Modeling Financial Contagion Using Mutually Exciting Jump Processes*

Yacine Aït-Sahalia
Department of Economics
Bendheim Center for Finance
Princeton University
and NBER

Julio Cacho-Diaz
Department of Economics
Bendheim Center for Finance
Princeton University

Roger J.A. Laeven
Dept. of Econometrics and OR
Tilburg University
and CentER

This Version: March 11, 2010

Abstract

Adverse shocks to stock markets propagate across the world, with a jump in one region of the world seemingly causing an increase in the likelihood of a different jump in another region of the world. To capture this effect mathematically, we introduce a model of asset return dynamics involving mutually exciting jump processes. In the model, a jump in one region of the world or one segment of the market increases the intensity of jumps occurring both in the same region (self-excitation) as well as in other regions (cross-excitation). The model generates the type of jump clustering that is observed empirically. Jump intensities then mean-revert until the next jump. We develop and implement an estimation procedure for this model. Our estimates provide evidence for self-excitation both in the US market as well as in other markets. Furthermore, we find that US jumps tend to get reflected quickly in most other markets, while statistical evidence for the reverse transmission is much less pronounced. Implications of our model are also investigated.

Keywords: Jumps; Contagion; Mutually exciting processes; GMM.

JEL classification: C13; G12.

*We are grateful to seminar and conference participants at Cornell, Princeton, Tilburg, Toulouse, City University of London, University of Melbourne, University of Technology Sydney, the AFMATH Conference in Brussels, the Cambridge-Princeton Conference and the TCF Workshop on Lessons from the Credit Crisis for their comments and suggestions. This research was funded in part by the NSF under grant SES-0850533 (Aït-Sahalia) and by the NWO under grants Veni-2006 and Vidi-2009 (Laeven). *Matlab* code to implement the estimation procedure developed in this paper is available from the authors upon request.

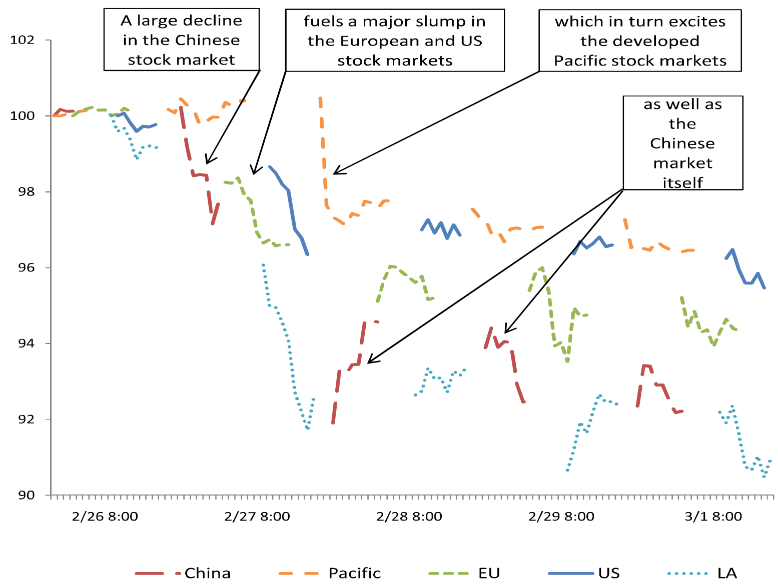

Figure 1: Mutual Excitation: Example I. This figure plots the cascade of declines in international equity markets experienced between February 26, 2007 and March 1, 2007 in the US; Latin America (LA); Developed European countries (EU); China; and Developed countries in the Pacific. Data are hourly. Source: MSCI MXRT international equity indices on Bloomberg.

1. Introduction

Asset market crashes are very unlikely to occur under standard Brownian-driven statistical models, at least with volatility variables calibrated to realistic values. Even more unlikely would be crashes that happen in not just one, but multiple markets around the world at nearly the same time. And, even more unlikely would be further large price moves that happen in close succession over the following days, like earthquake aftershocks. Yet, despite the predictions of standard models, recurring crises happen every decade or so, with sufficient ferocity to overwhelm the statistical assumptions embedded in traditional models used for trading, portfolio management and derivative pricing. These crises seldom have discernible economic causes or warnings, and they tend to propagate across the world with little regard for economic fundamentals in the affected markets.

Of course, jump processes can be employed to capture the large moves in asset markets that continuous models are unable to generate. But the interplay between the various jump terms across markets and over time is not trivial, and standard jump specifications are unable to replicate those patterns. Indeed, adverse shocks to asset markets in one economic sector or region of the world seem to increase the probability of observing successive adverse shocks, not only in the market originally affected but also in other markets. Figures 1 and 2 illustrate two such recent examples, which took place in February 2007 and October 2008, respectively. Figure 1 describes the aftermath of a sharp

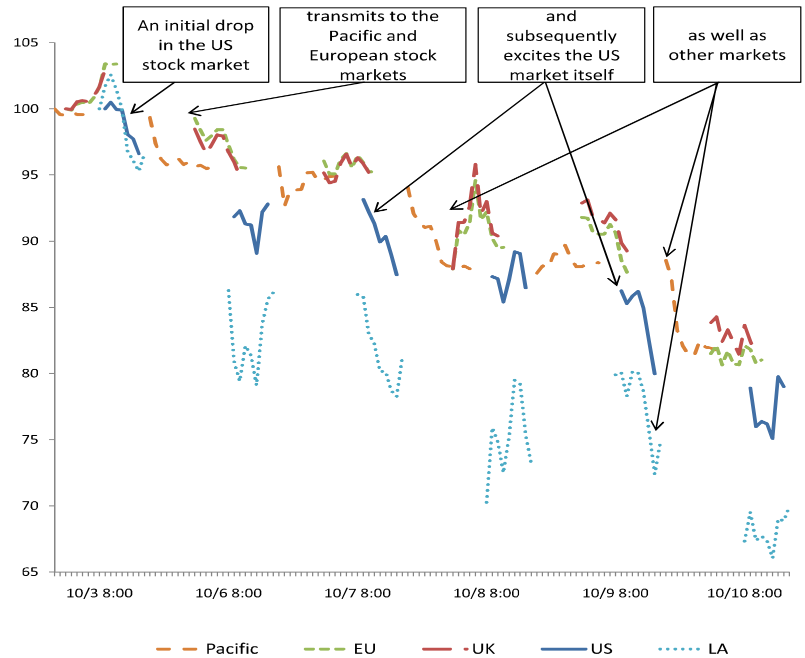

Figure 2: Mutual Excitation: Example II. This figure plots the cascade of declines in international equity markets experienced between October 3, 2008 and October 10, 2008 in the US; Latin America (LA); UK; Developed European countries (EU); and Developed countries in the Pacific. Data are hourly. Source: MSCI MXRT international equity indices on Bloomberg.

decline that originated in the Chinese stock market. Even though this drop was perhaps not the only information that caused the European and US markets to fall –there was some concomitant macroeconomic news on durable goods in the US as well– its effect was felt in other markets in the form of a major downfall that by far exceeded the typical market reaction to macroeconomic news releases. Figure 2 shows what happened on and in the few days that followed October 3, 2008, when the prospects of a \$700 billion bailout of the US financial sector were assessed in markets around the world. The figure shows that foreign markets followed the initial US drop in value by dropping in turn and that successive large price moves were then recorded over a period of days. More generally, when observed over a longer time period, as in Figure 3, where we plot daily stock index returns in six world regions, extreme moves tend to appear in clusters, both in time (horizontally) and in space (vertically).

These figures illustrate two key aspects of asset price jumps that we will model: they are clustered in time, and they tend to contaminate cross-sectionally to other regions. Jump clustering in time is a strong effect in the data. For example, from mid-September to mid-November 2008, the US stock market jumped by more than 5% on eight separate days. In the post-World War II era, jumps of magnitude greater than 5% of either sign have only happened on sixteen other days.

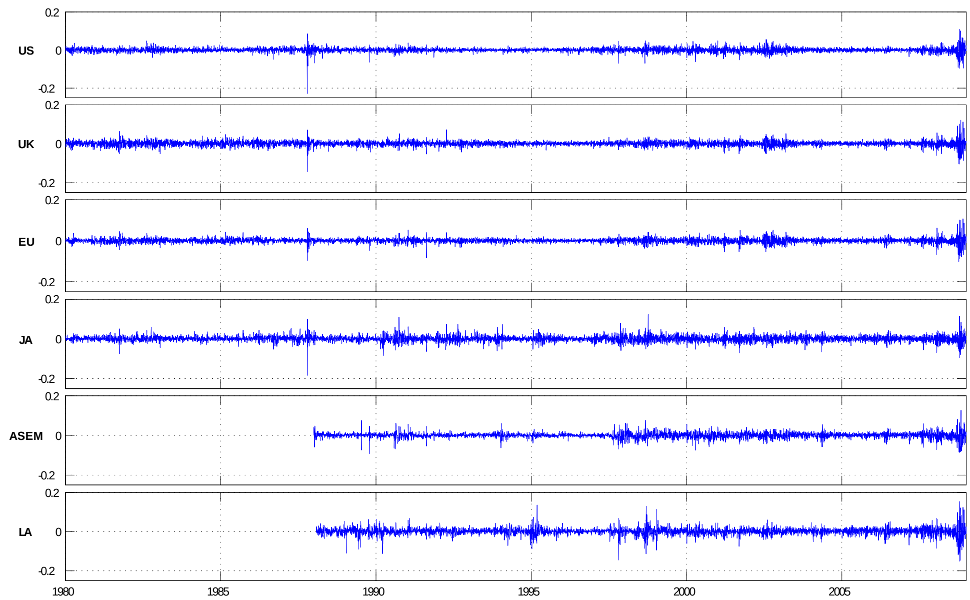

Figure 3: Stock Index Returns in the Six World Regions. This figure plots the log-returns of daily MSCI international equity index data. We consider six series: US; UK; Developed countries Europe (EU); Japan (JA); Emerging markets Asia (ASEM); and Latin America (LA). Sample period US, UK, EU and JA: January 1, 1980 to December 31, 2008; sample period ASEM: January 1, 1988 to December 31, 2008; sample period LA: January 29, 1988 to December 31, 2008. These series are the basis for our empirical analysis contained in Section 5. Descriptive statistics are in Table 1.

And intra-day fluctuations were even more pronounced: during the same two months, the range of intraday returns exceeded 10% during fourteen days. Cross-sectionally, the example in Figure 1 is one of the few cases in which the US stock market is affected by a non-US stock market jump. Usually, stock markets around the world appear to take their cues from the US market. A more typical situation is therefore that illustrated in Figure 2 where US news and market events lead successive market moves elsewhere around the world. As becomes apparent from inspection of the figures, the shocks do not occur simultaneously, especially when viewed at intradaily frequencies. It takes some time for the transmission, if at all, to take place.[1]

In fact, what characterizes a crisis from the point of view of the time series properties of observed returns is typically not the initial jump, which is often limited in scope and in many cases could easily be absorbed, but the amplification that takes place subsequently over hours or days, and the fact that other markets become affected as well.[2] To the best of our knowledge, existing time series

[1] Differences in time zones and trading hours require careful consideration; our treatment of the data is discussed later in the paper.

[2] There is lively debate in the literature regarding the meaning to give to the term "financial contagion," whether it should be distinguished from spillovers, etc.: see e.g., Forbes and Rigobon (2002) for a discussion. Without taking sides in that debate, we use the term "financial contagion" throughout the paper to mean both the cross-region

models used to represent financial crises are not able to generate all these features, dynamic and cross-sectional, together.

To model this type of financial propagation, we need to leave the widely applied class of Lévy jumps, which is the usual class of driving processes employed in the literature. Lévy processes, such as the compound Poisson process, have independent increments: as a result, they do not allow for any type of serial dependence, whence propagation of jumps over time as well as propagation of jumps across markets are key components we wish to capture. So, in this paper, we employ a model for asset return dynamics that captures the cross-sectional and serial dependence observed across stock markets around the world. Mutually exciting jump processes, known as Hawkes processes, are natural candidates for modeling this "contagion" phenomenon mathematically.[3] Basically, in a Hawkes process, a jump somewhere raises the probability of future jumps both in the same region and elsewhere. Jumps in asset returns therefore "self-excite" both in space and in time. In order for the asset returns process to be stationary, we then make the degrees of excitation of the various jumps, or jump intensities, mean revert until the next jump.

These jumps, by virtue of their self- and cross-excitation, introduce a feedback element. This aspect of the model can be thought of as playing the same role for jumps as ARCH does for volatility (see Engle (1982)). Namely, the ARCH model introduces feedback from returns to volatility and back: large returns lead to large volatilities which then make it more likely to observe large returns. In the absence of further excitation, volatility then reverts to its steady state level. Here, similarly, jumps lead to larger jump intensities, which then make it more likely to observe further jumps. In the absence of further excitation, jump intensities then revert to their steady state level.

Existing applications of Hawkes processes have typically employed them as pure point processes. As a result, the paths of the variables of interest will typically be piecewise constant, which is appropriate in many settings. Here, however, we wish to model asset prices. While our focus is on

transmission of shocks and the increase in the likelihood of successive shocks over time in the affected countries following an initial shock. While contagion in this broad sense can take place both during "good" times as well as during crisis times, the contagion phenomenon is more prevalent during crisis times.

[3] Hawkes processes were originally proposed as mathematical models to represent the transmission of contagious diseases in epidemiology. They have also found applications in neurophysiology and in the modeling of earthquake occurrences (see, e.g., Brillinger (1988) and Ogata and Akaike (1982)). In market microstructure, Bowsher (2007) employs them to jointly model transaction times and price changes at high frequency. Salmon and Tham (2007) study how trading activity at one maturity of the yield curve affects and is affected by trading at other maturities. In their model, the instantaneous volatility is a function of the price intensity which follows a self-exciting process. They find that the yields at different maturities appear to be driven by different trading clocks, which they model through a multivariate Hawkes process.

Self-exciting models are now also being employed to model joint defaults in portfolios of credit derivatives; see e.g., Azizpour and Giesecke (2008). Errais et al. (2010) show how portfolio credit risk can be evaluated in a model of correlated credit default timing based on self-exciting processes. In their setting, the jump-diffusion process is the partly latent risk factor impacting the jump intensity while the component jump process models financial loss due to default. Using affine processes, the conditional joint transform of the jump-diffusion risk factor process and the jump process can be computed explicitly, leading to the formulation of an affine jump-diffusion in an extended state space, that is amenable to tractable credit risk pricing applications.

