitrage opportunities within each market as well as internationally, and do not affect the jump intensity dynamics.

For the purpose of option pricing we define the density processes $\psi_{i,t}$ associated with the pricing kernel (5):

$$\psi_{i,t} = M_{i,t} \exp\left(\int_0^t r_{i,s} ds \right), \qquad i = 1, \ldots, m. \tag{6}$$

These are local martingales (which can be shown by applying Itô's formula), and hence each density process $\psi_{i,t}$ uniquely defines an equivalent martingale measure $\mathbb{Q}_i$ in market i. Applying Girsanov's theorem using the density process $\psi_{i,t}$, the index i follows, under $\mathbb{Q}_i$,[5]

$$\frac{dS_{i,t}^{\mathbb{Q}_i}}{S_{i,t}^{\mathbb{Q}_i}} = (r_{i,t} - q_{i,t})dt + \xi_{i,t}dW_{i,t}^{\mathbb{Q}_i} + J_{i,t}dN_{i,t} - \mathbb{E}^{\mathbb{Q}_i}[J_i]\lambda_{i,t}dt, \tag{7}$$

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

$$W_{i,t}^{\mathbb{Q}_i} = W_{i,t} + \int_0^t \eta_i \xi_{i,s} ds. \tag{8}$$

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 μ_i and $\mu_i^{\mathbb{Q}_i}$. The jump risk premium 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 defined by (6), the model for log-index dynamics is given by

$$\begin{cases} d\log S_{i,t}^{\mathbb{Q}_i} = \left(r_{i,t} - q_{i,t} - \frac{1}{2}\xi_{i,t}^2 - \mathbb{E}^{\mathbb{Q}_i}[J_i]\lambda_{i,t}\right)dt + \xi_{i,t}dW_{i,t}^{\mathbb{Q}_i} + Z_{i,t}dN_{i,t}, \\ \quad d\lambda_{i,t} = \kappa_i(\overline{\lambda}_i - \lambda_{i,t})dt + \sum_{j=1}^m \delta_{ij}dN_{j,t}, \quad J_{i,t} = e^{Z_{i,t}} - 1, \quad Z_{i,t} \overset{\mathbb{Q}_i}{\sim} \mathcal{N}(\mu_i^{\mathbb{Q}_i}, \sigma_i^2), \end{cases} \tag{9}$$

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

[5] In the sequel, only random objects that, upon the measure change, differ path-wise (i.e., as functions of $\omega \in \Omega$) are indexed by a superscript $\mathbb{Q}_i$.

as under the physical measure.

2.3 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, e.g., a Heston-type stochastic volatility specification. Alternatively, following Andersen et al. (2017), 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. That is, given close-to-maturity options in the empirical application and consistent estimates for the spot volatility, inference based on the semi-nonparametric model is more robust.

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

$$\begin{cases} \dfrac{d\tilde{S}^{\mathbb{Q}_i}_{i,s}}{\tilde{S}^{\mathbb{Q}_i}_{i,s}} = (r_{i,t} - q_{i,t})ds + v_{i,s}dW^{\mathbb{Q}_i}_{i,s} + J_{i,s}dN_{i,s} - \mathbb{E}^{\mathbb{Q}_i}[J_i]\lambda_{i,s}ds, & s \in [t, T], \\ \quad v_{i,s} = \xi_{i,t}\mathbb{1}_{\{t \le s \le T\}}, \qquad \tilde{S}^{\mathbb{Q}_i}_{i,t} = S_{i,t}. \end{cases} \tag{10}$$

In other words, the spot volatility is $v_{i,s}$ 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 $\tilde{S}^{\mathbb{Q}_i}_{i,s}$ is initialized at the true index price $S_{i,t}$ at time t. We also refer to 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 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 $\tilde{S}_{i,t}$ instead of $S_{i,t}$ leads to negligible approximation errors. 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 $\mathbb{P}$:

$$\begin{cases} \dfrac{d\tilde{S}_{i,s}}{\tilde{S}_{i,s}} = (r_{i,t} - q_{i,t} + \eta_i \xi_{i,t}^2)ds + v_{i,s}dW_{i,s} + J_{i,s}dN_{i,s} - \mathbb{E}^{\mathbb{Q}_i}[J_i]\lambda_{i,s}ds, \qquad s \in [t,T], \\ \quad v_{i,s} = \xi_{i,t}\mathbb{1}_{\{t \le s \le T\}}, \qquad \tilde{S}_{i,t} = S_{i,t}, \end{cases} \tag{11}$$

for $i = 1, \ldots, m$. The specifications (10)–(11) will serve as a basis for the estimation procedure developed in the next section, but with $T = t + 1$; i.e., the spot volatility is assumed constant over the period of a single day.

To obtain estimates of the spot volatility values $v_{i,t}$, we use a jump-robust spot volatility estimator based on high-frequency returns observed before time t. Appendix C 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.

2.4 The Bivariate Specification and Conditional Characteristic Function

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

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

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

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 (9), following the discussion in Section 2.2 and Appendix A, and are semi-nonparametrically approximated following Section 2.3.

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 et al. (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.[6]

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 1 below for a general result):

$$\phi^{\mathbb{Q}_1}(u_1, X_t, T-t; v_t, \theta) = e^{\alpha(T-t)+\beta_1(T-t)\log F_{1,t}+\beta_3(T-t)\lambda_{1,t}+\beta_4(T-t)\lambda_{2,t}}, \tag{13}$$

where $u_1 \in \mathbb{R}$ is the argument of the CCF, θ is the vector of parameters, and $\alpha(T-t)$ and

[6] In our empirical analysis, we use the COS method for this purpose, proposed by Fang and Oosterlee (2008), to price European options.

$\beta(T-t)$ are the solutions to the following system of ODEs:

$$\begin{cases} \dot{\beta}_1(s) = 0, \\ \dot{\beta}_3(s) = -(\exp(\mu_1^{\mathbb{Q}} + \frac{1}{2}\sigma_1^2) - 1)\beta_1 - \kappa_1\beta_3 + \exp(\mu_1^{\mathbb{Q}}\beta_1 + \frac{1}{2}\sigma_1^2\beta_1^2 + \delta_{11}\beta_3 + \delta_{21}\beta_4) - 1, \\ \dot{\beta}_4(s) = -\kappa_2\beta_4 + \exp(\delta_{12}\beta_3 + \delta_{22}\beta_4) - 1, \\ \dot{\alpha}(s) = -\frac{1}{2}v_{1,t}^2\beta_1 + \kappa_1\bar{\lambda}_1\beta_3 + \kappa_2\bar{\lambda}_2\beta_4 + \frac{1}{2}v_{1,t}^2\beta_1^2, \end{cases} \tag{14}$$

$0 \leq s \leq T-t$, with initial conditions $\beta_1(0) = \mathrm{i}u_1, \beta_3(0) = 0, \beta_4(0) = 0$ and $\alpha(0) = 0$; for notational convenience, the time dependence in $\beta(s)$ has been omitted from the right-hand side expressions in (14). 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 (14) is not possible due to the non-linear components involved in the ODE for $\beta_3(s)$ and $\beta_4(s)$. 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 full bivariate specification involves 16 parameters, i.e., 8 parameters for each market. Although highly non-linear and complex, the option pricing relation 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. The next section discusses this in detail.

