Supplementary Material to "Jump Contagion among Stock Market Indices: Evidence from Option Markets"

H. Peter Boswijk
Department of Quantitative Economics
University of Amsterdam
and Tinbergen Institute

Roger J. A. Laeven
Department of Quantitative Economics
University of Amsterdam, EURANDOM
and CentER

Andrei Lalu
Department of Quantitative Economics
University of Amsterdam
and Tinbergen Institute

Evgenii Vladimirov
Department of Quantitative Economics
University of Amsterdam
and Tinbergen Institute

June 23, 2023

Abstract

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

Appendix A Model Specification

A.1 Change of Measure

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

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

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

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

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

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

One can show that the stochastic discount factor $M_{i,t}$ in Eqn. (A.1) ensures that the deflated index processes $\mathcal{S}^i_{i,t} := M_{i,t}S_{i,t}\exp(\int_0^t q_{i,s}ds)$ 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 between markets i and j, i.e., the price in currency i of one unit of currency j. This is guaranteed—and hence arbitrage opportunities across all economies are ruled out—whenever the exchange rate dynamics $E_{ij,t}$ are such that $M_{j,t} = M_{i,t}E_{ij,t}$ (see, for example, Brandt and Santa-Clara (2002), Backus, Foresi, and Telmer (2001)).

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

to domestic pricing kernels:

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

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

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

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

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

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

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

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

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

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

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 $\tilde{B}^i_{j,t} := \exp(-\int_0^t r_{i,s} ds) E_{ij,t} \exp(\int_0^t r_{j,s} ds)$ and $\tilde{S}^i_{j,t} := \exp(-\int_0^t r_{i,s} ds) E_{ij,t} S_{j,t} \exp(\int_0^t q_{j,s} ds)$. By applying Itô's formula, the dynamics of these processes under $\mathbb{Q}_i$ can be characterized as follows:

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

Define $G^k_t := \int_0^t \left(e^{V^j_{k,s} - V^i_{k,s}} - 1 \right) dN_{k,s}$ and $H^j_t := \int_0^t \left(e^{Z_{j,s} + V^j_{j,s} - V^i_{j,s}} - 1 \right) dN_{j,s}$. Then, given the assumptions on the zero mean relative jump sizes in the pricing kernels, i.e., $a_{i,k} + \frac{1}{2} b^2_{i,k} = 0$ for $k, i = 1, \ldots, m$, it follows that

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}_i}[G^k_s] &= \mathbb{E}\left[\psi_{i,t} G^k_t \right] \\
&= \mathbb{E}\left[-\int_0^t \eta_i \xi_{i,s} G^k_s \psi_{i,s} dW_{i,s} + \int_0^t \psi_{i,s} \left(e^{V^j_{k,s}} - e^{V^i_{k,s}} + G^k_s \left(e^{V^i_{k,s}} - 1 \right) \right) dN_{k,s} \right] \\
&= 0, \\
\mathbb{E}^{\mathbb{Q}_i}[H^j_t] &= \mathbb{E}\left[\psi_{i,t} H^j_t \right] \\
&= \mathbb{E}\left[-\int_0^t \eta_i \xi_{i,s} H^j_s \psi_{i,s} dW_{i,s} + \int_0^t \psi_{i,s} \left(e^{Z_{j,s} + V^j_{j,s}} - e^{V^i_{j,s}} + H^j_s \left(e^{V^i_{j,s}} - 1 \right) \right) dN_{j,s} \right] \\
&= \mathbb{E}\left[\int_0^t \mathbb{E}^{\mathbb{Q}_j}[J_{j,t}] \psi_{i,s} \lambda_{j,s} ds \right].
\end{aligned}$$

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 $\tilde{B}^i_{j,t}$ and $\tilde{S}^i_{j,t}$ 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

$$\chi_{k,t} = \int_0^t J_{k,t} dN_{k,t} - \int_0^t \mathbb{E}[J_{k,s}]\lambda_{k,s} ds, \qquad k = 1, \ldots, m. \tag{A.4}$$

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

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

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 $U^i_k = e^{V^i_k} - 1$ with $V^i_k \sim \mathcal{N}(a_{i,k}, b^2_{i,k})$, we have

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

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.