Hawkes processes have also been proposed in the literature on social interactions to model the "viral propagation" of some phenomena, such as the viewing of YouTube videos (see Crane and Sornette (2008).)

the return dynamics in times of crisis, we also wish to provide a model in which asset prices move normally and financial crises are rare. We will therefore not consider Hawkes processes on their own, but rather include them on top of the usual asset price components consisting of a drift or expected return, and a continuous or volatility component.

The paper makes two separate contributions. First, we propose a model consisting of a mutually exciting jump component added to a continuous Brownian component with stochastic volatility, as well as a drift term, which we refer to as a *Hawkes jump-diffusion* model by analogy with the Poissonian jump-diffusion model familiar to financial economists since Merton (1976). In the model, mutually exciting jumps are there to capture crises, while the remaining components are there to represent the evolution of the asset returns in normal times. Some important financial implications of the model we propose are investigated. Second, we develop and implement an estimation procedure for this model. Our estimation procedure is based on the generalized method of moments (GMM) using integrated moments of the model, which we derive in closed-form. This closed-form feature has some important advantages: it means, among other aspects, that we can deal with multivariate models and that the estimation is sufficiently fast to conduct an extensive Monte Carlo study to test its accuracy.[4,5]

When Hawkes processes are considered to model credit defaults, unlike our setting, the jump events in credit risk (defaults) are directly observable from the data. Here, due to the fact that neither the stochastic volatilities nor the jumps and jump intensities are directly observable, the corresponding inference problem we face is particularly challenging, and we develop an estimation procedure which explicitly accounts for the latency of part of the state vector. In brief, we do this by first computing the conditional moments using the *full state vector*: asset returns, stochastic volatilities, jumps and jump intensities; then conditioning down by taking expected values over the *latent state variables*: volatilities, jumps and jump intensities. This results in expressions that depend only upon the *observable state variables*, namely the asset returns, but involve all the parameters of the model, including those driving the latent state variables.

We estimate our model using international stock index returns for six world regions. The empirical analysis indicates that both the US market as well as other markets strongly self-excite

[4] The study of multivariate asset return models with jumps has recently seen a lot of activity. For example, Ang and Chen (2002) and Das and Uppal (2004) who study a Poisson-driven jump-diffusion model (a multivariate version of the Merton (1976) model). Ang and Chen (2002) find empirically that it fails to capture persistence of covariance dynamics in the data and captures almost no asymmetric correlation effects. Interesting to note is that none of the models considered in Ang and Chen (2002), which in addition to the jump-diffusion model include an asymmetric GARCH model, a regime-switching Gaussian model and a regime-switching GARCH model, completely explains the extent of observed asymmetries in correlation. Hawkes jump-diffusion models of the type that we study exhibit both persistence of co-variability and asymmetric co-variability and are therefore expected to outperform the simpler Poisson-driven jump-diffusion model. And, while the Poisson-driven jump-diffusion model studied by Ang and Chen (2002) and Das and Uppal (2004) captures only systemic jump risk, the jump risk model we propose captures both systemic jump risk and idiosyncratic jump risk.

[5] In a univariate setting, Maheu and McCurdy (2004) and Yu (2004) provide clear evidence for jump clustering, supporting self-exciting jump models. Yu (2004) derives unconditional moments and autocovariances for the jump model proposed by Knight and Satchell (1998), which is of a different type than the univariate version of the mutually exciting model we study.

over time. Cross-sectionally, we find that US jumps tend to get reflected quickly in most other markets, while statistical evidence for the reverse transmission is much less pronounced.

As a model for asset prices that explicitly accounts for crises, our model is decidedly reduced-form. As such, it cannot explain the source(s) of the contagion that is observed in the data in times of crises, or get at the channels of transmission of that contagion, whether they are trade linkages, financial linkages, financial constraints, outflows of capital, herding behaviors, the fragility of the system, lack of coordinated responses, etc. In all likelihood, all are important to different degrees at different times in different crises, and there is an important literature, both theoretical and empirical, on these various mechanisms.[6] Our model is merely a framework to quantify the nature and extent of the observed contagion, not an attempt at isolating the transmission mechanism(s).

But, once estimated on the data, this reduced-form model can be employed as a description of the process driving the asset returns, notably their risk. In the Merton tradition, such reduced-form models are classically employed in finance for derivative pricing, portfolio choice or risk management, among others. In this context, the model is amenable to at least four important financial applications.

First, we discuss how the mutually exciting jump intensities of the model can be filtered out of the observed asset returns, and propose the resulting time series as a measure of market stress in lieu perhaps of volatility-based measurements such as the volatility index (VIX) which aggregates together both diffusive and jump risks. A second application of the model consists in quantifying the magnitude of the tails of asset returns. We evaluate the joint tails over the typical ten-day horizon that is employed in Value-at-Risk (VaR) calculations, and compute the systemic risk inherent in multiple assets jumping together in the same time period, comparing in particular the prediction of Poissonian jump-diffusions vs. Hawkes jump-diffusions. We find that most of the autocorrelation and cross-correlation patterns due to the mutually exciting component of the model tend to be important but die out after a few days, making the model ideally suited for this purpose. Third, it is possible to derive the optimal composition of an investor's portfolio which is subject to continuous Brownian moves and mutually exciting jumps. Evidently, controlling exposure to jumps should be done on the basis of the most realistic model for correlated shocks. This motivates the study of more subtle forms of propagation in the form not just of simultaneous jumps within or across sectors, but rather in the form of a jump in one sector causing an increase in the likelihood that a different jump will occur in another sector.[7] Finally, in terms of contingent claim pricing, we note that our model

[6] See, e.g., Calvo and Mendoza (2000), Chang and Velasco (2001), Caramazza et al. (2004), Dornbusch et al. (2000), Dungey and Gonzalez-Hermosillo (2005), Dungey and Martin (2007), Eichengreen et al. (1996), Fry et al. (2009), Gerlach and Smets (1995), Rigobon (2003), Glick and Rose (1999), Kaminsky and Reinhart (1998), Kaminsky and Reinhart (2000), Kodres and Pritsker (2002), Krugman (1979), Nikitin and Smith (2008), Obstfeld (1996), Pavlova and Rigobon (2008) and Rijckeghem and Weder (2001).

[7] Other papers have studied the impact of time-varying correlations and jumps on portfolio choice in a framework without mutual excitation: Ang and Bekaert (2002) consider a regime-switching model, consisting of one regime with low correlations and low volatilities and one regime with higher correlations, higher volatilities and lower conditional means. They find that the existence of a higher volatility bear market regime does not nullify the benefits of international diversification, as long as the investor dynamically rebalances his portfolio. Das and Uppal (2004) find empirically that the loss reduction in diversification due to transmission across equity markets is not substantial.

can be restricted to fit the rich class of affine jump-diffusion models, in their generalized version allowing for multiple jump types defined in Appendix B of Duffie et al. (2000). An affine special case of our model would therefore share in the tractable pricing implementation that results from an affine structure, following Errais et al. (2010) who showed how to conduct credit risk pricing in the context of an affine self-exciting process, based on the explicit joint transform of intensities and price process that arises in an affine model.

The rest of this paper is organized as follows: In Section 2, we present the model of asset returns, and discuss some of its properties. In Section 3, we develop our estimation procedure. In Section 4, we study the finite sample behavior of our estimators on the basis of a Monte Carlo study. Section 5 contains the empirical analysis. In Section 6, we investigate some implications of our model and its financial applications. Conclusions are in Section 7.

2. Asset Return Dynamics

In our model, asset prices are subject to Brownian volatility as well as jumps. Relative to the classical jump-diffusion model in finance, which originated in Merton (1976), we allow the jumps to be mutually exciting in the fashion of Hawkes (1971a) as a means of capturing the contagion phenomenon: that is, a jump somewhere raises the likelihood of getting a jump elsewhere in the near future.

2.1. Jump Dynamics: Mutually Exciting Processes

Hawkes (1971a) (see also Hawkes (1971b), Hawkes and Oakes (1974) and Oakes (1975)) introduced mutually exciting processes, which are special cases of path-dependent point processes. The intensities of a multivariate mutually exciting process depend on the paths of the point process; hence, the jump intensities are themselves stochastic processes and will be part of the state vector. The couple consisting of the jump process and its intensity remains a Markov process.

The jump intensities ramp up in response to jumps in one of the marginal point processes. We consider m point processes $N_{i,t}$, $i = 1, \ldots, m$. In our application, there will be one such jump process $N_{i,t}$ for each of the m regions of the world that we study. Of course, alternative breakdowns of the sample are possible, isolating sectors of the economy, for instance.

Similar to the familiar Poisson process, the Hawkes process is defined by its intensity process $\lambda_{i,t}$ which describes the $\mathcal{F}_t-$conditional mean jump rate per unit of time, namely

$$\begin{cases} \mathbb{P}\left[N_{i,t+\Delta} - N_{i,t} = 0 | \mathcal{F}_t\right] = 1 - \lambda_{i,t}\Delta + o(\Delta) \\ \mathbb{P}\left[N_{i,t+\Delta} - N_{i,t} = 1 | \mathcal{F}_t\right] = \lambda_{i,t}\Delta + o(\Delta) \\ \mathbb{P}\left[N_{i,t+\Delta} - N_{i,t} > 1 | \mathcal{F}_t\right] = o(\Delta) \end{cases} \tag{2.1}$$

Aït-Sahalia et al. (2009) derive a closed-from solution to the portfolio choice problem with standard systematic, not mutually exciting, jumps across asset returns.

but instead of being constant the jump intensities follow the dynamics

$$\lambda_{i,t} = \lambda_{i,\infty} + \sum_{j=1}^{m} \int_{-\infty}^{t} g_{i,j}\left(t-s\right) dN_{j,s}, \qquad i = 1, \ldots, m. \tag{2.2}$$

In other words, each previous jump $dN_{j,s}$, $s \in (-\infty, t)$, $j = 1, \ldots, m$, raises the jump intensities $\lambda_{i,t}$. Taken jointly, (N, λ) is a Markov process. The distribution of the jump processes $N_{j,s}$ is determined by that of the intensities $\lambda_{j,s}$. Each compensated process $N_{i,t} - \int_{-\infty}^{t} \lambda_{i,s} ds$ is a local martingale. Hawkes processes are related to, but essentially different from, doubly stochastic Poisson (or Cox) processes: In a doubly stochastic Poisson process, the jump intensities are also stationary stochastic processes but the conditioning σ-algebra in (2.1) is $\{N_{i,s}(s \le t) : \lambda_{i,s}(-\infty < s < \infty)\}$, i.e., the path of $\lambda_{i,t}$ is not affected by the path of $N_{i,t}$; see e.g., Karr (1991) for further details on doubly stochastic Poisson processes.

We assume that in (2.2) the constant parameters $\lambda_{i,\infty} \ge 0$ for all $i = 1, \ldots, m$, that the real-valued functions $g_{i,j}(u) \ge 0$ for all $u \ge 0$ and for all $i, j = 1, \ldots, m$. Notice that $\lambda_{i,\infty}, g_{i,j}$ being non-negative for all $i, j = 1, \ldots, m$ ensures the non-negativity of the intensity processes with probability one.

Let Λ_∞ denote the $m \times 1$ vector with components $\lambda_{i,\infty}$ and $\Gamma := \int_0^\infty G_u du$ the $m \times m$ matrix where G_u is the matrix with elements $g_{i,j}(u)$'s. Let $\lambda_i := \mathbb{E}\left[\lambda_{i,t}\right]$ denote the unconditional expected jump intensity, and Λ the $m \times 1$ vector with components λ_i. Since from (2.1), $\mathbb{E}\left[dN_{i,s}\right] = \lambda_i ds$, we see that

$$\lambda_i = \lambda_{i,\infty} + \sum_{j=1}^{m} \lambda_j \int_{-\infty}^{t} g_{i,j}\left(t-s\right) ds = \lambda_{i,\infty} + \sum_{j=1}^{m} \left(\int_0^\infty g_{i,j}\left(u\right) du \right) \lambda_j, \tag{2.3}$$

or in vector form $\Lambda = \Lambda_\infty + \Gamma \cdot \Lambda$. Therefore $\Lambda = \left(I - \Gamma\right)^{-1} \Lambda_\infty$, where I is the identity matrix, and we assume that all the elements of Λ are positive and finite. This will ensure stationarity of the model.

2.2. Full Dynamics: Mutually Exciting Jump-Diffusion

The mutually exciting jump process constitutes only part of our model of asset returns. We assume that asset log-returns follow the semimartingale dynamics

$$dX_{i,t} = \mu_i dt + \sigma_i dW_{i,t} + Z_{i,t} dN_{i,t}, \qquad i = 1, \ldots, m, \tag{2.4}$$

which consists of a drift term, a volatility term, and mutually exciting jumps. Here, $W_t := [W_{1,t}, \ldots, W_{m,t}]'$ is an m-dimensional vector of standard Brownian motions with constant correlation coefficients $\rho_{i,j}$, $i, j = 1, ..., m$, $Z := [Z_1, \ldots, Z_m]'$ is the vector of jump sizes, cross-sectionally and serially independently distributed with laws F_{Z_i} supported on $(-\infty, \infty)$, and $N_t := [N_{1,t}, \ldots, N_{m,t}]'$ is the vector of Hawkes processes just described. Throughout, all dynamics are with respect to the objective probability measure and not to an equivalent martingale measure. Z_i may be replaced by $\log\left(1 + \tilde{Z}_i\right)$ so that the discontinuous part in the differential of the stochastic exponential

$e^{X_{i,t}}$ becomes $\tilde{Z}_i e^{X_{i,t}} dN_{i,t}$. For notational convenience we write Z_i rather than $\log\left(1+\tilde{Z}_i\right)$. We will leave the distribution of jump sizes essentially unrestricted, so asymmetries such as negative jumps (crashes) being more likely than positive jumps (booms), can be built into the jump size distributions.

In the base model (2.4), the quantities μ_i and σ_i are constant parameters. In this case we assume that Σ, the $m \times m$-dimensional variance-covariance matrix of the risk coming from the continuous part of the model, with elements $\Sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j$, is a non-singular matrix. We always assume that the vector of Brownian motions W, the vector of jump sizes Z and the vector of jump processes N are mutually independent.