3 Estimation Procedure

In this section, we develop the estimation procedure we follow, exploiting rich synchronized datasets of stock market indices and corresponding option panels. The jump-robust spot volatility estimator, which we use to semi-nonparametrically approximate index dynamics, is described in Appendix C. 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 base our estimation procedure on GMM with a *continuum* of moment conditions (C-GMM), originally developed in Carrasco and Florens (2000) and Carrasco and Florens (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 efficient and reliable estimates. Unfortunately, the computational burden of C-GMM is exponentially increasing with the dimension of the state vector. Therefore, in Section 3.3, we introduce a partial-information version of multivariate implied-state C-GMM applied to the bivariate model, and establish its asymptotic properties.

Throughout this section, we assume that for each of the markets $i = 1, \ldots, m$, and at regular-interval observation times $t = 0, 1, \ldots, T$, 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 $\hat{v}_{i,t}$.

3.1 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 θ. Let us define the option pricing relation determining a stacked vector of option prices $p_t = (p'_{1,t}, \ldots, p'_{m,t})'$, with $\tau_t = (\tau'_{1,t}, \ldots, \tau'_{m,t})'$ the corresponding time to maturity and $k_t = (k'_{1,t}, \ldots, k'_{m,t})'$ the moneyness level, given the global state vector X_t, model parameters θ, and volatility estimates $\hat{v}_t$, as follows:

$$p_t = \mathcal{P}(F_t, \lambda_t, \theta, \hat{v}_t, \tau_t, k_t), \tag{15}$$

with $\mathcal{P} : \mathbb{R}_+^m \times \Lambda \times \Theta \times \mathbb{R}_+^m \times \mathbb{R}_+^{mn_\tau} \times \mathbb{R}_+^{mn_k} \to \mathbb{R}_+^{mn_\tau n_k}$. Here $\Lambda \subseteq \mathbb{R}_+^m$ is the domain of the jump intensities and Θ is a compact parameter space, such that the stationarity condition of the multivariate Hawkes process is satisfied.[7] 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_τ and n_k represent the number of different maturities and moneyness levels, respectively.

We exploit the option-pricing relation (15) to imply the latent jump intensities. In fact, under the assumption of correct model specification and if the true model parameters θ_0 and volatility processes ξ_t were known, we could recover the true jump intensities from this equation using the market observed forward prices F_t and option prices p_t. However, for any other parameter vector $\theta \in \Theta$, we can also back out a parameter-dependent proxy λ_t^θ for the unobserved jump intensities λ_t, by solving (15) for λ_t^θ.

Formally, let us define the domain of invertibility of the option-pricing relation $\Sigma \subset \mathbb{R}_+^{mn_\tau n_k} \times \Theta \times \mathbb{R}_+^m \times \mathbb{R}_+^{mn_\tau} \times \mathbb{R}_+^{mn_k}$, such that it is a maximal set for which a mapping $f : \Sigma \to \Lambda$ is uniquely defined by

$$p_t = \mathcal{P}(F_t, f(p_t, F_t, \theta, \hat{v}_t, \tau_t, k_t), \theta, \hat{v}_t, \tau_t, k_t). \tag{16}$$

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

$$\lambda_t^\theta = f(p_t, F_t, \theta, \hat{v}_t, \tau_t, k_t), \tag{17}$$

where we use the superscript θ 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 λ_t is retrieved based on the market-observables when evaluating the mapping (17) at the true model parameters θ_0 and using the true volatility process ξ_t (still assuming correct model specification).

In practice, the latent jump intensities are backed out by minimizing the difference between

[7] This requires the spectral radius of the matrix consisting of the entries (δ_{ij}/κ_i), $i, j = 1, 2$, to be less than unity.

the market-observed and model-implied option prices due to (16) 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:

$$\lambda_t^\theta = \arg\min_{\lambda} \sum_{i=1}^{m} \sum_{j=1}^{n_{i,t}} \Big(\mathcal{IV}(X_{i,t}, \theta, \hat{v}_{i,t}, \tau_{i,t}^j, k_{i,t}^j) - \mathcal{BSIV}(F_{i,t}, \tau_{i,t}^j, k_{i,t}^j) \Big)^2, \tag{18}$$

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 θ, jointly for all m economies. The model prices 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 (18) at every time point, independently of previous or later points, i.e., we do not impose or exploit 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.

It is worth emphasizing again that the backing-out procedure can only be applied to short-dated options, due to the use of spot volatility estimates in (17) and the corresponding approximated index return dynamics; see Section 2.3. For longer-dated options, the resulting approximation errors would start to play a larger role, and therefore, one has to take into account the dynamics of the volatility process over time. Treating the volatility process as constant has a minor effect on short-dated options, as the volatility is expected to mean-revert at a slow rate.

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

3.2 Parameter Estimation in a Full-Information Setting

As discussed in the previous section, the model admits a generalized affine jump-diffusion representation under the physical and risk-neutral probability measures. One of the great advantages of the class of affine jump-diffusions is that the CCF of the state vector X_T conditional on information available at time t is known in a closed form (up to the solution of an ODE system) as an exponentially affine function of X_t.

This CCF 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. Alternatively, 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 a C-GMM estimator extended to allow for implied state variables.

Because C-GMM requires a stationary Markovian state, we consider a state process $Y_t = (y_{1,t}, \ldots, y_{m,t}, \lambda_{1,t}, \ldots, \lambda_{m,t})'$, 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}:

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

with Δ 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

$$u(s, Y_t, Y_{t+1}; \hat{v}_t, \theta) := e^{\mathrm{i}s \cdot Y_{t+1}} - \phi(s, Y_t, \Delta; \hat{v}_t, \theta) \tag{19}$$

with an instrument function $m(r, Y_t)$, to obtain the moment function

$$\begin{aligned} h(r, s, Y_t, Y_{t+1}; \hat{v}_t, \theta) &:= m(r, Y_t) \cdot u(s, Y_t, Y_{t+1}; \hat{v}_t, \theta) \\ &= m(r, Y_t) \left(e^{\mathrm{i}s \cdot Y_{t+1}} - \phi(s, Y_t, \Delta; \hat{v}_t, \theta) \right), \end{aligned} \tag{20}$$

and hence the moment conditions

$$\mathbb{E}[h(r, s, Y_t, Y_{t+1}; \hat{v}_t, \theta_0)] = 0, \qquad \text{for all } r, s \in \mathbb{R}^{2m}. \tag{21}$$

The idea of GMM with a continuum of moments, developed in Carrasco and Florens (2000), Carrasco and Florens (2002) and Carrasco et al. (2007), is to use not a discrete finite set of vectors s as arguments for the moment conditions (21), but rather to employ a continuum of values of s. Furthermore, these authors show that also using a continuum of instruments of the form $m(r, Y_t) = e^{\mathrm{i}r \cdot Y_t}$ 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 (15) 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 (21) based on the state vector with implied intensities $Y_t^\theta = (y_{1,t}, \ldots, y_{m,t}, \lambda_{1,t}^\theta, \ldots, \lambda_{m,t}^\theta)'$.

Let us denote the sample analogue of the moment conditions (21) based on the state vector with implied intensities as

$$h_T(\tau; \hat{v}, \theta) := \frac{1}{T-1} \sum_{t=1}^{T-1} h(\tau, Y_t^\theta, Y_{t+1}^\theta; \hat{v}_t, \theta), \tag{22}$$

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

$$\langle f, g \rangle = \int f(\tau)\overline{g(\tau)}\pi(\tau)d\tau, \tag{23}$$

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 τ is given by

$$Q_T(\theta) = \|h_T(\tau; \hat{v}, \theta)\|^2 = \int h_T(\tau; \hat{v}, \theta)\overline{h_T(\tau; \hat{v}, \theta)}\pi(\tau)d\tau. \tag{24}$$

This is similar to minimizing the Euclidean norm in the case of GMM with a finite number of moments. Carrasco and Florens (2000) and Carrasco et al. (2007) prove the consistency of the GMM estimator based on the integral norm under some regularity conditions and Boswijk et al. (2015) show the consistency results when the implied state is employed in the estimation of univariate models.