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

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

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

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

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

for $i = 1, \ldots, m$. Thus indeed, 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 $\mathbb{Q}_i$ as under the physical measure.

A.2 The Bivariate Specification and Conditional Characteristic Function

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

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

given by

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

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

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

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

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

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

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

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

$0 \leq u \leq T-t$, with initial conditions $\beta_1(0) = \mathrm{i}s_1, \beta_3(0) = 0, \beta_4(0) = 0$ and $\alpha(0) = 0$; for notational convenience, the time dependence in $\beta(u)$ has been omitted from the right-hand side expressions in (A.9). Note that this ODE system does not involve the Brownian price of risk coefficients, the instantaneous correlation coefficient, or the jump size parameters of the second index. An explicit analytic solution of (A.9) is not possible due to the non-linear components involved in the ODE for $\beta_3(u)$ and $\beta_4(u)$. Therefore, we solve this system numerically. Recall that due to the adopted approximation, $v_{1,t}$ is fixed to its value at time t when we price an option expiring at time T. The marginal CCF for the second index, needed to price options on the second index, can be obtained in a similar way.

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

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

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

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

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

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

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.

Appendix B Data Selection and Processing

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

B.1 Option Data Selection

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

(i) retain recordings with message type "Trade" or "Quote";

(ii) retain recordings with a positive Transaction price or recordings with positive Bid and Ask prices;

(iii) for each distinct Reuters Instrument Code (RIC) symbol retain the last Bid, Ask and Transaction price in the reference interval;

(iv) select the Transaction price if available, otherwise calculate the mid Bid-Ask price.

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

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

(i) drop RIC symbols only if all of the following four conditions are met (concurrently):

  (a) the number of either Bid or Ask quotes recorded in the interval is less than or equal to 2;

  (b) there are no trade observations available in the interval;

  (c) the elapsed time between the last Bid and Ask is larger than 10 seconds;

  (d) the spread between last Bid and Ask is larger than 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.

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

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

More specifically, to imply forward prices, we collect for each day all the put-call pairs with the same strike price and maturity, subject to an additional constraint that there are at least two Bid and two Ask quotes for each option during the reference interval. The additional constraint on the number of quotes filters out illiquid options and ensures we obtain reliable forwards. After implying forward prices from all the available put-call pairs, we take the average of the forward prices implied from pairs with the same option maturity and use the resulting term structure of forward prices to calculate Black-Scholes implied volatilities. For this last step, we require risk-free interest rates for each market. In principle, these could also be backed out from box spreads built from the option sets available in each interval, but this would have required an overly complicated option pairing algorithm. We therefore opted to use publicly available data-sets with daily LIBOR-US, LIBOR-GBP and EURIBOR interest rate fixings. We have used linear interpolation for these fixings where needed to match the considered option's maturity.

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

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

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

B.3 Volatility Surface Interpolation

This sub-section provides details of the standard interpolation technique we use to construct the implied volatility data panel, used as input in the estimation procedure. Using an implied volatility option panel as input for the estimation procedure has two advantages. First, it ensures a homogeneous information set is used at each sample observation time-point to imply latent jump intensities from option prices as the grid of (relative) moneyness levels and option maturities is fixed.[1] Second, it reduces computational costs as obtaining model-implied option prices for a fixed set of maturities is computationally less-demanding. We first provide details of the filters employed to select the option price quotes from which implied volatilities are calculated. Next, we provide more information about the interpolation procedure and summary statistics for the resulting implied volatility surfaces.

Defining the moneyness level, k, as the strike-to-forward ratio, i.e., $k = K/F$, we designate an option as an out-of-the-money (OTM) option if it has moneyness level $k > 1.02$ for call options and $k < 0.98$ for put options. We consider options to be close to at-the-money (ATM) if $0.98 \leq k \leq 1.02$. We designate an option as in-the-money (ITM) if it is not OTM or close to ATM. We use call options to imply volatilities when $k > 1$, unless a particular call option has a spread which is more than twice as large as its put counterpart, or the put counterpart was quoted closer to the temporal reference point. A mirrored condition is applied for $k \leq 1$. These conditions trade off the liquidity of relevant options against the synchronicity of the data

[1] 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.

points used as inputs for building volatility smiles. When building implied volatility smiles, we make sure that for each volatility smile the call (put) prices (calculated for all options using put-call parity) are monotonically decreasing (increasing) functions of k.