In an extension of the model (2.4), we allow for stochastic volatility:

$$dX_{i,t} = \mu_i dt + \sqrt{V_{i,t}} dW_{i,t}^X + Z_{i,t} dN_{i,t}, \tag{2.5}$$

where the instantaneous variance follows the Heston (1993) model

$$dV_{i,t} = \kappa_i(\theta_i - V_{i,t})dt + \eta_i \sqrt{V_{i,t}} dW_{i,t}^V, \tag{2.6}$$

with κ_i, θ_i, η_i constant parameters. Model (2.4) corresponds to the special case where the initial value is $V_{i,0} = \theta_i$ and $\eta_i = 0$. Note that $V_{i,t}$ is a local variance rather than a local standard deviation; while keeping this in mind we will refer to $V_{i,t}$ as the stochastic volatility variable. $V_{i,t}$ follows the square root process of Feller (1951) and is bounded below by zero. The boundary value zero cannot be achieved as long as Feller's condition, $2\kappa_i\theta_i \geq \eta_i^2$, is satisfied.

The extended model allows for correlations among the individual Brownian motions driving equations (2.6) as well as between the individual Brownian motions driving equations (2.5) and (2.6), thereby capturing a potential leverage effect. But in the presence of systematic jumps, the Brownian correlation, even if it increases, will not play a major role: in times of crisis, jumps swamp everything else.[8]

Our estimation procedure can cover general stochastic volatility specifications, and we show below how to derive closed-form expressions for observable moments of the log-returns. But we need to settle on a specific model with interpretable parameters to do the actual parametric estimation, and for this purpose we assume that the volatility follows the model (2.6).

Our model, with mutually exciting jump processes added to a continuous Brownian component with (possibly stochastic) volatility as well as a drift, will be referred to as a *Hawkes jump-diffusion*. Hawkes jump-diffusion processes will generate observed time-varying correlations and maximal correlations around crisis times, due to the systematic jumps. Isolating a change in the structure of the Brownian variance-covariance matrix Σ becomes difficult in this context, because linear correlation measures weigh equally all returns; as a result, they will tend to average out any

[8] The linear correlation measure of dependence is not an appropriate measure for the type of financial propagation observed in crises (see Bae et al. (2003) and Longin and Solnik (2001) for a discussion). Abandoning the correlation framework, Bae et al. (2003) use a multinomial logistic regression model to model and evaluate joint occurrences of large absolute value returns.

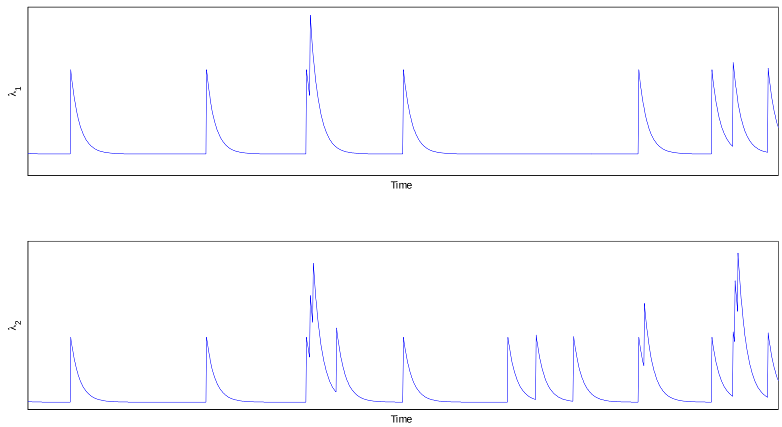

Figure 4: Jump Intensity Processes: A Sample Path. This figure plots a sample path of the intensity processes $(\lambda_{1,t}, \lambda_{2,t})$ generated by a bivariate mutually exciting Hawkes process with exponential decay (see Section 2.3).

contagion that happens over a limited number of days among the comparatively large number of days where no jumps occur.[9] Similarly, the observed clustering of extreme returns is generated by Hawkes jump-diffusion processes in a different manner from pure stochastic volatility models where periods of high volatility can also lead to clusters of larger absolute returns, although not of a magnitude and rate of occurrence compatible with what is actually observed in the data.

2.3. Mean Reversion in Jump Intensities

The tractability of the jump part of our model, and hence the possibility of estimating it, depends on the parameterization of the intensity processes $\lambda_{i,t}$. A special case of interest occurs when

$$g_{i,j}(t-s) = \beta_{i,j} e^{-\alpha_i (t-s)}, \qquad s < t, \quad i, j = 1, \ldots, m, \tag{2.7}$$

with $\alpha_i > 0$, $\beta_{i,j} \geq 0$ for all $i, j = 1, \ldots, m$. In this case, a jump in asset prices causes the intensities to jump up, and then the intensity decays exponentially back: $\lambda_{i,t}$ jumps by $\beta_{i,j}$ whenever a shock in sector j occurs, and then decays back towards a level $\lambda_{i,\infty}$ at speed α_i. Figure 4 illustrates a sample path of the jump intensity processes in the bivariate mutually exciting case. Asset return dynamics of equity markets that are highly interrelated may exhibit jumps that take place almost

[9] Analyzing correlation structures between asset returns, Goetzmann et al. (2005) find that (linear) correlations across world markets vary significantly over time. Evidence from capital markets history suggests that periods of poor market performance are associated with high correlations. They find that average correlations went up and reached a peak in the 1930's (Great Depression) that has been unequaled until the modern era. Bekaert and Harvey (1995) and Bekaert and Harvey (2000) provide evidence that market integration and financial liberalization change the comovement of emerging markets stock returns with the global market factor.

simultaneously. In our model this corresponds to the case in which mutual excitation is very severe (large $\beta_{i,j}$, $i \neq j$).

With this model, the Γ matrix is given by

$$\Gamma = \begin{pmatrix} \frac{\beta_{11}}{\alpha_1} & \dots & \frac{\beta_{1m}}{\alpha_1} \\ \vdots & \ddots & \vdots \\ \frac{\beta_{m1}}{\alpha_m} & \dots & \frac{\beta_{mm}}{\alpha_m} \end{pmatrix}. \tag{2.8}$$

Under exponential decay (2.7), each jump intensity has the mean-reverting dynamics

$$d\lambda_{i,t} = \alpha_i \left(\lambda_{i,\infty} - \lambda_{i,t}\right) dt + \sum_{j=1}^{m} \beta_{i,j} dN_{j,t}. \tag{2.9}$$

Other specifications of the intensity process may also be considered and it is possible to obtain the covariance density matrix and spectral density for general $g_{i,j}$ functions. In even more general specifications, the intensity may depend not only on the amount of time elapsed since a jump event but also on the size of past jump events.

The jump part of the model is able to generate the two features we are after: First, jump activity is variable over time, with many (of the few) jumps typically being concentrated in short periods of time; this is jump clustering. Second, adverse shocks propagate across the world in a contagious way, with adverse events in one region of the world seemingly increasing the likelihood of shocks in other regions of the world; this is jump propagation. Under exponential decay, the element-by-element Laplace transform $V_N^*(s)$ of the covariance density matrix $V_N(\tau)$ is given by

$$V_N^*(s) = \left(I - G^*(s)\right)^{-1} \left(G^*(s)\text{Diag}(\Lambda) + \beta V_N^{*\,\prime} \{\alpha\}/\{\alpha + s\}\right), \tag{2.10}$$

with $G(\tau)$ the $m \times m$-matrix given by

$$G(\tau) = \begin{pmatrix} \beta_{11} e^{-\alpha_1 \tau} & \cdots & \beta_{1m} e^{-\alpha_1 \tau} \\ \vdots & \ddots & \vdots \\ \beta_{m1} e^{-\alpha_m \tau} & \cdots & \beta_{mm} e^{-\alpha_m \tau} \end{pmatrix}, \tag{2.11}$$

and $G^*(s)$ its element-by-element Laplace transform. The indices of $\{a\}$ and $\{a + s\}$ correspond to the element-by-element indices of β.

With stochastic volatilities and stochastic jump intensities added to the state vector, our model can be restricted to be part of the class of generalized affine jump-diffusion processes. To see this, consider a generalized affine jump-diffusion A in a state space $D \subset \mathbb{R}^{3\times m}$, defined as a strong solution to the stochastic differential equation

$$dA_t = \mu^A(A_t)dt + \sigma^A(A_t)dW_t^A + \sum_{j=1}^{m} dJ_{j,t}, \tag{2.12}$$

where $\mu^A : D \to \mathbb{R}^{3\times m}$, $\sigma^A : D \to \mathbb{R}^{(3\times m)\times(3\times m)}$, W^A is a Brownian motion in $\mathbb{R}^{3\times m}$, and J_i, $i = 1, \dots, m$, are pure jump processes with jump intensities $\lambda_{i,t}^A = \lambda_i^A(A_t)$, for some $\lambda_i^A : D \to [0, \infty)$, and with fixed jump size distributions on $\mathbb{R}^{3\times m}$. It is possible to restrict a process A of the form (2.12) to be affine, by considering the special case where $\mu^A, \sigma^A\sigma^{A\prime}$ and λ_t^A are affine on D. Then our Hawkes jump-diffusion model with exponential decay can be restricted to be affine by setting $A_t = [X_t, V_t, \lambda_t]'$ with the corresponding $\mu^A, \sigma^A\sigma^{A\prime}$ and λ_t^A being affine.

2.4. Jump Size Distribution

We have noted that the analysis can proceed without assumptions on the distribution of the jump magnitudes, and in fact, we will provide expressions for the moments of the process as functions of the generic moments of the jump size Z_t, which we denote $M[Z,k] := \mathbb{E}\left[Z_t^k\right]$. In order to estimate a specific parametric model, however, we need to parameterize these moments to reduce the dimensionality of the parameter space, given that the distribution of jumps of finite activity is difficult to pin down since large jumps are by nature rare.

For this purpose, we will assume that Z_i is a scalar random variable with cumulative probability distribution

$$F_{Z_i}(x) = \begin{cases} p_i e^{-\gamma_{i,1}(-x)}, & -\infty < x \leq 0; \\ p_i + (1-p_i)(1-e^{-\gamma_{i,2}x}), & 0 < x < \infty; \end{cases} \tag{2.13}$$

where $\gamma_{i,1}, \gamma_{i,2} > 0$ and $0 \leq p_i \leq 1$, $i = 1, \ldots, m$. The corresponding density is

$$f_{Z_i}(x) = \begin{cases} p_i \gamma_{i,1} e^{-\gamma_{i,1}(-x)}, & -\infty < x \leq 0; \\ (1-p_i)\gamma_{i,2} e^{-\gamma_{i,2}x}, & 0 < x < \infty. \end{cases} \tag{2.14}$$

One easily verifies that

$$\mathbb{E}\left[Z_i^k\right] = (-1)^k \frac{k!\ p_i}{\gamma_{i,1}^k} + \frac{k!\ (1-p_i)}{\gamma_{i,2}^k}, \qquad k = 1, 2, \ldots. \tag{2.15}$$

This jump size distribution is also used by Kou (2002) in the context of option pricing.

In our empirical study below we use equity index data. As such, our analysis does not suffer from survivorship bias. However, survivorship bias, when present (for example, when individual stock returns are used), can easily be dealt with in this model. Explicitly modeling the survival and death process involves introducing a point mass at $Z_i = -\infty$ in the distribution of the jump size for log-returns.

3. Estimation Procedure

Neither the point processes $N_{i,t}$ and the intensity processes $\lambda_{i,t}$, nor the stochastic volatilities $V_{i,t}$, $i = 1, \ldots, m$, are directly observable. Instead, what is observed are asset prices, hence log-returns $X_{i,t}$. It turns out that we can derive explicit expressions for the moments of the log-returns that are implied by the model, as a function only of the observable state variables, in effect integrating out the unobservable state variables. We will therefore develop a GMM-based estimation procedure.

3.1. Explicit Expressions for the Moments: Markov Infinitesimal Generator

The key to the estimation procedure is the availability in closed-form of the main moments for the model. Having the moment conditions in closed-form means that the effort involved in minimizing the GMM criterion function becomes minimal, despite the number of parameters and moment functions. The moment conditions that we will use to carry out GMM estimation of the model are:

$$\begin{cases} \mathbb{E}\left[\Delta X_{i,t}\right] \\ \mathbb{E}\left[\left(\Delta X_{i,t} - \mathbb{E}\left[\Delta X_{i,t}\right]\right)^r\right], \qquad r = 2, \ldots, 4 \\ \mathbb{E}\left[\Delta X_{i,t}\Delta X_{j,t} - \mathbb{E}\left[\Delta X_{i,t}\right]\mathbb{E}\left[\Delta X_{j,t}\right]\right], \qquad i \neq j \\ \mathbb{E}\left[\Delta X_{i,t+\tau}\Delta X_{j,t} - \mathbb{E}\left[\Delta X_{i,t}\right]\mathbb{E}\left[\Delta X_{j,t}\right]\right], \qquad \tau > 0. \end{cases} \tag{3.1}$$

As we will see, the reason for using these moment functions is two-fold. First, they are "natural" in the sense of being easily interpretable: variance, kurtosis, autocovariances. Second, and more importantly, each moment function plays a specific role in identifying parts of the model.

Throughout the analysis, we use unconditional moments as opposed to conditional moments. To compute the unconditional moments, we first compute the conditional moments using the full state vector: asset returns, stochastic volatilities, jumps and jump intensities. Next, we take expected values, integrating out the latent state variables: volatilities, jumps and jump intensities. Doing so, we obtain expressions that depend only upon the observable state variables: asset returns. These expressions, however, contain all the parameters of the model, including those related to the latent state variables.

To compute the conditional moments, we use the explicit expression of the generator of the Markov process (2.4) or (2.5)-(2.6) driven by the processes (W, N, λ). Specifically, the computation of the moment functions reduces to the evaluations of conditional expectations of functions of the form $f(y_1, y_0, \delta)$ where y_1 and y_0 denote log-returns separated by some function of the sampling interval δ. We assume that sampling is equidistant in time. We need to evaluate expressions of the form $\mathbb{E}_{Y_1,Y_0}\left[f(Y_1, Y_0, \Delta)\right]$. The dynamics of the system depends upon additional latent variables which at the same instants we denote by $\xi_1 = (v_1, l_1)$ and $\xi_0 = (v_0, l_0)$, respectively, where v is the volatility and l the jump intensity. The standard infinitesimal generator $\mathcal{A}$ is the operator which, when applied to a function $g(y_1, y_0, \xi_1, \xi_0, \delta)$ in its domain, returns the function $\mathcal{A} \cdot g$ given in full generality by

$$\begin{aligned} \mathcal{A} \cdot g = & \frac{\partial g}{\partial \delta} + \sum_{i=1}^{m} \mu_i^X \frac{\partial g}{\partial y_{1,i}} + \sum_{i=1}^{m} \mu_i^V \frac{\partial g}{\partial v_{1,i}} + \sum_{i=1}^{m} \mu_i^\lambda \frac{\partial g}{\partial l_{1,i}} \\ & + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \sigma_{ij}^{YY} \frac{\partial^2 g}{\partial y_{1,i}\partial y_{1,j}} + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \sigma_{ij}^{YV} \frac{\partial^2 g}{\partial y_{1,i}\partial v_{1,j}} + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \sigma_{ij}^{VV} \frac{\partial^2 g}{\partial v_{1,i}\partial v_{1,j}} \\ & + \sum_{i=1}^{m} l_{1,i} \int_{z_i} \left\{ g(y_1 + z_i, y_0, v_1, l_1 + \beta_i, v_0, l_0, \delta) - g(y_1, y_0, v_1, l_1, v_0, l_0, \delta) \right\} f_{Z_i}(z_i) dz_i, \end{aligned} \tag{3.2}$$

where μ^X is the vector of expected log-returns, μ^V the vector of stochastic volatility drifts, μ^λ the vector of jump intensity drift, σ^{YY} the variance-covariance matrix of log-returns, σ^{YV} the variance-covariance matrix of interactions between returns and volatilities, σ^{VV} the variance-covariance matrix of stochastic volatilities and finally $\beta_i = [\beta_{i,1}, ..., \beta_{i,m}]'$ the vector of excitation parameters for the i^{th} jump term. In the case of the mean-reverting jump intensity model described in Section 2.3, we have $\mu_i^{\lambda} = \alpha_i \left(\lambda_{i,\infty} - l_{1,i}\right)$.