Note that minimizing (24) defines a first-step C-GMM estimator; a second step, leading to full efficiency, was further developed in Carrasco et al. (2007). However, the use of the implied state variable Y_t^θ 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 formally establish efficiency gains from a second step, we focus on first-step implied-state C-GMM estimation. We will establish the corresponding asymptotic distributional properties, which account for the effect of implied-state moments on estimation uncertainty.

In practice the criterion function (24) has to be evaluated numerically using quadrature methods. Carrasco et al. (2007) show that introducing optimal instruments of the form $m(r, Y_t) = e^{\mathrm{i}r \cdot Y_t}$ does not increase the computational complexity, because all elements associated with the index r have an analytical form. Therefore, the numerical integration of (24) is

of the same dimension as the state vector.[8]

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 (24) for the bivariate model, which we detail in the following sub-section.

3.3 Parameter Estimation in a Partial-Information Setting

Singleton (2001) notes that although full 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 loss in asymptotic efficiency is traded off against the computational simplicity relative to the full ML-CCF approach.

Fortunately, a similar idea 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 (24), we exploit a partial-information estimator based on the sum of m 2-dimensional integrals. Although this approach can be described in the general multivariate setting, here we apply it directly to the bivariate model introduced in Section 2.4.

We start by providing the closed-form CCF for the bivariate model under $\mathbb{P}$.

Proposition 1 *The conditional characteristic function of the state vector* $Y_t = (y_{1,t}, y_{2,t}, \lambda_{1,t},$

[8] In other words, the instruments can be integrated out from the criterion function using a property of Fourier transforms. See the appendix in Carrasco et al. (2007) or the online appendix in Boswijk et al. (2015).

$\lambda_{2,t})'$ *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:*

$$\begin{cases} \dot{\beta}_1(u) = 0 \\ \dot{\beta}_2(u) = 0 \\ \dot{\beta}_3(u) = -(\exp(\mu_1^{\mathbb{Q}_1} + \frac{1}{2}\sigma_1^2) - 1)\beta_1 - \kappa_1\beta_3 + \exp(\mu_1\beta_1 + \frac{1}{2}\sigma_1^2\beta_1^2 + \delta_{11}\beta_3 + \delta_{21}\beta_4) - 1 \\ \dot{\beta}_4(u) = -(\exp(\mu_2^{\mathbb{Q}_2} + \frac{1}{2}\sigma_2^2) - 1)\beta_2 - \kappa_2\beta_4 + \exp(\mu_2\beta_2 + \frac{1}{2}\sigma_2^2\beta_2^2 + \delta_{12}\beta_3 + \delta_{22}\beta_4) - 1 \\ \dot{\alpha}(u) = (\eta_1 - \frac{1}{2})\hat{v}_{1,t}^2\beta_1 + \frac{1}{2}\hat{v}_{1,t}^2\beta_1^2 + \kappa_1\bar{\lambda}_1\beta_3 + (\eta_2 - \frac{1}{2})\hat{v}_{2,t}^2\beta_2 + \frac{1}{2}\hat{v}_{2,t}^2\beta_2^2 + \kappa_2\bar{\lambda}_2\beta_4 \\ \qquad + \varrho_t\hat{v}_{1t}\hat{v}_{2t}\beta_1\beta_2 \end{cases} \tag{25}$$

with initial conditions $\beta(0) = \mathrm{i}s$ *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 $\beta_1(u) = \mathrm{i}s_1$ and $\beta_2(u) = \mathrm{i}s_2$, 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.

Let us further 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, respectively, and $Y_t^{(3)} = (y_{1,t}, \lambda_{2,t})$ 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,

$$\phi^{(i)}(v, Y_t, \Delta; \hat{v}_t, \theta) := \phi\left(s^{(i)}, Y_t, \Delta; \hat{v}_t, \theta\right) = e^{\alpha^{(i)}(\Delta)+\beta_3^{(i)}(\Delta)\lambda_{1,t}+\beta_4^{(i)}(\Delta)\lambda_{2,t}}, \tag{26}$$

where $\alpha^{(i)}(\Delta), \beta_3^{(i)}(\Delta), \beta_4^{(i)}(\Delta)$ are the solutions to the ODE system (25) 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 described in (21) we can consider four sets of "marginal" moment conditions stacked in a vector form, that is:

$$\mathbb{E}[\mathbf{h}(\tau, t; \theta_0)] = \mathbb{E}\left[\begin{pmatrix} h^{(1)}(\tau; \hat{v}_t, \theta_0) \\ h^{(2)}(\tau; \hat{v}_t, \theta_0) \\ h^{(3)}(\tau; \hat{v}_t, \theta_0) \\ h^{(4)}(\tau; \hat{v}_t, \theta_0) \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{27}$$

with

$$h^{(i)}(\tau; \hat{v}_t, \theta) = m(r, Y_t^{(i)}) \left(e^{\mathrm{i}s \cdot Y_t^{(i)}} - \phi^{(i)}(s, Y_{t+1}; \hat{v}_t, \theta) \right),$$

where $i = 1, 2, 3, 4$, $\tau = (r, s)'$ and $r, s \in \mathbb{R}^2$.

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

$$Q_T(\theta) = \int \mathbf{h_T}(\tau; \hat{v}, \theta)' \overline{\mathbf{h_T}(\tau; \hat{v}, \theta)} \pi(\tau) d\tau = \sum_{i=1}^{4} \int h_T^{(i)}(\tau; \hat{v}, \theta) \overline{h_T^{(i)}(\tau; \hat{v}, \theta)} \pi(\tau) d\tau, \tag{28}$$

where by $h_T^{(i)}(\tau; \hat{v}, \theta)$ we denote the sample counterpart of the i^{th} component of the moment condition (27).

Intuitively, we decompose the joint criterion function (24) into the sum of the four criterion functions based on the marginals.[9] As in the LML-CCF suggested by Singleton (2001), there is some potential loss of efficiency with this approach. However, we significantly gain in terms of computational power as the estimation problem in (28) requires four numerical evaluations

[9] Alternatively, one could see the criterion function (28) as a specific choice of the probability density function $\pi(\tau)$ in (24). 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 (24) with this choice of density functions results in the decomposed criterion function (28).

of 2-dimensional integrals, rather than the numerical integration over 4-dimensional space as in (24).

Furthermore, there is sufficient information in the decomposed criterion function (28) that allows us to identify all parameters in the bivariate model. Indeed, if the cross-excitation parameters δ_{12} and δ_{21} were zero, and the Brownian motions $(W_{1,t}, W_{2,t})$ were uncorrelated, then we could split the estimation 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. Moreover, having the cross-market marginals $Y_t^{(3)}$ and $Y_t^{(4)}$ provides additional identification power if the cross-excitation parameters are non-zero. Importantly, 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 (28). We support this discussion with the simulation analysis detailed in the next sub-section.

Overall, unlike a regular GMM estimation procedure, GMM with a continuum of moments allows reaching asymptotic efficiency as in ML estimation. However, the computational burden of C-GMM exponentially increases with the increase of dimensionality in the state vector. Therefore, we consider the C-GMM setting based on partial information. Moreover, the use of a continuum of moments allows us to effectively exploit all information from the marginal CCFs, which in turn potentially embodies a substantial amount of information about the full state vector. Therefore, the loss in asymptotic efficiency in the partial-information setting relative to the full C-GMM criterion function is likely to be small compared to the gains in terms of computational simplicity.