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

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

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

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

To build the input for our estimation procedure, we calibrate the SVI model at every time point for two volatility slices using a quasi-explicit calibration approach as per De Marco and Martini (2009). For each day we choose two volatility slices such that times-to-maturity for the first slice $\tau_1 \leq \tau$ and for the second $\tau_2 > \tau$, and τ_1, τ_2 are the closest available maturities 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 B.1 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.

Appendix C Estimation Procedure

C.1 Jump-Robust Volatility Estimation

This appendix provides the details of the jump-robust volatility estimation procedure. We assume that for each day $t = 1, \ldots, T$, we observe $n + 1$ intra-day equity prices at equidistant

Table B.1: 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.

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.[2]

[2]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).

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

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

C.2 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, respectively.

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

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

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

(i) $\mathbf{h}_t(\tau; v, \theta)$ *is continuously differentiable w.r.t.* 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 **C.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*

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

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

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

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

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

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

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

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

We are now equipped to state the following proposition:

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

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

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* (C.2)*.*

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

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

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

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

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

where the first elements on the right-hand sides of both equations capture only the direct dependence of the moment function on 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

$$\begin{aligned}
&\sqrt{T}(\hat{\theta} - \theta_0) \\
&= -\left\langle \nabla_\theta \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \nabla_\theta \mathbf{h}_T(\tau, \bar{v}, \bar{\theta}) \right\rangle^{-1} \left\langle \nabla_\theta \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \sqrt{T}\mathbf{h}_T(\tau, v, \theta_0) \right\rangle \\
&\quad - \left\langle \nabla_\theta \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \nabla_\theta \mathbf{h}_T(\tau, \bar{v}, \bar{\theta}) \right\rangle^{-1} \left\langle \nabla_\theta \mathbf{h}_T(\tau, \hat{v}, \hat{\theta}), \sqrt{T}\nabla_v \mathbf{h}_T(\tau, \bar{v}, \bar{\theta})(\hat{v} - v) \right\rangle. \qquad \text{(C.3)}
\end{aligned}$$

The second term on the right-hand side of (C.3) vanishes asymptotically by Assumption

C.4, which is seen as follows. For fixed τ, we have

$$\sqrt{T}\nabla_v \mathbf{h}_T(\tau, \bar{v}, \bar{\theta})(\hat{v} - v) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\partial \mathbf{h}_t(\tau, \bar{v}_t, \bar{\theta})}{\partial \bar{v}_t}(\hat{v}_t - v_t). \tag{C.4}$$

Theorem 8.7 of Aït-Sahalia and Jacod (2014) implies that, as $n \to \infty$, $\{\ell_n^{1/2}(\hat{v}_t - v_t)\}_{t\geq 1}$ converges stably to a sequence $\{\zeta_t\}_{t\geq 1}$ that is, conditionally on $\mathcal{F}_T$, independent Gaussian with mean zero and finite $\mathcal{F}_T$-measurable variances. Rewriting the corresponding Edgeworth expansion (see Yoshida (2013)) as $\hat{v}_t - v_t = \ell_n^{-1/2}\zeta_t + \ell_n^{-1} r_{nt}$, with $r_{nt} = \mathcal{O}_p(1)$, we find that the right-hand side of (C.4) can be written as

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

As $(T, n) \to \infty$, the first term is $\mathcal{O}_p(\ell_n^{-1/2})$ due to a (large T) stable central limit theorem. Because the mean of $\{r_{nt}\}_{t\geq 1}$ may be non-zero, the second term is $\mathcal{O}_p(T^{1/2}\ell_n^{-1}) = \mathcal{O}_p(\sqrt{T/n})$. Therefore, the condition $T/n \to 0$ as $(T, n) \to \infty$ guarantees that (C.4) converges in probability to 0, so that the second right-hand side term in (C.3) is asymptotically negligible.