The usefulness of the infinitesimal generator for our purpose lies in the fact that

$$\begin{aligned}\mathbb{E}_{Y_1,\xi_1}\left[g(Y_1, Y_0, \xi_1, \xi_0, \Delta)|Y_0, \xi_0\right] &= \exp\left(\Delta\mathcal{A}\right) \cdot g(Y_0, Y_0, \xi_0, \xi_0, 0) \\ &= \sum_{j=0}^{J} \frac{\Delta^j}{j!}\left(\mathcal{A}^j \cdot g\right)(Y_0, Y_0, \xi_0, \xi_0, 0) + O_p\left(\Delta^{J+1}\right), \qquad (3.3)\end{aligned}$$

where subscripts in $\mathbb{E}_{Y_1,\xi_1}$ indicate the random variables that the expected value operates on, and $\mathcal{A}^j \cdot g$ is defined recursively by $\mathcal{A}^j \cdot g = \mathcal{A} \cdot (\mathcal{A}^{j-1} \cdot g)$ for all $j \geq 1$. As we will see in the theorems that follow, (3.2) and its iterates $\mathcal{A}^j \cdot g$, hence the terms in (3.3), can be evaluated in closed-form for the moment functions of interest given in (3.1). While the starting moment function $f(y_1, y_0, \delta)$ does not depend directly upon the latent variables (ξ_1, ξ_0), $\mathcal{A} \cdot f$ and successive iterates will, since the coefficients $\{\mu^X, \mu^V, \mu^{\lambda}, \sigma^{YY}, \sigma^{YV}, \sigma^{VV}, \beta_i\}$ will in general depend on the full state variables $[X, V, \lambda]$.

In other words, we can obtain the conditional expectation of g, using the full state vector including its unobservable components, $\mathbb{E}_{Y_1,\xi_1}\left[g(Y_1, Y_0, \xi_1, \xi_0, \Delta)|Y_0, \xi_0\right]$. We then need to "condition down" by integrating out the unobservable state variables in order to produce moment functions that can be fitted to the log-returns data. All the expectations are taken with respect to the law of the process at the true parameter values. From the law of iterated expectations, we have that

$$\begin{aligned}\mathbb{E}_{Y_1,Y_0,\xi_1,\xi_0}\left[g(Y_1, Y_0, \xi_1, \xi_0, \Delta)\right] &= \mathbb{E}_{Y_1,Y_0,\xi_1,\xi_0}\left[\mathbb{E}_{Y_1,\xi_1}\left[g(Y_1, Y_0, \xi_1, \xi_0, \Delta)|Y_0, \xi_0\right]\right] \\ &= \sum_{j=0}^{J} \frac{\Delta^j}{j!}\mathbb{E}_{Y_0,\xi_0}\left[\left(\mathcal{A}^j \cdot g\right)(Y_0, Y_0, \xi_0, \xi_0, 0)\right] + O_p\left(\Delta^{J+1}\right), \qquad (3.4)\end{aligned}$$

so the last step in the necessary calculations will involve computing unconditional expectations with respect to the stationary law of the state variables.

3.2. The Univariate Case

For ease of exposition, we first provide the expressions of these moments in the univariate case $m = 1$. In this situation, we are looking at a single asset with stochastic volatility and jumps that self-excite, meaning that future jump intensities depend upon the history of their own past jumps:

$$\begin{cases} dX_t = \mu dt + \sqrt{V_t} dW_t^X + Z_t dN_t \\ dV_t = \kappa(\theta - V_t)dt + \eta\sqrt{V_t} dW_t^V \\ d\lambda_t = \alpha\left(\lambda_\infty - \lambda_t\right)dt + \beta dN_t \end{cases} \qquad (3.5)$$

with $\mathbb{E}\left[dW_t^X dW_t^V\right] =: \rho^V dt$ and $\lambda := \mathbb{E}\left[\lambda_t\right] = \alpha\lambda_\infty/(\alpha - \beta)$. Note that classical compound Poisson process jumps are obtained when $\beta = 0$ and $\lambda_0 = \lambda_\infty$. Then $\lambda_t = \lambda_\infty = \lambda$ at all t.

At this stage, we can leave the distribution of the jump size essentially unrestricted, and provide expressions as functions of the moments of the jump size Z_t, for which we write $M[Z,k] := \mathbb{E}\left[Z_t^k\right]$. Our first main result states that:

Theorem 1. *For the univariate model (3.5), the moments are given in closed-form up to order* Δ^2 *by the following expressions*

$$\begin{aligned}
\mathbb{E}\left[\Delta X_t\right] &= (\mu + \lambda M\left[Z,1\right])\Delta + o(\Delta^2) \\
\mathbb{E}\left[(\Delta X_t - \mathbb{E}\left[\Delta X_t\right])^2\right] &= (\theta + \lambda M\left[Z,2\right])\Delta + \frac{\beta\lambda\left(2\alpha-\beta\right)}{2(\alpha-\beta)} M\left[Z,1\right]^2 \Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta X_t - \mathbb{E}\left[\Delta X_t\right])^3\right] &= \lambda M\left[Z,3\right]\Delta \\
&\quad + \frac{3}{2}\left(\eta\theta\rho^V + \frac{(2\alpha-\beta)\beta\lambda M\left[Z,1\right] M\left[Z,2\right]}{(\alpha-\beta)}\right)\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta X_t - \mathbb{E}\left[\Delta X_t\right])^4\right] &= \lambda M\left[Z,4\right]\Delta + \left(\frac{3\theta\eta^2}{2\kappa} + 3\theta^2 + 6\theta\lambda M\left[Z,2\right]\right. \\
&\quad + 3\lambda\left(\lambda + \frac{(2\alpha-\beta)\beta}{2(\alpha-\beta)}\right) M\left[Z,2\right]^2 \\
&\quad \left. + \frac{2(2\alpha-\beta)\beta\lambda M\left[Z,1\right] M\left[Z,3\right]}{(\alpha-\beta)}\right)\Delta^2 + o(\Delta^2)
\end{aligned}$$

while the autocorrelation function of the process is given by

$$\mathbb{E}\left[(\Delta X_t - \mathbb{E}\left[\Delta X_t\right])(\Delta X_{t+\tau} - \mathbb{E}\left[\Delta X_{t+\tau}\right])\right] = \frac{\beta\lambda(2\alpha-\beta)}{2(\alpha-\beta)} e^{-(\alpha-\beta)\tau} M\left[Z,1\right]^2 \Delta^2 + o(\Delta^2)$$

for all $\tau > 0$.

Intuitively, the identification of the parameters is achieved as follows: the higher order moments (3 and 4) isolate the jump parameters at the leading order, while the variance puts them on an equal footing with the diffusive parameters.[10] The autocovariance isolates the jump parameters. Indeed, if the model had no jump component, then from (3.5) with $Z_t \equiv 0$, the law of iterated expectations implies that

$$\begin{aligned}
\mathbb{E}\left[(\Delta X_t)\left(\Delta X_{t+\tau}\right)\right] &= \mathbb{E}\left[(\Delta X_t)\,\mathbb{E}\left[(\Delta X_{t+\tau})\,|\mathcal{F}_{t+\tau}\right]\right] = \mathbb{E}\left[(\Delta X_t)\left(\mu\Delta\right)\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[(\Delta X_t)\,|\mathcal{F}_t\right](\mu\Delta)\right] = \mu^2\Delta^2
\end{aligned}$$

and so $\mathbb{E}\left[(\Delta X_t - \mathbb{E}\left[\Delta X_t\right])(\Delta X_{t+\tau} - \mathbb{E}\left[\Delta X_{t+\tau}\right])\right] = 0$.

Thus any autocovariance is due to the jump component. Further, if the jump component is Poissonian, then the increments would be independent and it is the self-excitation of the jumps

[10] We use regular moments as opposed to absolute moments. The use of absolute moments –especially absolute moments of order less than 1 (see Aït-Sahalia (2004))– may be considered in addition, especially for estimating parameters of the volatility process. Since we are mainly interested in the parameters of the jump process we have chosen to consider only regular moments in the interest of simplicity and parsimony.

that gives rise to the autocorrelation. Thus the observed autocovariance of the increments isolates the self-exciting component of the model.

As noted above, the model reduces to a Poissonian jump-diffusion if $\beta = 0$. Then we indeed have

$$\begin{aligned}
\mathbb{E}\left[\Delta X_t\right] &= (\mu + \lambda M\left[Z,1\right])\Delta + o(\Delta^2) \\
\mathbb{E}\left[(\Delta X_t - \mathbb{E}\left[\Delta X_t\right])^2\right] &= (\theta + \lambda M\left[Z,2\right])\Delta + o(\Delta^2) \\
\mathbb{E}\left[(\Delta X_t - \mathbb{E}\left[\Delta X_t\right])^3\right] &= \lambda M\left[Z,3\right]\Delta + \frac{3}{2}\eta\theta\rho^V\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta X_t - \mathbb{E}\left[\Delta X_t\right])^4\right] &= \lambda M\left[Z,4\right]\Delta + 3\left(\frac{\theta\eta^2}{2\kappa} + (\theta + \lambda M\left[Z,2\right])^2\right)\Delta^2 + o(\Delta^2)
\end{aligned}$$

and the model does not generate any autocorrelation.

It is clear from the explicit expressions of these moments that there will be difficulty in practice identifying the stochastic volatility parameters: there are simply too many latent processes: V_t, N_t, λ_t; while those are integrated out as part of the development of the unconditional moments, their corresponding parameters remain. So they are theoretically identified, but this identification can be tenuous in practice. This is an unavoidable consequence of the unobservability of those state variables. This will be especially so in the multivariate context where there are many potential correlation coefficients: between each asset's log-return and all the volatilities, among the different asset log-returns and among the different volatilities. To avoid this problem, we will restrict some of the parameters to be identical, and identify only the level of the volatility.

3.3. The Bivariate Case

We now turn to the bivariate case, $m = 2$. We consider the general case in the Appendix, but restrict attention here to the more tractable triangular excitation case, where $\beta_{12} = 0$, with state-independent volatilities. That is, the base model is

$$\begin{cases}
dX_{1t} = \mu_1 dt + \sigma_1 dW_t^1 + Z_{1t} dN_{1t} \\
dX_{2t} = \mu_2 dt + \sigma_2 dW_t^2 + Z_{2t} dN_{2t} \\
d\lambda_{1t} = \alpha_1\left(\lambda_{1\infty} - \lambda_{1t}\right)dt + \beta_{11} dN_{1t} \\
d\lambda_{2t} = \alpha_2\left(\lambda_{2\infty} - \lambda_{2t}\right)dt + \beta_{21} dN_{1t} + \beta_{22} dN_{2t}
\end{cases} \tag{3.6}$$

with $\mathbb{E}[dW_t^1 dW_t^2] =: \rho dt$.

Theorem 2. *For the bivariate model (3.6), the mean and variance of the log-returns are given in closed-form up to order Δ^2 by the following expressions*

$$\begin{aligned}
\mathbb{E}\left[\Delta X_{1t}\right] &= (\mu_1 + \lambda_1 M[Z_1, 1])\Delta + o(\Delta^2) \\
\mathbb{E}\left[\Delta X_{2t}\right] &= (\mu_2 + \lambda_2 M[Z_2, 1])\Delta + o(\Delta^2)
\end{aligned}$$

and

$$
\begin{aligned}
\mathbb{E}\left[(\Delta X_{1t}-\mathbb{E}\left[\Delta X_{1t}\right])^2\right] &= \left(\sigma_1^2+\lambda_1 M[Z_1,2]\right)\Delta+\tfrac{(2\alpha_1-\beta_{11})\beta_{11}\lambda_1}{2(\alpha_1-\beta_{11})}M[Z_1,1]^2\Delta^2+o(\Delta^2)\\
\mathbb{E}\left[(\Delta X_{2t}-\mathbb{E}\left[\Delta X_{2t}\right])^2\right] &= (\sigma_2^2+\lambda_2 M\left[Z_2,2\right])\Delta+\Bigg(\tfrac{\beta_{21}^2\left(\alpha_1^2+(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})\right)\lambda_1}{2(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})}\\
&\quad+\tfrac{(2\alpha_2-\beta_{22})\beta_{22}\lambda_2}{2(\alpha_2-\beta_{22})}\Bigg)M[Z_2,1]^2\Delta^2+o\left(\Delta^2\right)\\
\mathbb{E}\left[(\Delta X_{1t}-\mathbb{E}\left[\Delta X_{1t}\right])(\Delta X_{2t}-\mathbb{E}\left[\Delta X_{2t}\right])\right] &= \rho\sigma_1\sigma_2\Delta\\
&\quad+\tfrac{\beta_{21}\left(\alpha_1^2+(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})\right)\lambda_1}{2(\alpha_1-\beta_{11})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})}M\left[Z_1,1\right]M\left[Z_2,1\right]\Delta^2+o\left(\Delta^2\right).
\end{aligned}
$$

As expected, the variance places the contributions from the diffusive and jump components of the model on the same level. However, the leading terms of the higher order moments isolate the jump component:

Theorem 3. *The higher order moments of the bivariate model are given up to order* Δ^2 *by the following expressions:*

$$
\begin{aligned}
\mathbb{E}\left[(\Delta X_{1t}-\mathbb{E}\left[\Delta X_{1t}\right])^3\right] &= \lambda_1 M[Z_1,3]\Delta+\tfrac{3(2\alpha_1-\beta_{11})\beta_{11}\lambda_1}{2(\alpha_1-\beta_{11})}M\left[Z_1,1\right]M\left[Z_1,2\right]\Delta^2+o(\Delta^2)\\
\mathbb{E}\left[(\Delta X_{2t}-\mathbb{E}\left[\Delta X_{2t}\right])^3\right] &= \lambda_2 M[Z_2,3]\Delta+\tfrac{3}{2}\Bigg(\tfrac{(2\alpha_2-\beta_{22})\beta_{22}\lambda_2}{(\alpha_2-\beta_{22})}\\
&\quad+\tfrac{\left(\alpha_1^2+(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})\right)\lambda_1\beta_{21}^2}{(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})}\Bigg)M\left[Z_2,1\right]M\left[Z_2,2\right]\Delta^2+o(\Delta^2)\\
\mathbb{E}\left[(\Delta X_{1t}-\mathbb{E}\left[\Delta X_{1t}\right])^2(\Delta X_{2t}-\mathbb{E}\left[\Delta X_{2t}\right])\right] &= \tfrac{\beta_{21}\left(\alpha_1^2+(\alpha_2-\beta_{22})(\alpha_1-\beta_{11})\right)\lambda_1}{2(\alpha_1-\beta_{11})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})}\\
&\quad\times M\left[Z_1,2\right]M\left[Z_2,1\right]\Delta^2+o(\Delta^2)\\
\mathbb{E}\left[(\Delta X_{1t}-\mathbb{E}\left[\Delta X_{1t}\right])(\Delta X_{2t}-\mathbb{E}\left[\Delta X_{2t}\right])^2\right] &= \tfrac{\beta_{21}\left(\alpha_1^2+(\alpha_2-\beta_{22})(\alpha_1-\beta_{11})\right)\lambda_1}{2(\alpha_1-\beta_{11})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})}\\
&\quad\times M\left[Z_1,1\right]M\left[Z_2,2\right]\Delta^2+o(\Delta^2)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\left[(\Delta X_{1t}-\mathbb{E}\left[\Delta X_{1t}\right])^4\right] &= \lambda_1 M[Z_1,4]\Delta+\Big(3\sigma_1^4+6M\left[Z_1,2\right]\lambda_1\sigma_1^2\\
&\quad+\tfrac{3M[Z_1,2]^2\lambda_1(2\alpha_1(\beta_{11}+\lambda_1)-\beta_{11}(\beta_{11}+2\lambda_1))}{2(\alpha_1-\beta_{11})}+\tfrac{2(2\alpha_1-\beta_{11})\beta_{11}\lambda_1}{(\alpha_1-\beta_{11})}M\left[Z_1,1\right]M\left[Z_1,3\right]\Big)\Delta^2+o(\Delta^2)\\
\mathbb{E}\left[(\Delta X_{2t}-\mathbb{E}\left[\Delta X_{2t}\right])^4\right] &= \lambda_2 M[Z_2,4]\Delta+\Big(3\sigma_1^4+6M\left[Z_2,2\right]\lambda_2\sigma_2^2\\
&\quad+3M\left[Z_2,2\right]^2\left(\eta_{22}+\beta_{22}\lambda_2\right)+4\left(\eta_{22}+\left(\beta_{22}-\lambda_2\right)\lambda_2\right)M\left[Z_2,1\right]M\left[Z_2,3\right]\Big)\Delta^2+o(\Delta^2).
\end{aligned}
$$

The other contemporaneous cross-terms that are not listed above,

$$
\mathbb{E}\left[(\Delta X_{1t}-\mathbb{E}\left[\Delta X_{1t}\right])^j(\Delta X_{2t}-\mathbb{E}\left[\Delta X_{2t}\right])^k\right],
$$

of order $j+k=4$, are of smaller order $O(\Delta^2)$.

The expressions above depend upon expected values and variances of the jump intensities, which are given by

$$
\begin{aligned}
\lambda_1 &:= \mathbb{E}\left[\lambda_{1t}\right] = \tfrac{\alpha_1\lambda_{1,\infty}}{\alpha_1-\beta_{11}}\\
\lambda_2 &:= \mathbb{E}\left[\lambda_{2t}\right] = \tfrac{\alpha_1\beta_{21}\lambda_{1,\infty}+\alpha_1\alpha_2\lambda_{2,\infty}-\alpha_2\beta_{11}\lambda_{2,\infty}}{(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})}
\end{aligned}
$$

and

$$
\begin{aligned}
\eta_{11} &:= \mathbb{E}\left[\lambda_{1t}^2\right] = \lambda_1^2+\tfrac{\beta_{11}^2\lambda_1}{2(\alpha_1-\beta_{11})}\\
\eta_{22} &:= \mathbb{E}\left[\lambda_{2t}^2\right] = \lambda_2^2+\tfrac{\beta_{22}^2\lambda_2}{2(\alpha_2-\beta_{22})}+\tfrac{\left(\alpha_1^2+(\alpha_2-\beta_{22})\alpha_1+\beta_{11}(\beta_{22}-\alpha_2)\right)\lambda_1\beta_{21}^2}{2(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})}\\
\eta_{12} &:= \mathbb{E}\left[\lambda_{1t}\lambda_{2t}\right] = \lambda_2\lambda_1+\tfrac{(2\alpha_1-\beta_{11})\beta_{11}\beta_{21}\lambda_1}{2(\alpha_1-\beta_{11})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})}.
\end{aligned}
$$

The final result provides the autocovariance function of the bivariate model:

Theorem 4. *For the bivariate model (3.5), the autocovariances of the log-returns at lag $\tau > 0$,*

$$C_{i,j}(\tau) := \mathbb{E}\left[(\Delta X_{j,t} - \mathbb{E}\left[\Delta X_{j,t}\right])(\Delta X_{i,t+\tau} - \mathbb{E}\left[\Delta X_{i,t+\tau}\right])\right],$$

are given in closed-form up to order Δ^2 by the following expressions, which are driven by the mutually exciting component of the model:

$$\begin{aligned}
\mathbb{C}_{1,1}(\tau) &= e^{-\tau(\alpha_1-\beta_{11})} a_{11,1} M\left[Z_1,1\right]^2 \Delta^2 + o(\Delta^2)\\
\mathbb{C}_{1,2}(\tau) &= e^{-\tau(\alpha_1-\beta_{11})} a_{12,1} M\left[Z_1,1\right] M\left[Z_2,1\right] \Delta^2 + o(\Delta^2)\\
\mathbb{C}_{2,1}(\tau) &= \left(e^{-\tau(\alpha_1-\beta_{11})} a_{21,1} + e^{-\tau(\alpha_2-\beta_{22})} a_{21,2}\right) M\left[Z_1,1\right] M\left[Z_2,1\right] \Delta^2 + o(\Delta^2)\\
\mathbb{C}_{2,2}(\tau) &= \left(e^{-\tau(\alpha_1-\beta_{11})} a_{22,1} + e^{-\tau(\alpha_2-\beta_{22})} a_{22,2}\right) M\left[Z_2,1\right]^2 \Delta^2 + o(\Delta^2)
\end{aligned}$$

where

$$\begin{aligned}
a_{11,1} &= \frac{(2\alpha_1-\beta_{11})\beta_{11}\lambda_1}{2(\alpha_1-\beta_{11})}\\
a_{12,1} &= \frac{(2\alpha_1-\beta_{11})\beta_{11}\beta_{21}\lambda_1}{2(\alpha_1-\beta_{11})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})}\\
a_{21,1} &= \frac{(2\alpha_1-\beta_{11})\beta_{11}\beta_{21}\lambda_1}{2(\alpha_1-\beta_{11})(-\alpha_1+\alpha_2+\beta_{11}-\beta_{22})}\\
a_{21,2} &= \frac{\beta_{21}\left(\alpha_1^2-(\alpha_2-\beta_{22})^2\right)\lambda_1}{(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})(\alpha_1-\alpha_2-\beta_{11}+\beta_{22})}\\
a_{22,1} &= \frac{(2\alpha_1-\beta_{11})\beta_{11}\lambda_1\beta_{21}^2}{2(\alpha_1-\beta_{11})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})(-\alpha_1+\alpha_2+\beta_{11}-\beta_{22})}\\
a_{22,2} &= \frac{(2\alpha_2-\beta_{22})\beta_{22}\lambda_2}{2(\alpha_2-\beta_{22})} + \frac{\left(\alpha_1^2-(\alpha_2-\beta_{22})^2\right)\lambda_1\beta_{21}^2}{2(\alpha_2-\beta_{22})(\alpha_1+\alpha_2-\beta_{11}-\beta_{22})(\alpha_1-\alpha_2-\beta_{11}+\beta_{22})}.
\end{aligned}$$

As in the univariate case, the diffusive component of the model generates no autocorrelation. Also, the autocorrelation structure is asymmetric, reflecting the fact that the excitation from jumps in sector 1 to jumps in sector 2 is not the same as that from jumps in sector 2 to jumps in sector 1.

3.4. Interval-Based Moment Functions

GMM can be carried out using the instantaneous moments above, on short time intervals (daily). However, in order to improve accuracy and reduce the number of moment conditions to be used, we calculate in addition the unconditional interval-based moments ($i, j = 1, 2; 0 \le s_1 < s_2 < s_3 < s_4$)

$$\begin{cases}
\mathbb{E}\left[\int_{s_1}^{s_2} dX_{i,t}\right]\\
\mathbb{E}\left[\left(\int_{s_1}^{s_2} dX_{i,t} - \mathbb{E}\left[\int_{s_1}^{s_2} dX_{i,t}\right]\right)^2\right]\\
\mathbb{E}\left[\int_{s_1}^{s_2} dX_{i,t} \int_{s_1}^{s_2} dX_{j,u} - \mathbb{E}\left[\int_{s_1}^{s_2} dX_{i,t}\right]\mathbb{E}\left[\int_{s_1}^{s_2} dX_{j,u}\right]\right], \qquad i \neq j\\
\mathbb{E}\left[\int_{s_3}^{s_4} dX_{i,u} \int_{s_1}^{s_2} dX_{j,t} - \mathbb{E}\left[\int_{s_3}^{s_4} dX_{i,u}\right]\mathbb{E}\left[\int_{s_1}^{s_2} dX_{j,t}\right]\right]
\end{cases}$$

which can be applied on time intervals of arbitrary lengths. We are also able to compute these integrated moment conditions in closed-form.

The explicit formulae we derive for the moment conditions are contained in the Appendix. The (lengthy) closed-form expressions for the integrals $I^{\tau}_{i,j,N}$ and $I_{i,j,N}$ defined in the Appendix as well as for the derivatives of the moment conditions (needed to compute Ω; see (3.11) below) are not displayed in this paper to save space. These expressions are available in computer form from the authors upon request.

3.5. Array of Autocovariances and Cross-Covariances

To implement GMM with these interval-based moment conditions, a choice has to be made on the "array" of autocovariances and cross-covariances (i.e., the number of autocovariances and cross-covariances, and the length of the intervals of integration) to be used. Autocorrelograms and cross-correlograms of the data can be used as exploratory tools for determining the total length of the time period on which autocovariances and cross-covariances are to be considered.

Finding the "optimal" array of autocovariances and cross-covariances, striking a balance between accuracy and computational feasibility, has been one of our objectives in an extensive Monte Carlo study, which we will present below. Also, in case the sampling interval Δ that we use to compute the first four moments and the instantaneous autocovariances and cross-covariances is different from the sampling interval Δ that we use to compute (lagged) autocovariances and cross-covariances, we use a varying number of observations for different elements in the GMM procedure and as such the scalar N needs to be replaced by a vector in the relevant places.

3.6. GMM Estimation

Given the expressions for the moment functions above, we proceed to estimate the parameters of the model using standard GMM. Let $Y_{n\Delta} := X_{n\Delta} - X_{(n-1)\Delta}$, $n = 1, \ldots, N$, $\Delta > 0$, with $N\Delta = T$, denote the log-returns on the interval $[0, T]$ and let $\theta \in \Theta$ denote our d-dimensional parameter vector. To estimate θ we consider a vector of M moment conditions $h(y, \Delta, \theta)$, $M \geq d$, continuously differentiable in θ. Let θ_0 denote the true value of θ and suppose that $\mathbb{E}\left[h(Y_{n\Delta}, \Delta, \theta_0)\right] = 0$. This is the key requirement for consistency of the GMM estimator. The closed-form expressions for the moments derived above ensure that it will be satisfied, since we will use for each component of the vector h the difference between the corresponding sample moment of the log-returns and its closed-form expression derived under the model.

We assume that the log-returns are stationary, which is the case for our model under the parameter restrictions discussed above. Let $g_T(y, \Delta, \theta)$ denote the sample average of $h(y_{n\Delta}, \Delta, \theta)$, that is $g_T(y, \Delta, \theta) := \frac{1}{N} \sum_{n=1}^{N} h(y_{n\Delta}, \Delta, \theta)$. Then the GMM estimator $\hat{\theta}_T$ is the value of $\theta \in \Theta$ that minimizes the quadratic form

$$g_T(y, \Delta, \theta)' \, W_T \, g_T(y, \Delta, \theta), \tag{3.7}$$

where W_T is an $M \times M$ positive definite weight matrix assumed to converge in probability to a positive definite limit W. The system is exactly identified if $M = d$, in which case the choice of W_T becomes irrelevant.

The weight matrix W_T can be chosen optimally in the following way: Suppose that the process

$\{h(Y_{n\Delta}, \Delta, \theta_0)\}_{n=-\infty}^{\infty}$ is strictly stationary with mean zero and v-th autocovariance matrix $\Gamma_v :=$ $\mathbb{E}\left[h(Y_{n\Delta}, \Delta, \theta_0)\ h(Y_{(n-v)\Delta}, \Delta, \theta_0)'\right]$. Assuming that these autocovariances are absolutely summable (i.e., that the sequence of partial absolute sums has a finite limit), let $S := \sum_{v=-\infty}^{\infty} \Gamma_v$. We note that S is the asymptotic variance of the sample average of $h(y_{n\Delta}, \Delta, \theta_0)$:

$$S = \lim_{N\to\infty} N\ \mathbb{E}\left[g_{T_N}(Y, \Delta, \theta_0)\ g_{T_N}(Y, \Delta, \theta_0)'\right]. \tag{3.8}$$

The optimal weight matrix turns out to be S^{-1}. A non-negative definite estimator of S is

$$\hat{S}_T = \hat{\Gamma}_{0,T} + \sum\nolimits_{v=1}^{q} \left(1 - \frac{v}{q+1}\right)\left(\hat{\Gamma}_{v,T} + \hat{\Gamma}'_{v,T}\right), \tag{3.9}$$

with

$$\hat{\Gamma}_{v,T} = \frac{1}{N} \sum\nolimits_{n=v+1}^{N} h(y_{n\Delta}, \Delta, \tilde{\theta})\ h(y_{(n-v)\Delta}, \Delta, \tilde{\theta})'. \tag{3.10}$$

Here $\tilde{\theta}$ is an initial consistent estimate of θ_0, which can be obtained by minimizing (3.7) with $W_T = I$.