The full estimation procedure is then as follows. First, we calculate the jump-robust volatility estimates from high-frequency equity index prices. Then, starting with an initial parameter set θ, we back-out the latent jump intensity processes λ_t^θ from the joint dynamics of option prices and equity indices. Next, we evaluate the C-GMM criterion function, which is based on the state vector with implied intensities as if the full state vector was observable. After updating the parameter set θ, we repeat the last two stages, iterating until the criterion

function is minimized.

We discuss in detail the asymptotic properties of the estimation procedure and the calculation of the standard errors in Appendix B (see, in particular, Proposition 2), and we provide Monte Carlo simulation results in the next sub-section.

3.4 Simulation Results

We analyze the finite-sample performance of the partial-information estimation procedure described in the previous sub-section in a Monte Carlo simulation study for the bivariate model described in Section 2.4.

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 2.3, 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:

$$d\xi_{i,t}^2 = \nu_i(\overline{\xi_i^2} - \xi_{i,t}^2)dt + \sigma_{\xi,i}\xi_{i,t}\left(\rho_{\xi,i}dW_{i,t} + \sqrt{1-\rho_{\xi,i}^2}dW_{i,t}^{\xi}\right), \tag{29}$$

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 (29) 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 (29) 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 (18) (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 (28). 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 simulation results are provided in Table 1. We report the true parameter values

Table 1: Simulation results for the bivariate model

	$\mu_1^{\mathbb{Q}_1}$	σ_1	κ_1	$\bar{\lambda}_1$	δ_{11}	δ_{12}	μ_1	η_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
	$\mu_2^{\mathbb{Q}_2}$	σ_2	κ_2	$\bar{\lambda}_2$	δ_{22}	δ_{21}	μ_2	η_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. 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%, 90% quantiles are presented on separate rows. The following parameters are used to simulate the stochastic volatility processes: $\nu_1 = \nu_2 = 4.8$, $\overline{\xi_1^2} = \overline{\xi_2^2} = 0.015$, $\sigma_{\xi,1} = \sigma_{\xi,2} = 0.22$, $\rho_{\xi,1} = \rho_{\xi,2} = -0.6$.

used in the simulations and the corresponding Monte Carlo means, standard deviations and quantiles of the estimates. Overall, 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, η_1 and η_2, are less precise, due to the fact that their identification is based solely on the return dynamics.

The Monte Carlo simulations were also used to investigate the reliability of the asymptotic standard errors, as derived in Appendix B. In practice, these standard errors appear to be sensitive to the step size used in the calculation of numerical gradients. Therefore, we report, in the empirical results in Section 5, standard errors based on a step size chosen such that the Monte Carlo standard deviations were in line with the (average) asymptotic standard errors in the simulations.

4 Data

The data collection process for this paper has been challenging and elaborate. 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 1 January 2006 to 13 August 2015. In particular, we exploit a very large sample of intra-day tick-by-tick observations in the underlying data-set to obtain 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 use the rich panel of daily data to estimate bivariate as well as univariate versions of the model. Although procedures required for obtaining such homogeneous panels of daily data are perhaps not typically discussed in a dedicated section, we find that the sampling of tick-by-tick data underlying the option panel observations is a non-trivial and potentially novel exercise, which warrants some additional description. Therefore, the remainder of this section is concerned with the data collection process.

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 which denote the time at which the price data was received by the Thomson Reuters from the exchange's servers. As the use of official exchange-determined "close" prices would not be possible because the options are traded in different time zones, we use the synchronization procedure outlined in the next paragraphs.

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*. For example, 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 driven by daylight saving time adjustments used in the US and in Europe.

We impose several rules and filters in the data selection routine such that the collected data are: (i) similar to the end-of-day option price determination done by trading venues, (ii) as close as possible to the reference point to achieve synchronicity between observations from different markets, and (iii) provide reliable information. For specific details we refer the interested reader to Appendix D which further outlines the selection procedure. Table 2 provides the descriptive statistics for the filtered option sample for each of the three markets.

In addition to the filtered option data, we use short-term interbank lending interest rates for each relevant currency,[10] which we interpolate to match option time-to-maturity. Circumventing the need to specify and calibrate dividend yield dynamics, we follow Aït-Sahalia and Lo (1998) and back out forward prices from put-call parity pairs and estimate our model using log-forward returns. The details on forward price calculations are also provided in Appendix D.

Further data processing was carried out to construct a (homogeneous) panel of Black-Scholes implied volatilities for a dense grid of option maturities and strike prices. 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.[11] Second, it reduces computational costs as obtaining model-implied option prices for a fixed set of maturities is computationally less-demanding.

To construct the homogeneous option panel, the sample implied volatility points for each

[10] We use LIBOR-US, LIBOR-GBP and EURIBOR short-term interest rates for options on S&P 500 futures, and the FTSE 100 and DAX 30 indices, respectively.

[11] 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 from ATM options. Using a fixed moneyness and maturity grid therefore improves the likelihood that information about tail events is extracted from options.

Table 2: 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: Number of option contracts						
$0.75 < k \leq 0.85$	1 375	2 968	9 505	16 000	23 413	33 499
$0.85 < k \leq 0.92$	6 646	8 914	18 139	21 831	36 986	41 122
$0.92 < k \leq 0.98$	18 743	16 670	20 637	20 843	40 197	40 236
$0.98 < k \leq 1.03$	23 127	18 285	17 300	17 332	34 234	34 250
$1.03 < k \leq 1.10$	10 911	13 052	17 838	21 861	28 560	37 994
$1.10 < k \leq 1.20$	1 258	2 721	4 764	9 875	6 782	15 024
Total	62 060	62 610	88 183	107 742	170 172	202 125
Panel B: Sample mean of implied volatility (%)						
$0.75 < k \leq 0.85$	38.7	35.3	36.6	33.1	39.3	32.7
$0.85 < k \leq 0.92$	31.2	27.2	30.4	27.2	29.0	25.5
$0.92 < k \leq 0.98$	22.0	20.8	24.1	23.1	21.8	20.8
$0.98 < k \leq 1.03$	17.2	16.9	20.0	20.1	16.7	17.0
$1.03 < k \leq 1.10$	18.8	17.1	19.2	18.2	16.7	15.4
$1.10 < k \leq 1.20$	31.0	24.6	26.0	21.1	26.7	20.2
Total	21.2	20.6	25.1	23.8	24.1	22.0
Panel C: Sample standard deviation of implied volatility (%)						
$0.75 < k \leq 0.85$	9.8	10.6	7.0	7.4	10.6	8.6
$0.85 < k \leq 0.92$	10.9	8.9	8.4	7.3	9.0	7.7
$0.92 < k \leq 0.98$	8.7	7.6	7.9	7.1	8.5	7.6
$0.98 < k \leq 1.03$	7.9	6.9	7.9	7.0	8.5	7.6
$1.03 < k \leq 1.10$	9.2	7.6	7.8	6.7	8.8	7.6
$1.10 < k \leq 1.20$	11.6	9.9	9.8	7.9	11.4	8.8
Total	10.3	9.3	9.9	8.7	11.9	9.8

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 1 January 2006 to 13 August 2015. The filters employed in the data selection procedure are detailed in Appendix D. Observations are bucketed into two categories for time-to-maturity, τ, and into six categories with respect to the moneyness level, defined as strike-to-forward ratio $k = K/F$.

index option were interpolated over a fixed set of moneyness and option maturities.[12] Having experimented with different techniques (among which cubic splines, polynomial fit and kernel smoothing), 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

[12]Interpolating implied volatilities is a common procedure. See, for instance, Broadie, Chernov, and Johannes (2007) and Bardgett, Gourier, and Leippold (2019). For each sample observation point we have first interpolated volatility slices in the moneyness dimension and then linearly in total variance to a fixed expiration time τ, which we set equal to 40 days for all three markets. For the estimation procedure, we sample from the resulting interpolated volatility fit up to 13 option implied volatilities with fixed (relative) moneyness levels evenly spaced between 85% and 109%. SVI interpolation procedure details are provided in Appendix D.