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

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

Together with consistency and Slutsky's lemma, this yields the 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 $\mathbf{A}$:

$$\begin{aligned}
\hat{\mathbf{A}}_T &= \left\langle \nabla_\theta \mathbf{h}_T(\cdot, \hat{v}, \hat{\theta}), \nabla_\theta \mathbf{h}_T(\cdot, \hat{v}, \hat{\theta}) \right\rangle \\
&= \int \nabla_\theta \mathbf{h}_T(\tau, \hat{v}, \hat{\theta})' \overline{\nabla_\theta \mathbf{h}_T(\tau, \hat{v}, \hat{\theta})} \pi(\tau) d\tau \\
&= \sum_{i=1}^{k} \int \nabla_\theta h_T^{(i)}(\tau, \hat{v}, \hat{\theta}) \overline{\nabla_\theta h_T^{(i)}(\tau, \hat{v}, \hat{\theta})} \pi(\tau) d\tau.
\end{aligned} \tag{C.5}$$

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

$$\mathbf{K}_T\mathbf{f}(\tau_1) = \int \mathbf{k}_T(\tau_1,\tau_2)\mathbf{f}(\tau_2)\pi(\tau_2)d\tau_2, \tag{C.6}$$

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

$$T^{-1}\hat{\mathbf{A}}_T^{-1}\hat{\mathbf{B}}_T\hat{\mathbf{A}}_T^{-1}, \tag{C.7}$$

where

$$\begin{aligned}
\hat{\mathbf{B}}_T &= \Big\langle \nabla_\theta \mathbf{h}_T(\cdot,\hat{v},\hat{\theta}), \mathbf{K}_T\nabla_\theta\mathbf{h}_T(\cdot,\hat{v},\hat{\theta})\Big\rangle \\
&= \int \nabla_\theta \mathbf{h}_T(\tau_1,\hat{v},\hat{\theta})'\overline{\mathbf{K}_T\nabla_\theta\mathbf{h}_T(\tau_1,\hat{v},\hat{\theta})}\pi(\tau_1)d\tau_1 \\
&= \int \nabla_\theta \mathbf{h}_T(\tau_1,\hat{v},\hat{\theta})'\overline{\int \mathbf{k}_T(\tau_1,\tau_2)\nabla_\theta\mathbf{h}_T(\tau_2,\hat{v},\hat{\theta})\pi(\tau_2)d\tau_2}\pi(\tau_1)d\tau_1 \\
&= \sum_{i=1}^{k}\sum_{j=1}^{k}\int \nabla_\theta h_T^{(i)}(\tau_1,\hat{v},\hat{\theta})\overline{\int k_T^{(ij)}(\tau_1,\tau_2)\nabla_\theta h_T^{(j)}(\tau_2,\hat{v},\hat{\theta})\pi(\tau_2)d\tau_2}\pi(\tau_1)d\tau_1.
\end{aligned} \tag{C.8}$$

C.3 Simulation Results

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

Our estimation procedure is designed for the semi-nonparametric specification, in which spot volatilities $\xi_{i,s}$ are "freezed" to their values at time t for some short time interval. In other words, we approximate the stochastic volatilities by the processes $v_{i,s} = \xi_{i,t}$ for $s \in [t,T]$. As has been discussed in Section 2.2, this approximation has negligible errors when pricing options with short expiration time. However, in order to take this approximation into account in our Monte Carlo analysis, we simulate state vector series jointly with the stochastic volatility

processes $\xi_{i,t}$ from a fully parametric specification. In particular, we use the Heston (1993) volatility process:

$$d\xi_{i,t}^2 = \nu_i(\overline{\xi_i^2} - \xi_{i,t}^2)dt + \sigma_{\xi,i}\xi_{i,t}\left(\rho_{\xi,i}dW_{i,t} + \sqrt{1-\rho_{\xi,i}^2}dW_{i,t}^{\xi}\right), \tag{C.9}$$