Under standard regularity conditions (see Hansen (1982)), $\sqrt{T_N}(\hat{\theta} - \theta_0)$ converges in law to $N(0, \Omega)$ where

$$\Omega^{-1} := \Delta^{-1} D' W D (D' W S W D)^{-1} D' W D, \tag{3.11}$$

and D is the gradient of $E\left[h(Y_{n\Delta}, \Delta, \theta)\right]$ with respect to θ' evaluated at θ'_0. When W_T is chosen optimally, $W = S^{-1}$ and (3.11) reduces to $\Omega^{-1} = \Delta^{-1} D' S^{-1} D$.

3.7. Testing for the Presence of Contagion

If desired, it is possible to test for the presence of contagion in the context of our model. Whether or not contagion occurs boils down here to testing the joint hypothesis that all the coefficients of mutual excitation $\beta_{i,j}$'s are 0. We can further separate between testing for self- or time-series contagion: diagonal $\beta_{i,i} = 0$ and testing for cross-sectional contagion: off-diagonal $\beta_{i,j} = 0$, $i \neq j$. Since the inference is based on standard GMM given the relevant moment functions, GMM-based testing tools apply to test these parameter restrictions. We will employ below for this purpose the χ^2 statistic that follows from (3.11).

4. Finite Sample Behavior: A Monte Carlo Study

One major advantage of the proposed estimation method is that it is numerically tractable, so that large numbers of Monte Carlo simulations can be conducted to determine the finite sample distribution of the estimators. From the onset, one should be aware of the fact that one cannot expect degrees of accuracy similar or even close to what we are used to when estimating, e.g., a standard stochastic volatility model. By definition, extreme events occur infrequently and there is a positive probability that no jump occurs in any given finite time interval. When no jump occurs,

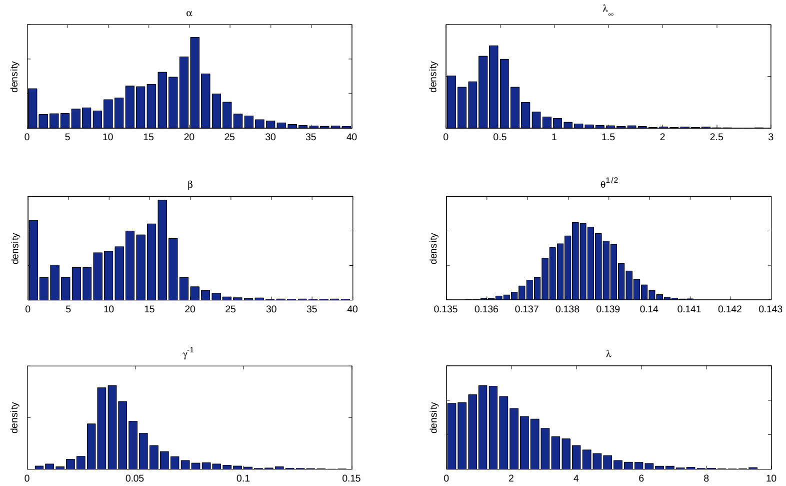

Figure 5: Monte Carlo Results: Univariate Case. This figure plots the empirical distribution functions of the parameter estimators for the univariate Hawkes jump-diffusion model, obtained from 5,000 simulated sample paths.

there is no identification whatsoever, while when few jumps occur identification is expected to be weak. This is a consequence of the classical peso problem. Furthermore, we are effectively after subtle effects concerning the jumps of the process, not just whether they are present or not, but their finer structure: whether they self-excite, whether they cross-excite, etc. Finally, the estimation procedure can rely only on the time series of the log-returns of the various assets, since all other state variables are latent (volatility, jumps, jump intensities.) In other words, we are asking for a lot out of the data. So one should not expect miracles from a time series of necessarily finite length.

Nevertheless, what emerges out of the Monte Carlos is evidence that, using the estimation methodology outlined above, the parameters of the data generating process can be recovered with sufficient degree of precision in a realistic context.

Before taking our model to financial data, we study by means of extensive Monte Carlo simulations the "optimal" array of autocovariances and cross-covariances, and the corresponding degree of accuracy of our estimation method for various sets of (fictitious) parameter values.

In general, we find that for daily data, the optimal array of moment conditions, striking a balance between accuracy and computational feasibility, is to use daily means and instantaneous (co)variances and to use daily, weekly or monthly autocovariances and cross-covariances depending on the degree of excitation. We find that using autocovariances and cross-covariances between periods of the same length rather than of different lengths produces the most stable results. Visual inspection of the autocorrelograms and cross-correlograms is used to determine the number of autocovariances and cross-covariances we need to include in our GMM estimation. In the bivariate case, the restrictions that $\alpha_1 = \alpha_2$ and $\lambda_{1,\infty} = \lambda_{2,\infty}$ appear to be necessary to have a good degree

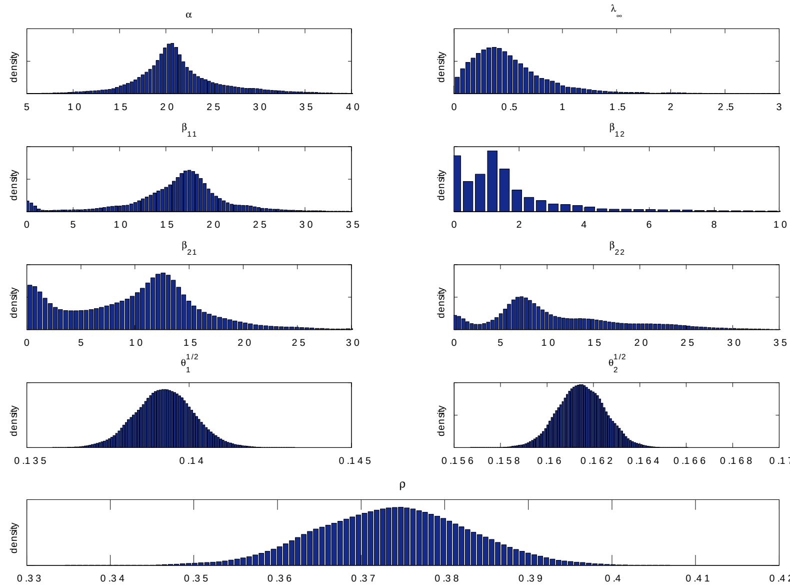

Figure 6: Monte Carlo Results: Bivariate Case. This figure plots the empirical distribution functions of the parameter estimators for the bivariate Hawkes jump-diffusion model, obtained from 5,000 simulated sample paths.

of identification.

To facilitate disentangling diffusion from jumps we adopt a two stage procedure: we first estimate the parameters from the continuous part of our model using truncated data. We suppose that truncation removes (most of) the jumps so that the truncated data can be viewed as being generated by the continuous part of our model. We fit the truncated data using moment conditions that are derived from the continuous part of our model, ignoring the discontinuous jump part. Then in a second stage we treat the obtained parameter estimates for the continuous part of our model as fixed and given and identify the parameters of the Hawkes process using the full set of moment conditions. With these types of moment conditions the estimates that we obtain from Monte Carlo simulated data are usually quite accurate. Due to the truncation, the volatility and correlation parameters are slightly biased downwards (but we will correct for this in our empirical analysis contained in Section 5).

The small selection of Monte Carlo results that we present below is obtained in a setting that mimics that of our empirical analysis described in Section 5. In particular, we impose similar parameter restrictions, use a similar sample period and use parameter values similar to the parameter estimates obtained from real data.

We first consider the univariate version of the Hawkes jump-diffusion model. Figure 5 plots the small sample distribution of our estimators for the univariate Hawkes jump-diffusion process parameters obtained from 5,000 simulated sample paths. The true parameter values are $\alpha = 20.3$, $\beta = 17.1$, $\lambda_\infty = 0.40$, $\sqrt{\theta} = 0.14$, $1/\gamma = 0.030$, $\mu = 0$ and $p = 1$. We use the interval-based expressions for the first moment and the autocovariances, derived in the Appendix, and use the leading terms of the third and fourth moments, contained in Theorem 1. Because μ cannot be estimated consistently in finite time, we fix it in our GMM estimation and we further fix p.

In the bivariate case, which we are ultimately interested in, due to the cross-covariances (and under the restrictions that $\alpha_1 = \alpha_2 =: \alpha$ and $\lambda_{1,\infty} = \lambda_{2,\infty} =: \lambda_\infty$), we might expect even better identification than in the univariate case. Figure 6 plots the small sample distribution of our estimators for the bivariate Hawkes jump-diffusion process parameters obtained from 5,000 simulated sample paths. The true parameter values are $\alpha = 20.3$, $\beta_{11} = 17.1$, $\beta_{12} = 1.2$, $\beta_{21} = 13.1$, $\beta_{22} = 7.1$, $\lambda_\infty = 0.40$, $\sqrt{\theta_1} = 0.14$, $\sqrt{\theta_2} = 0.17$, $\rho = 0.39$, $1/\gamma_1 = 0.030$, $1/\gamma_2 = 0.027$, $\mu_1 = 0.21$, $\mu_2 = 0.20$ and $p_1 = p_2 = 1$. Because on a fixed time horizon μ_1 and μ_2 can never be estimated consistently, we fix these parameters in our GMM estimation. Also, we fix $1/\gamma_1$, $1/\gamma_2$ and p_1, p_2 so that we need not consider the (approximated) third and fourth moments for identification, and rely solely on the interval-based moments derived in the Appendix. Doing so we can focus on the identification of the Hawkes process parameters, which are our prime interest in this paper.

The Monte Carlo results show that the population parameters of the Hawkes jump-diffusion model can be identified with sufficient degree of precision from data generated by the presupposed Hawkes jump-diffusion model. We can also analyze the situation in which the data generating process is in fact a Poissonian jump-diffusion (with constant jump intensities) but is presupposed to be a Hawkes jump-diffusion (with stochastic jump intensities). We find that our estimation methodology is robust in this respect, finding parameter estimates for $\alpha, \beta_{11}, \beta_{12}, \beta_{21}$ and β_{22} that are close to and statistically not significantly different from zero. With β's that are exactly zero the Hawkes jump-diffusion model reduces to a Poissonian jump-diffusion model.

5. Empirical Analysis

We now take the model to real data. While the same model could equally well be fitted to various sectors of the economy, or different stocks on the same sector, in the following we focus on capturing patterns of contagion between stock indices around the world.

5.1. Data

We use Morgan Stanley Capital International (MSCI) international equity index data. We will study six indices: US; Latin America (LA); UK; Developed countries Europe (EU); Japan (JA); Emerging markets Asia (ASEM). Daily data are available from January 1, 1980 (for US, UK, EU and JA), from January 29, 1988 (for LA) and from January 1, 1988 (for ASEM). Summary statistics are in Table 1. Autocorrelograms and cross-correlograms are in Figure 7.

Panel A	US	UK	EU	JA
Mean	0.000278	0.000235	0.000270	0.000228
St. Deviation	0.011037	0.012360	0.010893	0.014114
Skewness	-1.385245	-0.375330	-0.364734	-0.103119
Excess Kurtosis	32.929671	10.259949	10.936663	8.481328
Panel B	US	LA	ASEM	
Mean	0.000234	0.000527	0.000157	
St. Deviation	0.010958	0.017487	0.013179	
Skewness	-0.280164	-0.479455	-0.309408	
Excess Kurtosis	10.741348	10.572756	7.289829	

Table 1: Descriptive Statistics: MSCI International Equity Indices. This table reports descriptive statistics for the log-returns of daily MSCI international equity index data. Panel A: Sample period: January 1, 1980 to December 31, 2008. Panel B: Sample period US and Latin America: January 29, 1988 to December 31, 2008; sample period Emerging Markets Asia: January 1, 1988 to December 31, 2008.

We observe from Table 1 that for all the regions under study the Excess Kurtosis is substantially larger than that for a Gaussian distribution. Jumps in equity index returns can cause such Excess Kurtosis. Notice the difference in Excess Kurtosis between the US return series of panel A and the US return series of panel B. This difference is mainly due to the 1987 crisis, the returns of which are included in the longer sample period but not in the shorter sample period.

The plots in Figure 7 exhibit substantial correlations between US returns on day j and returns of other regions of the world on the following day $j + 1$, except for Latin America where this phenomenon is not pronounced. It becomes apparent from the plots that autocorrelations and cross-correlations die out quickly, so that in five days time (or even fewer) most correlation has disappeared. As discussed in Section 3, the existence of these short-run correlations is where the identification of the self- and cross-excitation parameters comes from.

5.2. Data Sequencing

The model predicts that the empirical autocorrelation and cross-correlation is decreasing and convex, hence that the instantaneous correlation is the largest. Figure 7 plots the empirical autocorrelations and cross-correlations using the (raw) MSCI data. Due to the fact that different markets operate in different time zones, and the empirical observation that transmission of some markets (US) seems to be stronger than transmission of other markets (non-US) we find a kink in some of the cross-correlations, for example, US-JA.