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 et al. (2007). Its dynamics and parametrization are not in any way related to our model specification.

For the non-parametric spot volatility estimates, which constitute another input in our estimation procedure, we exploit high-frequency index return data. More specifically, to obtain spot volatility estimates, we use one-minute index return series[13] 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 2.3.

5 Empirical Analysis

In this section, we first describe our estimation results for the three pairs of stock market indices and next we consider several applications drawing from these estimates. To gauge the effect of jump contagion featured in the bivariate model specification, we also report in Section 5.1 estimation results for the (nested) univariate model specification where the cross-excitation channel is absent. The applications in Section 5.2 further highlight the statistical and economic importance of the jump cross-excitation effect.

5.1 Model Estimation Results

Parameter estimates for the bivariate models are provided in Table 3. Each bivariate model is estimated using the partial-information implied-state C-GMM procedure developed in Section 3.3. For each index pair, we use synchronized daily data for the corresponding stock market indices and their options panels, following Section 4, where the synchronicity between

[13]Non-parametric volatility estimates based on one-minute returns are not systemically higher than the ones based on two- and five-minutes returns. Also, one-minute return series exhibit insignificant positive autocorrelation. All of these features suggest that microstructure noise is not strongly present in the series of returns.

Table 3: Bivariate model estimation results for FTSE 100, DAX 30 and S&P 500

	$\mu^{\mathbb{Q}}$	σ	κ	$\bar{\lambda}$	δ^s	δ^c	μ	η
FTSE	-0.126	0.020	4.063	0.353	1.638	2.506	-0.038	2.186
	(0.003)	(0.015)	(0.308)	(0.008)	(0.087)	(0.063)	(0.012)	(4.652)
DAX	-0.131	0.027	3.482	0.418	2.190	1.244	-0.025	2.680
	(0.002)	(0.011)	(0.14)	(0.039)	(0.125)	(0.059)	(0.016)	(3.459)
S&P	-0.148	0.036	3.320	0.296	2.501	0.517	-0.039	2.173
	(0.002)	(0.003)	(0.013)	(0.002)	(0.008)	(0.004)	(0.032)	(4.028)
FTSE	-0.131	0.033	3.216	0.283	1.709	2.119	-0.041	2.051
	(0.001)	(0.004)	(0.016)	(0.002)	(0.011)	(0.006)	(0.024)	(4.731)
S&P	-0.135	0.036	3.781	0.261	2.257	1.788	-0.037	1.977
	(0.004)	(0.012)	(0.214)	(0.032)	(0.142)	(0.122)	(0.032)	(4.049)
DAX	-0.138	0.039	4.235	0.394	2.287	1.658	-0.035	2.119
	(0.004)	(0.015)	(0.212)	(0.038)	(0.107)	(0.062)	(0.032)	(3.21)

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 δ^s parameters capture self-excitation for each index based on pairwise estimation (i.e., $\delta^s_i = \delta_{ii}$, $i = 1, 2$), while the δ^c parameters capture cross-excitation for each pair (i.e., $\delta^c_i = \delta_{ij}$, $i, j = 1, 2$, $i \neq j$). Standard errors are reported in parentheses.

markets is crucial for the identification of jump contagion in space.

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 equal to δ^s which ranges from 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.[14] Estimates of the cross-excitation parameter δ^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 cross-excitation effect from Germany (perhaps as a proxy for continental Europe) to the US is stronger than suggested by conventional wisdom. The cross-excitation effect from DAX to FTSE has the largest cross-

[14]The self-excitation of jumps is broadly in line with the findings in Boswijk et al. (2015) and Du and Luo (2019), who studied univariate self-excitation models with parametric volatility dynamics in the US market index using weekly data.

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. To the best of our knowledge, this is the first study that documents the presence of jump contagion across major stock market indices using spot index time-series data jointly with option data panels.

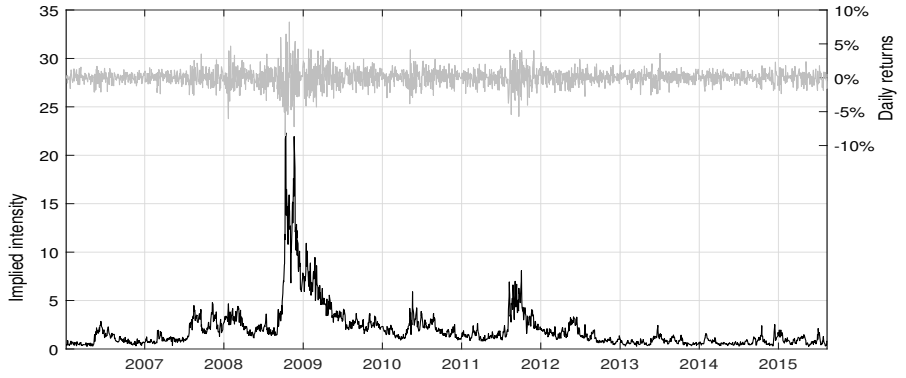
Using the model parameter estimates in Table 3, we imply the latent jump intensities for each index from the corresponding sets of option prices. Figure 2 shows plots of the implied jump intensities, along with the index log-forward returns for reference purposes. To back out the jump intensity for the 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. It may be emphasized once more that by treating the diffusive volatility non-parametrically, the estimation results, and hence the fitted jump intensities, are robust to volatility misspecification.

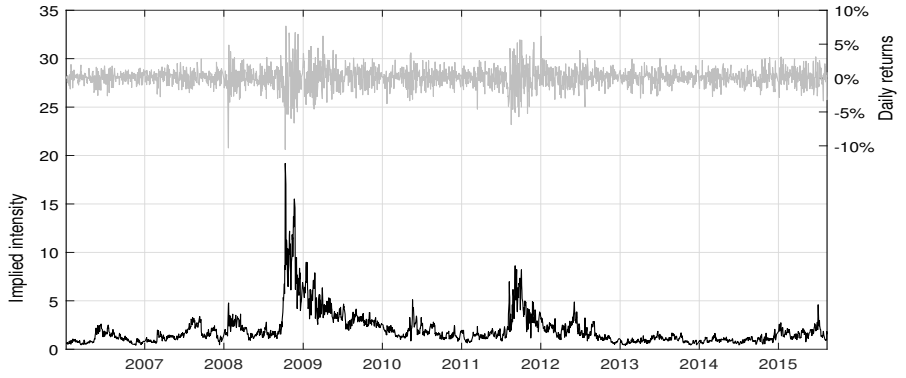
In our model set-up, the jump risk premia are driven by the difference in means between the jump sizes under the physical and risk-neutral probability measures, i.e., they are specified as $(\mathbb{E}[J] - \mathbb{E}^{\mathbb{Q}}[J])\lambda_t$ per unit of time. The estimated jump risk premium coefficients, $\mathbb{E}[J] - \mathbb{E}^{\mathbb{Q}}[J]$, 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 gauge the effect of cross-excitation in the jump components across markets, we also provide estimation results for the univariate model specification. The univariate model can be