where the drift term allows for mean-reversion in the volatility process and $W_{i,t}^{\xi}$ is a standard Brownian motion uncorrelated with the Brownian motions $W_{j,t}^{\xi}$, for $j \neq i$, and $W_{i,t}$ in the corresponding index dynamics. Therefore, the Brownian component in (C.9) and in the corresponding index are correlated with constant coefficient $\rho_{\xi,i}$, which captures the leverage effect. Note that although $W_{1,t}^{\xi}$ and $W_{2,t}^{\xi}$ are independent, the Brownian part in one stochastic volatility process is not independent of the Brownian component in the other volatility due to the contemporaneous correlation between $W_{1,t}$ and $W_{2,t}$ in the index dynamics, which in turn we fix to $\varrho = 0.6$ in our simulation study. Finally, when estimating the semi-nonparametric model, we use the true process $v_{i,s} = \xi_{i,t}$ for $s \in [t, T]$.

We simulate the state vector series from the bivariate model specification coupled with the stochastic volatility processes (C.9) 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

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

	$\mu_1^{\mathbb{Q}_1}$	σ_1	κ_1	$\overline{\lambda}_1$	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	$\overline{\lambda}_2$	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, and the semi-nonparametric approximation of Section 2.2. Each iteration consists of 1500 time points including simulated stock prices and 8 option prices for each time observation. True parameters and Monte Carlo sample means, standard deviations and 25%, 50%, 75% quantiles are presented on separate rows. The following parameters are used to simulate the stochastic volatility processes: $\nu_1 = \nu_2 = 4.8$, $\overline{\xi_1^2} = \overline{\xi_2^2} = 0.015$, $\sigma_{\xi,1} = \sigma_{\xi,2} = 0.22$, $\rho_{\xi,1} = \rho_{\xi,2} = -0.6$.

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 (10) (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 (18). Therefore, we run the Monte Carlo simulation with 100 replications, to obtain an (admittedly somewhat crude) indication of the finite-sample performance of the estimators.

The bivariate model specification involves 16 parameters we wish to identify, i.e., 8 pa-

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

The simulation results are provided in Table C.1. We report the true parameter values used in the simulations and the corresponding Monte Carlo means, standard deviations and quantiles of the estimates. Overall, notwithstanding the challenging nature of the econometric problem, the results indicate a good finite-sample performance of our partial-information estimation procedure for the bivariate model. In particular, the self- and cross-excitation parameters, which are of central interest, are estimated with good precision. As is usual, estimates of the Brownian prices of risk, η_1 and η_2, are less precise, due to the fact that their identification is based solely on the return dynamics.

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

The Monte Carlo simulations were also used to investigate the reliability of the asymptotic standard errors, as derived in Appendix C.2. 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

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

	$\mu_1^{\mathbb{Q}_1}$	σ_1	κ_1	$\overline{\lambda}_1$	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	-0.131	0.028	5.722	0.996	2.968	1.078	-0.040	2.522
std	0.015	0.013	0.546	0.278	0.478	0.207	0.013	3.325
25%	-0.138	0.019	5.430	0.812	2.691	0.963	-0.048	0.343
50%	-0.129	0.028	5.714	0.984	3.027	1.088	-0.041	2.599
75%	-0.122	0.032	6.040	1.150	3.235	1.204	-0.034	5.065
	$\mu_2^{\mathbb{Q}_2}$	σ_2	κ_2	$\overline{\lambda}_2$	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	-0.130	0.029	4.838	1.063	2.080	3.068	-0.039	1.759
std	0.014	0.009	0.513	0.443	0.312	0.544	0.011	3.364
25%	-0.135	0.024	4.578	0.782	1.853	2.837	-0.048	-0.868
50%	-0.127	0.028	4.907	1.070	2.103	3.129	-0.040	2.396
75%	-0.122	0.035	5.121	1.271	2.285	3.362	-0.033	4.324

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

Table D.1: Univariate model estimation results for FTSE 100, DAX 30 and S&P 500

	$\mu^{\mathbb{Q}}$	σ	κ	$\overline{\lambda}$	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.

the simulations.

Appendix D Applications

We illustrate the statistical and economic implications of jump contagion in three applications. In the classical GMM tradition, we provide throughout this appendix comparisons between parametric models that delineate only a subset of econometric relationships that are of particular interest. Specifically, we focus on the jump contagion channel that plays a central role in this paper.