To get a qualitative insight in the direction of jump transmissions, we sort daily US returns to find the most extreme declines (over 3.0% in a single day) in the US stock market in the period January 1, 1980 to December 31, 2008. If the inter-arrival time of these "jumps" was less than

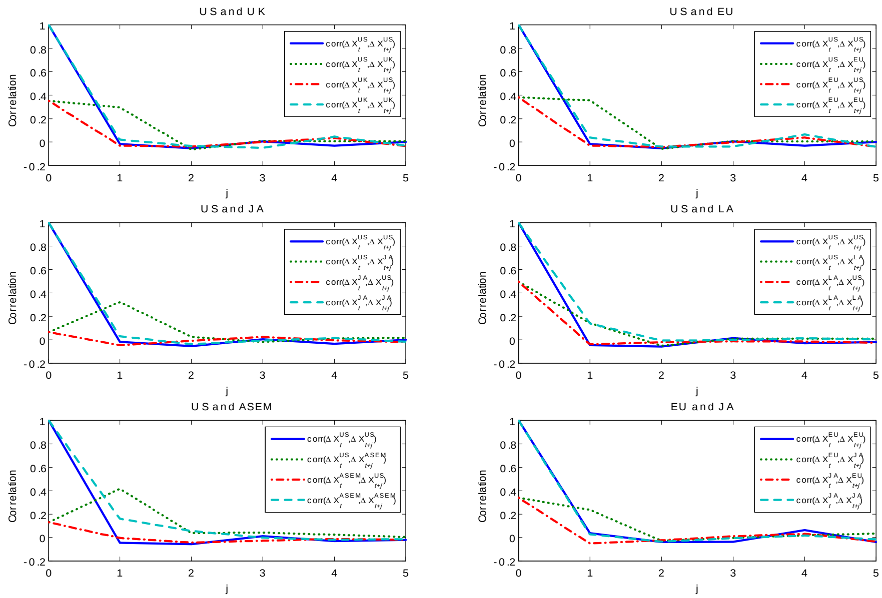

Figure 7: Autocorrelations and Cross-Correlations (Raw Data). This figure plots autocorrelations and cross-correlations for the log-returns of the daily MSCI international equity data of Table 1. The unit of the index j is days.

6 weeks, we grouped the returns as being one event, or a related sequence of events. We end up with 63 declines and 23 groups. We read the analysis in the press in the days following each event (statements such as "Tokyo opened lower *after* Wall Street closed down 3%" vs. "Wall Street opened lower *following* a rout in European markets") to confirm the sequencing: where and when the event started, and whether transmission (contagion) took place following one of these events.

Tables 2, 3 and 4 summarize our findings. Generally speaking, we find that it is either (and primarily) a US news announcement that causes the US market to slump, and often such a decline transmits contagiously, or it is a non-US event, typically an adverse shock in emerging markets, and then other regions of the world decline once the US start declining. There are a few exceptions to this general pattern of transmission. In order to cope with the (institutionally introduced) different time zones and hours of operation of the markets around the world some adjustments need to be made to the raw data. We detail below the adjustments made.

5.2.1. US and Japan, and US and Emerging Markets Asia

To get the sequencing right given the time difference, we calculate the instantaneous correlation by leading the Japanese return series by one day. To calculate the lagged cross-correlations we lead the Japanese return series by one day when computing $\mathrm{corr}(\Delta X_t^{\mathrm{US}}, \Delta X_{t+j}^{\mathrm{JA}})$, $j = 1, 2, \ldots$, and we do not

Group	Date	Event	Starts in	Transmits?
1	October 14, 1987	News on US trade deficit	US	Yes
	October 16, 1987	„	„	„
	October 19, 1987	Black Monday	Hong Kong	Yes
	October 22, 1987	„	„	„
	October 26, 1987	„	„	„
	November 30, 1987	Approaching US recession and accelerating inflation	US	Yes
	December 3, 1987	„	„	„
	January 8, 1988	„	„	„
2	September 9, 2008	Fannie Mae and Freddie Mac placed into conservatorship	US	No
	September 15, 2008	Bankruptcy of Lehman Brothers and sale of Merrill Lynch	US	Yes
	September 17, 2008	„	„	„
	September 22, 2008	Goldman Sachs and Morgan Stanley converted to bank holding companies	US	No
	September 29, 2008	Four bailouts in Europe and US House of Representatives voted against $ 700 billion rescue plan	Europe / US	Yes
	October 2, 2008	Prospects of the $700 billion US dollars bailout became reality after approval on October 1 (Senate) and October 3 (House of Representatives) "The financial markets, however, were not enthusiastic. ... weighed down by another round of bleak economic data, including a report showing that 159,000 jobs were lost in September..." (October 3, 2008; The New York Times)	US	Yes
	October 6, 2008	Major financial crisis in Iceland and several European governments guarantee bank deposits	Europe	Yes
	October 7, 2008	„	„	„
	October 9, 2008	Simultaneous cuts of interest rates by Central Banks	US	Yes
	October 15, 2008	Losses in Europe precipitated further losses in the US	Europe	Yes
	October 21, 2008	Stock markets continued to decline worldwide during the week of October 20, 2008.		
	October 22, 2008	„	„	„
	October 24, 2008	Losses in Europe precipitated further losses in the US. Stock markets declined sharply worldwide amidst growing fears among investors that a global recession is imminent if not already settled in. The panic was partly fueled by remarks by Alan Greenspan that the crisis is “a once-in-a-century credit tsunami”.	Europe / US	Yes
	October 27, 2008	„	US	Yes
	November 5, 2008	Bad news on economic activity	US	Yes
	November 6, 2008	„	„	„
	November 12, 2008	The prospects of a government rescue for US auto makers dwindled	US	No
	November 14, 2008	„	„	„
	November 19, 2008	Proposed federal bailouts of US auto makers failed	US	Yes
	November 20, 2008	„	„	„
	December 1, 2008	"The declines on Wall Street came after stocks fell in Europe..." (December 1, 2008; The Wall Street Journal) Announcement that US economy had officially entered recession in December 2007	Europe / US	Yes
	December 4, 2008	„	„	„

Table 2: Narrated Table of Major US Stock Market Drops.

lead the Japanese return series (i.e., use the original timing) when computing $\text{corr}(\Delta X_t^{\text{JA}}, \Delta X_{t+j}^{\text{US}})$, $j = 1, 2, \ldots$. The same applies for the pair US and emerging markets Asia. This sequencing procedure is symmetric in the sense that both $\text{corr}(\Delta X_t^{\text{US}}, \Delta X_{t+1}^{\text{JA}})$ and $\text{corr}(\Delta X_t^{\text{JA}}, \Delta X_{t+1}^{\text{US}})$ correspond to a 36 hours lag.

Group	Date	Event	Starts in	Transmits?
3	August 4, 1998	"Coming on a day with no significant bad news, the fall marks a shift in investors' mood." (August 5, 1998; The Wall Street Journal)	US	No
	August 27, 1998	"Global markets tumbled on speculation that Yeltsin may resign, along with an indefinite suspension of ruble trading and fear Russia may partially return to Soviet-style economics." (August 28, 1998; The Wall Street Journal)	Russia	Yes
	August 31, 1998	„	„	„
	October 1, 1998	"Greenspan painted a frightening picture of the damage that Long-Term Capital's failure could have inflicted on global economies, as he defended its rescue. He also suggested that more such threats may lurk in the markets." (October 2, 1998; The Wall Street Journal) "The Japanese economy, the world's second largest, appears to be on the brink of depression." (October 1, 1998; Financial Times)	Japan / US	Yes
4	October 27, 1997	Asian flu	Thailand	Yes
5	October 13, 1989	"Friday's sell-off was triggered by the collapse of UAL's buy-out plan and a big rise in producer prices." (October 16, 1989; The Wall Street Journal)	US	Yes
6	April 14, 2000	US News	US	Yes
7	September 17, 2001	WTC terrorist attack	US	Yes
	September 20, 2001	„	„	„
8	September 11, 1986	Worries about higher interest rates and renewed inflation in US	US	Yes
9	March 12, 2001	"Europe's stock markets cracked under the pressure of the plunging Nasdaq Stock Market in the U.S. and other technology-industry anxieties." (March 13, 2001; The Wall Street Journal)	US	Yes
	April 3, 2001	"In Tokyo, the Nikkei Stock Average fell 4.1% as investors dumped shares following Friday's declines in the U.S. markets." (April 10, 2001; The Wall Street Journal)	US	Yes
10	April 4, 1988	Fed increases rates	US	Yes

Table 3: Narrated Table of Major US Stock Market Drops, continued.

5.2.2. US and UK, and US and Developed Countries Europe

To get the sequencing right, we calculate the instantaneous correlation by averaging the instantaneous and 1-day lagged UK returns. To calculate the lagged cross-correlations we average the j-day lagged and $j+1$-day lagged UK returns when computing $\text{corr}(\Delta X_t^{\text{US}}, \Delta X_{t+j}^{\text{UK}})$, $j = 1, 2, \ldots$, and we do not modify the UK series (i.e., use the original timing) when computing $\text{corr}(\Delta X_t^{\text{UK}}, \Delta X_{t+j}^{\text{US}})$, $j = 1, 2, \ldots$. The same applies for the pair US and developed countries Europe. We finally note that the pair US and Latin America does not require any modification because they operate in the same time zone.

Group	Date	Event	Starts in	Transmits?
11	July 10, 2002	"Stocks tumbled, hurt by waning confidence in the market and in corporate integrity. Stock markets plunged around the world following analyst downgrades and the S&P 500 shake-up." (July 11, 2002; The Wall Street Journal)	US	Yes
	July 19, 2002	Internet bubble burst	"	"
	July 22, 2002	Internet bubble burst	"	"
	August 1, 2002	Bearish economic data from the US; Internet bubble burst	US	Yes
	August 5, 2002	Internet bubble burst	"	"
	September 3, 2002	**"Foreign markets usually take their cue from the U.S. Yesterday, the opposite happened."** (September 4, 2002; The Wall Street Journal)	Japan	Yes
	September 19, 2002	Internet bubble burst	US	Yes
	September 27, 2002	"European markets began the day sharply lower in reaction to a late selloff in the U.S. on Friday." (October 1, 2002; The Wall Street Journal)	US	Yes
12	January 4, 2000	The prospect of rising US interest rates shook investor confidence in the technology shares	US	Yes
13	October 25, 1982	Decline of Nasdaq	US	Yes
14	November 15, 1991	US news (statements Bush)	US	Yes
15	March 24, 2003	"Waning optimism about quick Iraq war" (March 26, 2003; The Wall Street Journal)	US	Yes
16	December 20, 2000	US news (statements Fed)	US	Yes
17	February 27, 2007	See Introduction of this paper	China / US	Yes
18	August 6, 1990	Mideast conflict	US	Yes
19	February 5, 2008	US news on service sector	US	Yes
20	July 7, 1986	Worries about US economy	US	Yes
21	March 8, 1996	US jobs data	US	Yes
22	June 6, 2008	"Concerns about reliability of world (Red. oil) supplies; surge, ..., sends stock market plunging and raises fears that US economy could be in for period of inflation and slow growth" (June 7, 2008; The Wall Street Journal)	US	Yes
23	January 9, 1998	Indonesian crises	Indonesia	Yes

Table 4: Narrated Table of Major US Stock Market Drops, continued.

5.3. Empirical Results

We use $4m + m(m-1)/2 + m^2 \times$ (#lags of autocovariances and cross-covariances) moment conditions to identify the parameters of our model, where in the bivariate case that we consider here $m = 2$: the interval-based first moments, autocovariances and cross-covariances, and the third and fourth moments approximated at the leading order in Δ. Dictated by our Monte Carlo results, it is assumed that $\alpha_1 = \alpha_2 =: \alpha$ and that $\lambda_{1,\infty} = \lambda_{2,\infty} =: \lambda_\infty$, and we fix $p_1 = p_2 = 1$. We further use 3 lags of daily autocovariances and cross-covariances. This leaves us with 13 parameters to be identified by 21 moment conditions. GMM parameter estimates are in Table 5.

We note the following:

- The model is fairly parsimonious, but as in any rich multivariate model, relatively large numbers of parameter must be allowed at least until we have a sense of which parameters are not needed, and the initial empirical analysis can then be sensitive to starting values.

1	US	US	US	US	US
2	UK	EU	JA	LA	ASEM
α	20.3**	21.8**	21.3***	24.2***	23.5***
	(9.3)	(9.2)	(6.5)	(4.9)	(0.8)
β_{11}	17.1**	18.6***	17.0**	17.8**	16.2***
	(6.7)	(6.2)	(7.9)	(7.6)	(4.9)
β_{12}	1.2	1.2	1.9	3.0	3.9
	(2.9)	(2.3)	(4.0)	(8.4)	(3.3)
β_{21}	13.1***	8.2**	4.0	5.6	8.6***
	(3.9)	(3.5)	(5.9)	(4.1)	(1.6)
β_{22}	7.2	13.6**	17.3***	18.5**	14.3***
	(4.9)	(6.8)	(5.0)	(7.6)	(1.8)
λ_∞	0.40***	0.31***	0.37**	1.80***	3.14***
	(0.081)	(0.066)	(0.159)	(0.252)	(0.156)
λ_1	4.63	3.88	4.81	19.83	29.54
λ_2	5.25	4.70	6.78	27.56	35.85
$\sqrt{\theta_1}$	0.146***	0.148***	0.146***	0.142***	0.138***
	(0.0132)	(0.0123)	(0.0153)	(0.0063)	(0.0049)
$\sqrt{\theta_2}$	0.175***	0.155***	0.199***	0.211***	0.164***
	(0.0078)	(0.0073)	(0.0059)	(0.0094)	(0.0052)
ρ	0.396***	0.456***	0.408***	0.757***	0.632***
	(0.0493)	(0.0401)	(0.0693)	(0.0748)	(0.0503)
μ_1	0.215***	0.199***	0.218***	0.370***	0.470***
	(0.0458)	(0.0412)	(0.0418)	(0.0395)	(0.0465)
μ_2	0.204***	0.188***	0.245***	0.774***	0.578***
	(0.0383)	(0.0346)	(0.0461)	(0.0751)	(0.0606)
$1/\gamma_1$	0.0307***	0.0326***	0.0304***	0.0156***	0.0138***
	(0.00726)	(0.00737)	(0.00669)	(0.00169)	(0.00138)
$1/\gamma_2$	0.0272***	0.0250***	0.0273***	0.0231***	0.0150***
	(0.00342)	(0.00306)	(0.00339)	(0.00189)	(0.00087)

Table 5: Parameter Estimates for the Bivariate Hawkes Jump-Diffusion Model. This table reports the GMM parameter estimates for the 13 parameters of the bivariate Hawkes jump-diffusion model under exponential decay; standard errors are in parentheses. *, **, and *** indicate significance at the 90%, 95%, and 99% confidence levels, respectively.