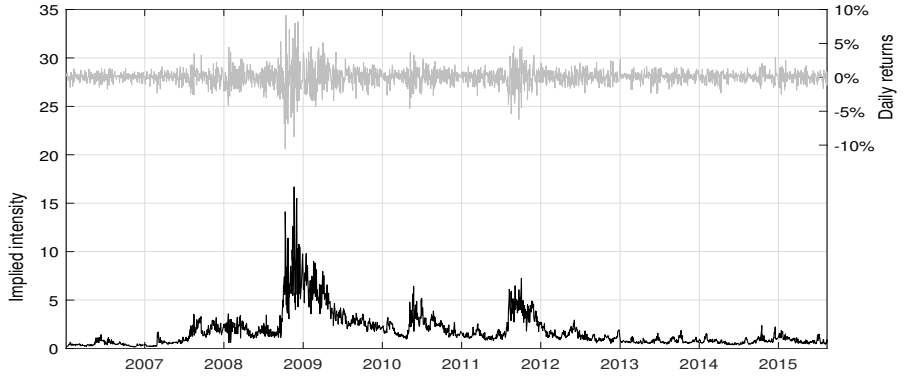
Figure 2: Time-series of the option-implied jump intensities



(a) FTSE 100



(b) DAX 30



(c) S&P 500

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.

Table 4: Univariate model estimation results for FTSE 100, DAX 30 and S&P 500

	$\mu^{\mathbb{Q}}$	σ	κ	$\bar{\lambda}$	δ	μ	η
FTSE	-0.127	0.030	2.132	0.318	1.798	-0.030	2.379
	(0.001)	(0.005)	(0.005)	(0.002)	(0.003)	(0.023)	(4.716)
DAX	-0.137	0.032	3.207	0.486	2.132	-0.029	2.109
	(0.003)	(0.013)	(0.086)	(0.012)	(0.08)	(0.026)	(3.436)
S&P	-0.161	0.043	2.445	0.305	2.176	-0.038	2.216
	(0.002)	(0.008)	(0.02)	(0.002)	(0.025)	(0.057)	(4.342)

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

seen as a nested version of the bivariate specification, where 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 3.2. The estimation results of the univariate models for the FTSE 100, DAX 30 and S&P 500 stock market indices are provided in Table 4, and will be used for gauging purposes in the next subsection.

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 Model Applications

We illustrate the statistical and economic implications of jump contagion in three applications.

5.2.1 Distribution of index returns

First, 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 4, 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 model estimates.

Since the simulated distribution of log-returns is conditional on the (initial) jump intensity

Table 5: 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	-7.47	-3.17	-2.10	-0.79	0.11	1.00	2.27	-0.52	6.27	0.0073
Univariate - FTSE	-6.34	-3.14	-2.10	-0.77	0.12	1.00	2.28	-0.28	4.46	0.0079
Bivariate - S&P	-7.90	-3.14	-2.08	-0.77	0.12	1.01	2.27	-0.59	7.17	0.0073
Univariate - S&P	-8.34	-3.13	-2.09	-0.76	0.13	1.02	2.30	-0.72	9.00	0.0079
	(b) Euro Debt Crisis: $\lambda_{1,0} = \lambda_{2,0} = 5$									
Bivariate - FTSE	-12.98	-7.69	-3.09	0.32	1.36	2.33	3.68	-2.11	11.34	0.1249
Univariate - FTSE	-10.80	-6.26	-2.18	0.31	1.33	2.29	3.64	-1.72	10.03	0.1238
Bivariate - S&P	-13.29	-7.69	-2.71	0.50	1.53	2.49	3.84	-2.23	12.80	0.1242
Univariate - S&P	-14.27	-8.41	-2.50	0.67	1.68	2.64	4.02	-2.38	14.18	0.1206
	(c) S&P Shock: $\lambda_{1,0} = 20, \lambda_{2,0} = \bar{\lambda}_2$									
Bivariate - FTSE	-9.56	-3.76	-2.07	-0.68	0.23	1.14	2.45	-1.00	9.04	0.0186
Univariate - FTSE	-6.66	-3.12	-2.10	-0.78	0.11	1.00	2.28	-0.39	5.85	0.0081
Bivariate - S&P	-15.88	-8.54	-2.99	3.79	5.89	7.10	8.76	-1.75	7.40	0.4872
Univariate - S&P	-16.63	-8.84	-3.01	4.54	6.57	7.78	9.65	-1.74	7.87	0.4872
	(d) FTSE Shock: $\lambda_{1,0} = \bar{\lambda}_1, \lambda_{2,0} = 20$									
Bivariate - FTSE	-16.00	-9.14	-3.77	2.78	5.13	6.35	7.92	-1.73	7.08	0.4835
Univariate - FTSE	-12.29	-6.37	-2.10	3.33	5.15	6.34	7.97	-1.54	6.87	0.4924
Bivariate - S&P	-8.59	-3.22	-2.06	-0.74	0.15	1.05	2.33	-0.76	8.15	0.0104
Univariate - S&P	-7.83	-3.14	-2.07	-0.76	0.13	1.01	2.31	-0.58	8.19	0.0074
	(e) 2008 Global Financial Crisis: $\lambda_{1,0} = 20, \lambda_{2,0} = 15$									
Bivariate - FTSE	-15.36	-8.89	-4.03	2.28	4.00	5.15	6.68	-1.83	7.72	0.3756
Univariate - FTSE	-11.83	-6.61	-2.55	2.45	3.89	4.99	6.53	-1.66	7.61	0.3693
Bivariate - S&P	-15.88	-8.54	-2.99	3.81	5.92	7.13	8.79	-1.75	7.38	0.4893
Univariate - S&P	-16.63	-8.85	-3.01	4.54	6.57	7.78	9.65	-1.74	7.87	0.4870

This table displays the empirical quantiles, 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 correlation between Brownian increments is set to be 0.6.

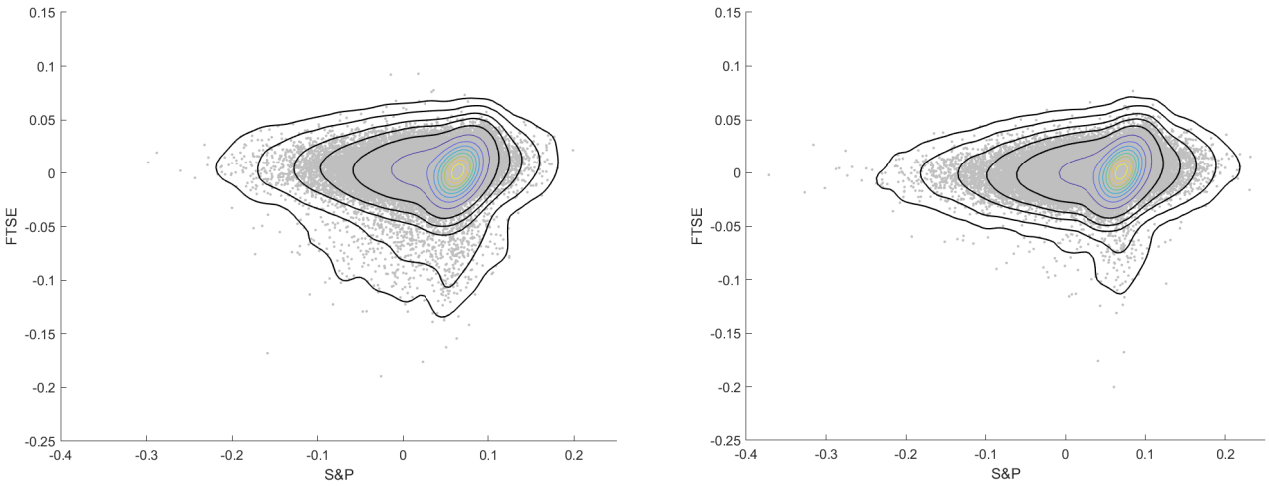
values, we consider five different scenarios to illustrate the effect of jump contagion. Under the base scenario (a), the initial values of the intensities are given by the corresponding estimates of the base rates $\bar{\lambda}_1$ and $\bar{\lambda}_2$, while in scenarios (b)–(e) we assume the initial values to be similar to levels implied from our model during the 2008 Global Financial Crisis and the Euro Debt Crisis. Table 5 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, while this is generally not the case for the S&P series, with the exception of scenario (d). A natural explanation for this is that in the bivariate model the spillover of jumps from the S&P 500 index to FTSE 100 is much more pronounced than vice versa, while the jump size parameters imply on average more negative jump sizes under the univariate specification than under the bivariate model. Scenario (d) assumes a large asymmetry in the level of intensities, with the intensity for S&P set to the base rate, showing that although the cross-excitation from FTSE to S&P is four times smaller than the reverse cross-excitation, its effect becomes important in this scenario. Wider distributions imply larger values of standard risk measures used for risk capital calculations such as Value-at-Risk (VaR) and Expected Shortfall (ES).