To gauge the effect of cross-excitation in the jump components across markets, we first provide estimation results for the univariate model specification. The univariate model can be seen as a nested version of the bivariate specification, 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 4.2. The estimation results of the univariate models for the FTSE 100, DAX 30 and S&P 500 stock market indices are provided in Table D.1.

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.

D.1 Distribution of Index Returns

As a first application, we consider the effect of jump contagion on the (conditional) distribution of index returns, under the physical probability measure $\mathbb{P}$ used for risk management. For this purpose, we simulate forward prices for a pair of indices using the parameter estimates of the bivariate and univariate models from Tables 3 and D.1, 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.

Table D.2: 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: $\lambda_{1,0} = \overline{\lambda}_1, \lambda_{2,0} = \overline{\lambda}_2$									
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: $\lambda_{1,0} = 20, \lambda_{2,0} = \overline{\lambda}_2$									
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: $\lambda_{1,0} = \overline{\lambda}_1, \lambda_{2,0} = 20$									
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 (in percentages), skewness (S), kurtosis (K), and expected number of jumps implied by the conditional distribution of simulated log-returns for S&P 500 ("index 1") and FTSE 100 ("index 2"). The stock index price paths are simulated using bivariate and univariate model parameter estimates, conditional upon different values ("scenarios") of the latent jump intensities. The return horizon is $h = 10$ days. Volatilities are assumed to be constant throughout the horizon and are set to $v_{i,s} = 8.36\%$ for both indices, and the instantaneous correlation between Brownian increments is set to be 0.6.

Since the simulated distribution of log-returns is conditional on the (initial) jump intensity values, we consider five different scenarios to illustrate the effect of jump contagion. Under the base scenario (a), the initial values of the intensities are given by the corresponding estimates of the base rates $\overline{\lambda}_1$ and $\overline{\lambda}_2$, 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 D.2 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 for S&P. 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). To illustrate, translated into 10-days Value-at-Risk capital requirements at the 99% probability level, the effect of accounting for cross-excitation by the bivariate model, using stress scenarios similar to those in the Global Financial Crisis, implies a risk capital increase from about 6 to 9 cents for each dollar invested in the FTSE 100 index. This can be seen from panel (e), column (2) in Table D.2, as $1 - \exp(-0.0661) \approx 6\%$ and $1 - \exp(0.0889) \approx 9\%$.

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.

D.2 Two-Index Options

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

The following two-index option payoff types are considered:

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

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

Two-index option price data points are provided in Table D.3, 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

Table D.3: Two-index options

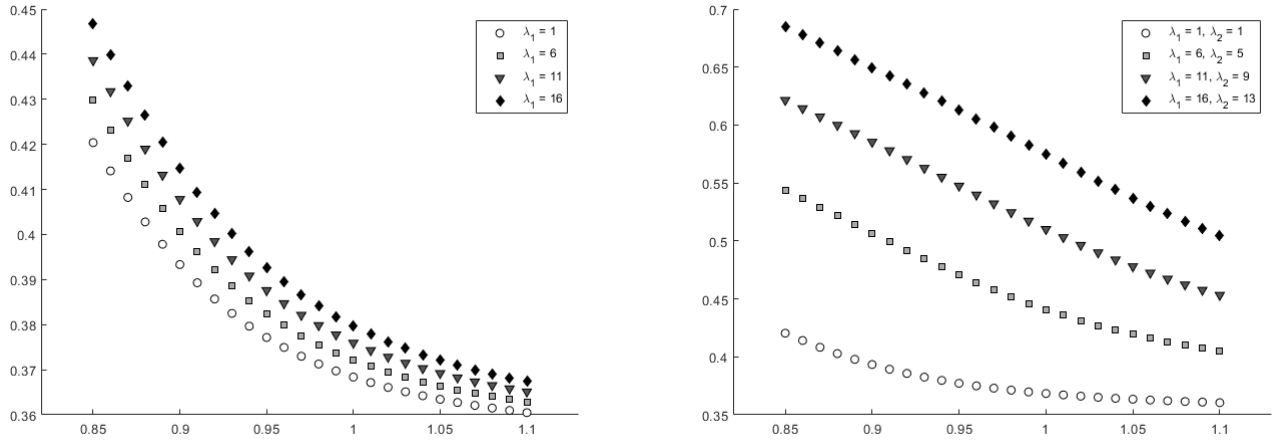
	Single Puts		Correlation		Put on max		Basket	
	S&P	FTSE	h=10	h=30	h=10	h=30	h=10	h=30
(a) $\lambda_{1,0} = \overline{\lambda}_1, \lambda_{2,0} = \overline{\lambda}_2$								
Bivariate	0.214	0.177	0.021	0.188	0.0018	0.0152	0.0042	0.0230
Univariate	0.235	0.182	0.001	0.044	0.0001	0.0050	0.0031	0.0171
(b) $\lambda_{1,0} = \lambda_{2,0} = 5$								
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) $\lambda_{1,0} = 20, \lambda_{2,0} = \overline{\lambda}_2$								
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) $\lambda_{1,0} = \overline{\lambda}_1, \lambda_{2,0} = 20$								
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.