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 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.

Next to Table 5, we provide the contour plots for scenario (c)—S&P Shock—in Figure 3.

Figure 3: Contour plots for scenario (c), S&P Shock, with $\lambda_{1,0} = 20, \lambda_{2,0} = \bar{\lambda}_2, h = 10$



(a) Bivariate estimates

(b) Univariate estimates

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.

Under this scenario, 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). Note that the effect of jump cross-excitation on return dependence cannot be fully captured in such contour plots, because jump clustering in our model specification is not confined to be contemporaneous in space.

5.2.2 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 6, 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 (c).

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. 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 6), we can see an effect due to the presence of cross-excitation in the bivariate model, although it is less pronounced than for the other two two-index option payoff types.

Table 6: 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.

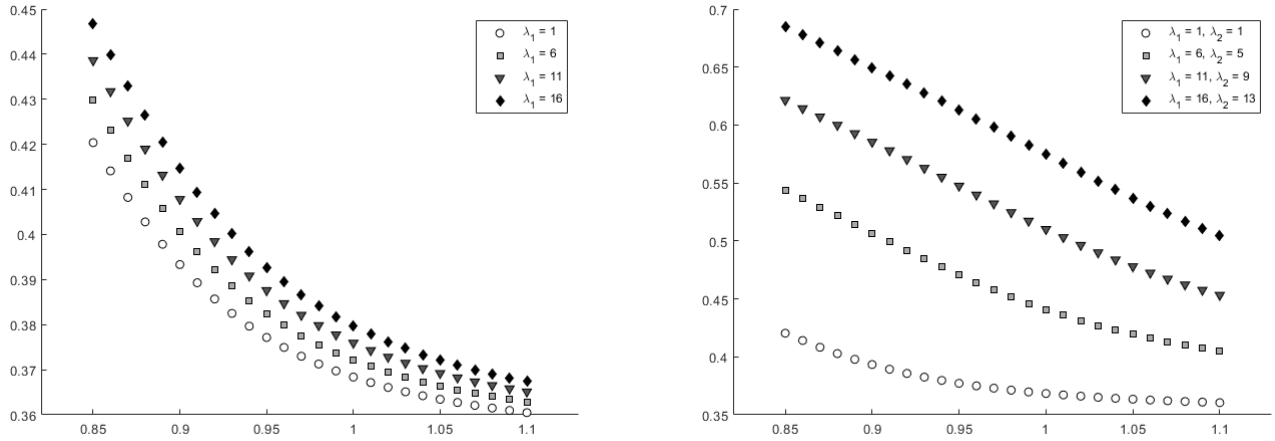
5.2.3 Comparative statics

Finally, with Figure 1 in mind, we are interested in the effect of cross-excitation on 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 4(a) shows changes to implied volatility smiles coming from the different assumed initial levels of the intensity process λ_1 (for fixed λ_2). Although this only captures a marginal

Figure 4: Cross-excitation effects of jumps on implied volatilities



(a) IV_2 when varying λ_1

(b) IV_2 when varying λ_1 and λ_2

Note: This figure plots option implied volatilities (IV_2) 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 λ_1 and λ_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.

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 λ_1 than ITM counterparts. Furthermore, we observe changes in the slopes of the implied volatility curve.

Figure 4(b) plots the implied volatilities when both intensity processes (λ_1 and λ_2) vary. This scenario mimics the occurrence of a jump in the US market: after a shock, the jump intensity λ_1 increases with the value of the self-excitation parameter, and λ_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 λ_2, we observe more pronounced shifts in the implied volatility smile than in Figure 4(a), corroborating once again the importance of jump contagion.

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 of the estimators.

We have estimated the bivariate specification of our model to rich, 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 the strongest 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.

Alter, A., & Beyer, A. (2014). The dynamics of spillover effects during the European sovereign debt turmoil. *Journal of Banking & Finance*, *42*, 134–153.

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.

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.

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.

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.

Boswijk, H. P., Laeven, R. J., & Lalu, A. (2015). Asset returns with self-exciting jumps: Option pricing and estimation with a continuum of moments. Working paper, 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.

Carrasco, M., & Florens, J.-P. (2002). Efficient GMM estimation using the empirical characteristic function. Working paper, Institut d'Économie Industrielle (IDEI), Toulouse.

Da Fonseca, J., Grasselli, M., & Tebaldi, C. (2008). A multifactor volatility Heston model. *Quantitative Finance*, *8*(6), 591–604.

De Marco, S., & Martini, C. (2009). Quasi-explicit calibration of Gatheral's SVI model. *Zeliade White Paper*, 1–15.

Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, *119*(534), 158–171.

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.

Ehrmann, M., Fratzscher, M., & Rigobon, R. (2011). Stocks, bonds, money markets and exchange rates: Measuring international financial transmission. *Journal of Applied Econometrics*, *26*(6), 948–974.

Engle, R. F., Ito, T., & Lin, W.-L. (1990). Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market. *Econometrica*, *58*(3), 525–542.

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.

Gourieroux, C., & Sufana, R. (2010). Derivative pricing with Wishart multivariate stochastic volatility. *Journal of Business & Economic Statistics*, *28*(3), 438–451.

Hamao, Y., Masulis, R. W., & Ng, V. (1990). Correlations in price changes and volatility across international stock markets. *The Review of Financial Studies*, *3*(2), 281–307.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, *58*(1), 83–90.

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.

Jacod, J., & Todorov, V. (2009). Testing for common arrivals of jumps for discretely observed multidimensional processes. *The Annals of Statistics*, *37*(4), 1792–1838.

Jebran, K., Chen, S., Ullah, I., & Mirza, S. S. (2017). Does volatility spillover among stock markets varies from normal to turbulent periods? Evidence from emerging markets of Asia. *The Journal of Finance and Data Science*, *3*(1-4), 20–30.

Karolyi, G. A. (1995). A multivariate GARCH model of international transmissions of stock returns and volatility: The case of the United States and Canada. *Journal of Business & Economic Statistics*, *13*(1), 11–25.

Kokholm, T. (2016). Pricing and hedging of derivatives in contagious markets. *Journal of Banking & Finance*, *66*, 19–34.

Koutmos, G., & Booth, G. G. (1995). Asymmetric volatility transmission in international stock markets. *Journal of International Money and Finance*, *14*(6), 747–762.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *Giornale dell'Istituto Italiano degli Attuari*, *64*, 19–47.

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.

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.

Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, *102*(1), 111–141.

Appendices

Appendix A 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.

Similar to Pan (2002), one can show that the stochastic discount factor $M_{i,t}$ in Eqn. (5) ensures that the deflated index processes $\mathcal{S}^i_{i,t} := M_{i,t}S_{i,t}\exp(\int_0^t q_{i,s}ds)$ 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:

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

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

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

where

$$\begin{aligned}\mathbb{E}[\exp(V^i_i + Z_i) - 1] &= \exp\left(a_{i,i} + \tfrac{1}{2}b^2_{i,i} + \mu_i + \rho_{i,i}b_{i,i}\sigma_i + \tfrac{1}{2}\sigma_i^2\right) - 1 \\ &= \exp\left(\mu_i^{\mathbb{Q}_i} + \tfrac{1}{2}\sigma_i^2\right) - 1 = \mathbb{E}^{\mathbb{Q}_i}[J_{i,t}],\end{aligned}$$

with $\mu_i^{\mathbb{Q}_i} = \mu_i + \rho_{i,i}b_{i,i}\sigma_i$. Therefore, the processes $\mathcal{S}^i_{i,t}$ 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 $\mathcal{S}^i_{j,t} := M_{i,t}E_{ij,t}S_{j,t}\exp(\int_0^t q_{j,s}ds)$ and $\mathcal{B}^i_{j,t} := M_{i,t}E_{ij,t}\exp(\int_0^t r_{j,s}ds)$ need to be local martingales, where $E_{ij,t}$ is the exchange rate be-

tween 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:

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

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

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

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

$$\frac{d\psi_{i,t}}{\psi_{i,t}} = -\eta_i\xi_{i,t}dW_{i,t} + \sum_{k=1}^{m}U_{k,t}^i dN_{k,t}. \tag{A.2}$$

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

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

are standard Brownian motions with the same instantaneous correlations as the original Brownian 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:

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

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 $k, i = 1, \ldots, m$, it follows that

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

and

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

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

$$\chi_{k,t} = \int_0^t J_{k,t}dN_{k,t} - \int_0^t \mathbb{E}[J_{k,s}]\lambda_{k,s}ds, \qquad k = 1,\ldots,m. \tag{A.3}$$

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)),

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

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

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

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

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

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 $k = 1, \ldots, m$, and thus does not change jump times.

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

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

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

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

Then the inner product of two (column) vector functions $\mathbf{f} = (f_1, \ldots, f_k)'$ and $\mathbf{g} = (g_1, \ldots, g_k)'$ 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, $\langle \mathbf{F}, \mathbf{G} \rangle := \int \mathbf{F}(\tau)' \overline{\mathbf{G}(\tau)} \pi(\tau) d\tau$, a $p \times d$ 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:

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

where $\tau = (r, s)'$ with $r, s \in \mathbb{R}^{2m}$, and $m(r, Y_t) = e^{\mathrm{i}r \cdot Y_t}$ 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

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

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

tively.

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 **B.1** *The stochastic process Y_t is a stationary Markov process.*

Assumption **B.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.* θ *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 $\mathbb{E}^{\theta_0}[\mathbf{h}_t(\tau; v_t, \theta_0)] = \mathbf{0}$*,* $\forall \tau \in \mathbb{R}^{2\times 2m}$ π*-almost everywhere, has a unique solution* θ_0 *in the interior of* Θ*.*

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 **B.3** *The sample moment conditions satisfy, as* $T \to \infty$*:*

(i) $\sup_{\theta\in\Theta} \|\mathbf{h}_T(\cdot, v, \theta) - \mathbb{E}^{\theta_0}[\mathbf{h}_t(\cdot, v_t, \theta)]\| \xrightarrow{P} \mathbf{0}$*;*

(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*

$$\mathbf{K} : L^2(\pi)^k \to L^2(\pi)^k, \qquad \mathbf{K}\mathbf{f}(\tau_1) := \int \mathbf{k}(\tau_1, \tau_2)\mathbf{f}(\tau_2)\pi(\tau_2)d\tau_2, \tag{B.1}$$

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 an $k \times k$ 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 **B.4** *Let the non-parametric volatility estimator $\hat{v}_t$ be defined from n high-frequency returns prior to time t, and*

(i) $\hat{v}_t \xrightarrow{P} v_t$, *as* $n \to \infty$;

(ii) $n \to \infty$ *as* $T \to \infty$, *such that* $T/n \to 0$.

Assumption B.4.(ii) is required for the estimation error in $\hat{v}_t$ to be negligible in the large-T 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 2 *Under Assumptions B.1–B.4, as* $T \to \infty$,

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

where

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

with $\mathbf{K}$ *as defined in* (B.1).

Proof: The consistency of the C-GMM procedure follows from Carrasco and Florens (2000) and Boswijk et al. (2015). Based on this, a mean value expansion of $\mathbf{h}_T(\tau, \hat{v}, \hat{\theta})$ 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 θ 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

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

The second term on the right-hand side of the expression above vanishes asymptotically by Assumption B.4. For the first term, Assumption B.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 desired 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 $\mathbf{A}$:

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

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

$$\mathbf{K}_T \mathbf{f}(\tau_1) = \int \mathbf{k}_T(\tau_1, \tau_2) \mathbf{f}(\tau_2) \pi(\tau_2) d\tau_2, \tag{B.2}$$

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

$$T^{-1} \hat{\mathbf{A}}_T^{-1} \hat{\mathbf{B}}_T \hat{\mathbf{A}}_T^{-1},$$

where

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

Appendix C Jump-Robust Volatility Estimation

This appendix provides the details of the jump-robust volatility estimation procedure. We assume that for each day $t = 1, \ldots, T$, we observe $n + 1$ intra-day equity prices at equidistant time points: $S_{t-1+j/n},\ j = 0, \ldots, n$ (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):

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

where r_n is some deterministic sequence, converging to 0 as $n \to \infty$, 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 α on the average of the previous five days' estimates for better option pricing performance.[15]

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 3, 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

[15] 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).

option panel.

Appendix D 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.

D.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 95% $\times$ median 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.

The first filter removes infrequently traded instruments which we deem likely to have illiquid quotes. The second filter aims to strike a balance between data synchronization and quote reliability.

D.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 datasets 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.

D.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. 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.

The standard SVI parametrization of implied total variance, $w(x, \tau)$, with time-to-expiry τ is given as a function of log-moneyness $x = \log(k) = \log(K/F)$ and a parameter set $\chi = \{a, b, \rho, m, \sigma\}$:

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

where $a \in \mathbb{R}, b \geq 0, |\rho| < 1, m \in \mathbb{R}, \sigma > 0$ and $a + b\sigma\sqrt{1 - \rho^2} \geq 0$. Our application regards the stochastic volatility inspired (SVI) model as an interpolation method akin to a polynomial fit. 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 (D.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 τ_1, τ_2 are the closest available maturities

Table 7: 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 1 January 2006 to 13 August 2015. The data are interpolated for each market, each day, and each maturity slice separately.

to τ. After having calibrated an SVI fit for these two volatility smiles, we interpolate between these slices linearly in total variance to τ, which we set equal to 40 days.

Table 7 reports the 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 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.