effect. To this point, a larger (in absolute terms) jump size mean and standard deviation under the risk-neutral measure for the S&P series in the univariate model relative to the bivariate counterpart, results in more expensive European puts on the S&P index under the univariate specification than under the bivariate model for all scenarios, except scenario (d).

Nevertheless, we clearly observe the effect that jump cross-excitation has on the pricing of, in particular, correlation and put on max options. The prices of these options are markedly higher under the bivariate model with non-zero cross-excitation than under the univariate model. The strongest effects are found in scenario (c), where the initial jump intensity in the US, which is the leading economy in this pair, is substantially larger than in the UK. We also observe that under the bivariate model, the prices of single puts in asymmetric scenarios are larger due to exposure to shocks in the other market. The results for basket options, which are relatively less sensitive to joint left tail events, depend upon the chosen weights. For the weights $w_1 = 0.3$ and $w_2 = 0.7$ (as in Table D.3), we can see an effect due to the presence of

Figure D.1: 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.

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

D.3 Comparative Statics of Implied Volatilities

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

In particular, we consider again the parameter estimates from the bivariate model for the pair S&P-FTSE and mimic a scenario in which jumps occur in the US market. We fix the volatility levels in both markets to 35% and consider short-dated options with an expiration period of 15 days. Figure D.1(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 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 D.1(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 D.1(a), corroborating once again the importance of jump contagion.

References

Aït-Sahalia, Y., & Jacod, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance*, *53*(2), 499–547.

Backus, D. K., Foresi, S., & Telmer, C. I. (2001). Affine term structure models and the forward premium anomaly. *The Journal of Finance*, *56*(1), 279–304.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37.

Bollerslev, T., & Todorov, V. (2011). Estimation of jump tails. *Econometrica*, *79*(6), 1727–1783.

Boswijk, H. P., Laeven, R. J., & Lalu, A. (2015). *Asset returns with self-exciting jumps: Option pricing and estimation with a continuum of moments* (Tech. Rep.). Amsterdam: University of Amsterdam and Tinbergen Institute.

Brandt, M. W., & Santa-Clara, P. (2002). Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. *Journal of Financial Economics*, *63*(2), 161–210.

Broadie, M., Chernov, M., & Johannes, M. (2007). Model specification and risk premia: Evidence from futures options. *The Journal of Finance*, *62*(3), 1453–1490.

Carrasco, M., & Florens, J.-P. (2000). Generalization of GMM to a continuum of moment conditions. *Econometric Theory*, *16*(6), 797–834.

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

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, *68*(6), 1343–1376.

Fang, F., & Oosterlee, C. W. (2008). A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, *31*(2), 826–848.

Gatheral, J. (2011). *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, *6*(2), 327–343.

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

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

Protter, P. E. (2005). *Stochastic Integration and Differential Equations* (2nd ed.). Springer.

Yoshida, N. (2013). Martingale expansion in mixed normal limit. *Stochastic Processes and their Applications*, *123*(3), 887–933.