To address this issue, a two-stage procedure is adopted: we first estimate the parameters $\{\mu_1, \mu_2, \theta_1, \theta_2, \rho\}$ from the continuous part of our model on the basis of truncated data. To this end, we use moments conditions that are derived from the continuous part of our model, ignoring the discontinuous jump part. Then in a second stage, we use the obtained parameter estimates to identify the parameters $\{\gamma_1, \gamma_2, \lambda_\infty, \alpha, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}\}$. Using the starting values obtained in this way, we then estimate *all* the parameters of the Hawkes jump-diffusion model simultaneously using the full set of moments conditions (doing so removes the slight downward bias in the correlation and volatility parameters). We repeat this multiple times,

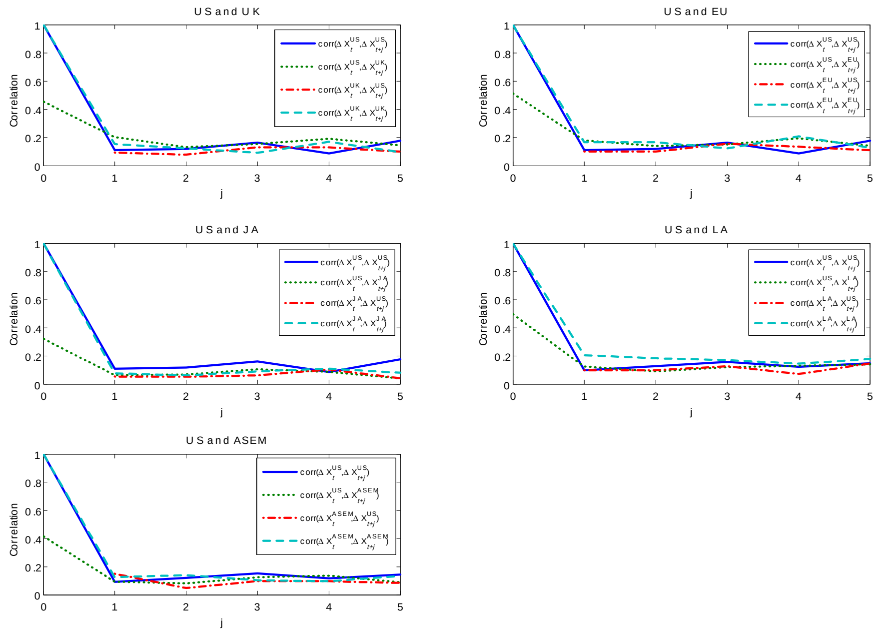

Figure 8: Autocorrelations and Cross-Correlations (Adjusted Data). This figure plots autocorrelations and cross-correlations for the adjusted log-returns of the daily MSCI international equity data with cutoffs. The sequencing adjustments made are described in detail in Section 5.2, while the cutoffs are listed in Section 5.3. The unit of the index j is number of days.

using parameter estimates as starting values, and continuing until starting values coincide with corresponding parameter estimates and a minimal value function is obtained.

- Rather than using the standard identity matrix as the weight matrix in first stage GMM, we put extra weights on the third and fourth moments to make them relatively comparable in magnitude to the other moments.

- When computing the empirical third and fourth moments as well as the empirical *lagged* autocovariances and cross-covariances, the population counterparts of which are completely determined by the jump component of the model, we apply one-sided cutoffs to the data (of -2% for US, UK, EU and JA; -2.5% for ASEM; and -3% for LA), considering only the returns (likely jumps) that are smaller than the given cutoffs. The magnitudes of the cutoffs are dictated by inspection of the corresponding higher order moments and lagged autocovariances and cross-covariances, selecting values for which these moments are robust to changes in the cutoff magnitudes. Figure 8 plots the resulting empirical autocorrelations and cross-correlations.

- As in the Monte Carlo study, visual inspection of the autocorrelograms and cross-correlograms is used to determine the number of autocovariances and cross-covariances we need to include (i.e., to determine the total length of the time period to be considered). In addition, "adaptive analysis" has been carried out: when Hawkes parameter estimates are large, the period considered should be small.

We find large estimated values for β_{11} and β_{22}, measuring the degree of self-excitation. They provide clear evidence that both the US market as well as other markets strongly self-excite. These estimates come along with relatively moderate estimates for the volatility levels. It seems that, due to the integrated nature of the moment conditions, the jump part of our model partially captures what was traditionally (with continuous return dynamics) modeled as (instantaneous) volatility.

Also, we typically find relatively large estimated values for β_{21}, measuring the degree of transmission from the US to other regions of the world. Large values for β_{21} imply that when the US jumps, there is a strong increase in the probability of a consecutive jump in another region of the world. From the empirical cross-correlation plots (see Figure 8) the effect seems to be mainly driven by transmission on the same day or the day following the day of occurrence of a US jump.[11] Estimates of β_{12}, measuring the degree of the reverse transmission, are positive but relatively small. Statistically, the estimated values of the β_{12}'s are not significantly different from zero with daily data.

In addition to the parameter estimates, we also computed the reconstructed autocovariances and cross-covariances that are obtained by substituting the parameter estimates into equations (B.7), (B.8) and (B.9), as well as the reconstructed third and fourth moments. Comparing the reconstructed covariances and higher order moments with their empirical counterparts, we find that the model does quite well in matching the data.

As alluded to in Section 3.7, our estimation methodology allows straightforwardly for testing for the presence of contagion. We test the following null hypotheses: $\mathcal{H}_0^{\mathrm{I}} : \beta_{i,j} = 0,\ i,j = 1,2$; $\mathcal{H}_0^{\mathrm{II}} : \beta_{i,i} = 0,\ i = 1,2$; $\mathcal{H}_0^{\mathrm{III}} : \beta_{i,j} = 0,\ i,j = 1,2,\ i \neq j$. We adopt a Wald chi-square test based on the GMM estimates. Test results are reported in Table 6. The null hypotheses are rejected in the overwhelming majority of the cases, providing clear evidence for excitation (rejection of $\mathcal{H}_0^{\mathrm{I}}$), self-excitation (rejection of $\mathcal{H}_0^{\mathrm{II}}$) and cross-excitation (rejection of $\mathcal{H}_0^{\mathrm{III}}$).

5.4. Robustness Check: Open-to-Close and Close-to-Open Returns

To mitigate the need to carry out sequencing adjustments to the raw data, which depend on the particular time zone in which the markets are operating, we also carry out an analysis with open-to-close (= "daylight" return) and close-to-open returns (= "night" return). With a close-to-open return computed from the previous afternoon's closing price to the next morning's opening price, we have an instantaneous correlation to measure at the daily frequency; see Becker et al. (1990) and Hamao et al. (1990) for other studies using open-to-close and close-to-open returns.

[11] Recall that, because of the sequencing procedure adopted (see Section 5.2), "the same day" for the pair US and JA (ASEM) means here date t for US and date $t+1$ for JA (ASEM). Similarly, "the day following the occurrence of a US jump" for the pair US and JA (ASEM) means here date t for US and date $t+2$ for JA (ASEM).

	Null Hypothesis	US UK	US EU	US JA	US LA	US ASEM	S&P500 Nikkei
$\mathcal{H}_0^{\mathrm{I}}$	No Excitation	***	***	***	***	***	***
$\mathcal{H}_0^{\mathrm{II}}$	No Self-Excitation	***	***	***	***	***	**
$\mathcal{H}_0^{\mathrm{III}}$	No Cross-Excitation	***	***		**	***	***

Table 6: Contagion Test Results. This table reports the rejection significance of the Wald chi-square test statistics where the respective null hypotheses specify complete absence of excitation ($\beta_{i,j} = 0,\ i, j = 1, 2$), absence of self-excitation ($\beta_{i,i} = 0,\ i = 1, 2$), and absence of cross-excitation ($\beta_{i,j} = 0,\ i, j = 1, 2,\ i \neq j$). *, **, and *** indicate rejection at the 90%, 95%, and 99% confidence levels, respectively. The underlying GMM parameter estimates are in Tables 5 and 8.

Panel A	S&P500	Nikkei
Mean	0.000241	-0.000153
St. Deviation	0.011382	0.014234
Skewness	-1.592521	-0.513620
Excess Kurtosis	34.612016	8.502970
Panel B	S&P500	Nikkei
Mean	0.000277	-0.000109
St. Deviation	0.011434	0.013400
Skewness	-1.581570	0.092117
Excess Kurtosis	35.468502	5.998946

Table 7: Descriptive Statistics Open-to-Close and Close-to-Open Data. This table reports descriptive statistics for the log-returns derived from the open-to-close and close-to-open equity index data. Panel A: The opening of the US market marks the beginning of a new day. Panel B: The opening of the Japanese market marks the beginning of a new day. Sample period: January 4, 1984 to December 31, 2008.

Daily returns can be computed either by summing the open-to-close return and the following close-to-open return or by summing the close-to-open return and the following open-to-close return. If the opening of the US market marks the beginning of a new day, the US daily returns will be computed by the former alternative while e.g., the Japanese daily returns will be computed by the latter alternative. If, however, the opening of the Japanese market marks the beginning of a new day, the US daily returns will be computed by the latter alternative while the Japanese daily returns will be computed by the former alternative. As we will see below these "natural" daily return definitions will introduce some asymmetry in the daily data.

We use open and close international equity index data from finance.yahoo.com. We will study two indices: S&P500 and Nikkei. Open-to-close and close-to-open data are available from January 4, 1984. Summary statistics are in Table 7. In Panel A, the opening of the US market marks the beginning of a new day, while in Panel B the opening of the Japanese market marks the beginning of a new day. Autocorrelograms and cross-correlograms are in Figure 9.

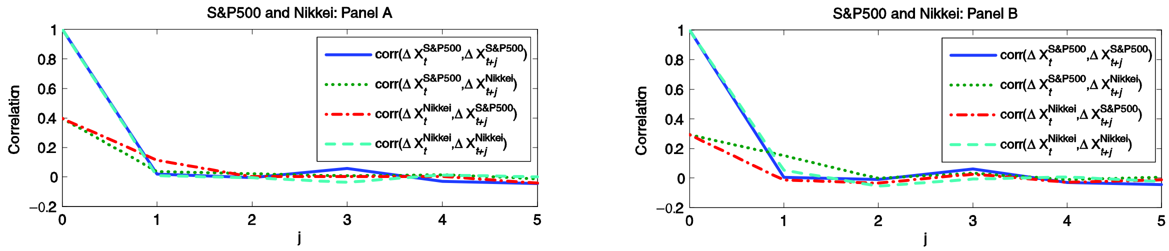

Figure 9: Autocorrelations and Cross-Correlations (Open-to-Close and Close-to-Open Data). This figure plots autocorrelations and cross-correlations for the log-returns derived from the open-to-close and close-to-open equity index data of Table 7. In Panel A the opening of the US market marks the beginning of a new day, while in Panel B the opening of the Japanese market marks the beginning of a new day. The unit of the index j is days.

From Figure 9 we observe some differences between autocorrelations and cross-correlations of Panel A and those of Panel B. Panel A exhibits moderate correlation between Nikkei returns on day j and S&P500 returns on the following day $j+1$, while the correlation between S&P500 returns on day j and Nikkei returns on the following day $j+1$ is relatively small. Conversely, Panel B exhibits substantial correlation between S&P500 returns on day j and Nikkei returns on the following day $j+1$, while the correlation between Nikkei returns on day j and S&P500 returns on the following day $j+1$ is approximately zero. This can be explained from the fact that for Panel A most transmissions from the US to Japan will already be reflected in the Japanese returns of the same day ("instantaneous"), hence a relatively small correlation between S&P500 returns on day j and Nikkei returns on the following day $j+1$. For Panel B, the transmissions from the US to Japan can only be reflected in the Japanese "night" returns of the same day or in the returns of the following days. It appears from the results that the Japanese night return does not significantly reflect the US cross-excitation, hence a substantial correlation between S&P500 returns on day j and Nikkei returns on the following day $j+1$. The converse applies to the transmission of shocks from Japan to the US.

Because the direction (causality) of transmissions on the same day cannot be identified with daily data (here composed from the intradaily open-to-close and close-to-open returns), GMM learns most about the direction of transmission from the covariances between day j and day $j+1$. Given the results of our qualitative analysis reported in Tables 2, 3 and 4, the most natural way to proceed now is to restrict attention to Panel B, letting the opening of the Japanese market mark the beginning of a new day. GMM parameter estimates are in Table 8. The GMM parameter estimates are obtained under the same assumptions as in Section 5.3.

The parameter estimates we find are consistent with the parameter estimates reported earlier in Section 5.3: there is clear evidence for self-excitation in the US market. The transmission from the US market to Japan is more pronounced than earlier, and the self-excitation in the Japanese market is still present, be it a little less pronounced. Again, we find positive but small

1	S&P500	1	S&P500
2	Nikkei	2	Nikkei
α	20.8***	$\sqrt{\theta_1}$	0.142***
	(3.7)		(0.0176)
β_{11}	17.6**	$\sqrt{\theta_2}$	0.195***
	(8.5)		(0.0064)
β_{12}	1.5	ρ	0.360***
	(4.6)		(0.0331)
β_{21}	10.8*	μ_1	0.206***
	(6.5)		(0.0452)
β_{22}	10.1*	μ_2	0.132***
	(5.8)		(0.0443)
λ_∞	0.38***	$1/\gamma_1$	0.0327***
	(0.070)		(0.00701)
λ_1	5.11	$1/\gamma_2$	0.0260***
λ_2	5.85		(0.00165)

Table 8: Parameter Estimates for the Bivariate Hawkes Jump-Diffusion Model: Open-to-Close and Close-to-Open Data. This table reports the GMM parameter estimates for the 13 parameters of the bivariate Hawkes jump-diffusion model under exponential decay; standard errors are in parentheses. *, **, and *** indicate significance at the 90%, 95%, and 99% confidence levels, respectively. The underlying data are the log-returns derived from open-to-close and close-to-open equity index data.

estimates, that are statistically not distinguishable from zero, for the reverse transmission from Japan to the US. We recall here that the sample period for the pair S&P500 and Nikkei is different from (shorter than) the sample period for the MSCI pair US and Japan considered earlier. Furthermore, while $\text{corr}(\Delta X_t^{\text{US}}, \Delta X_{t+1}^{\text{JA}})$ and $\text{corr}(\Delta X_t^{\text{JA}}, \Delta X_{t+1}^{\text{US}})$ correspond to a 36 hours lag, $\text{corr}(\Delta X_t^{\text{S\&P500}}, \Delta X_{t+1}^{\text{Nikkei}})$ and $\text{corr}(\Delta X_t^{\text{Nikkei}}, \Delta X_{t+1}^{\text{S\&P500}})$ correspond to a 24 hours lag. Contagion test results were reported in the last column of Table 6 for this mesurement of returns. Again the tests provide clear evidence for excitation, self-excitation as well as cross-excitation.

6. Implications of the Model

Our model is *reduced-form*: it cannot explain the source(s) of the contagion that is observed in the data, or get at the channels of transmission of that contagion, whether trade linkages, financial linkages, financial constraints, outflows of capital, herding behaviors, the fragility of the system, lack of coordinated responses, etc. On the other hand, as a description of the process driving the asset returns, the model can be employed as input for other purposes, in the Merton tradition. Three such applications which we now briefly outline consist in measuring market stress, risk management and optimal portfolio choice.

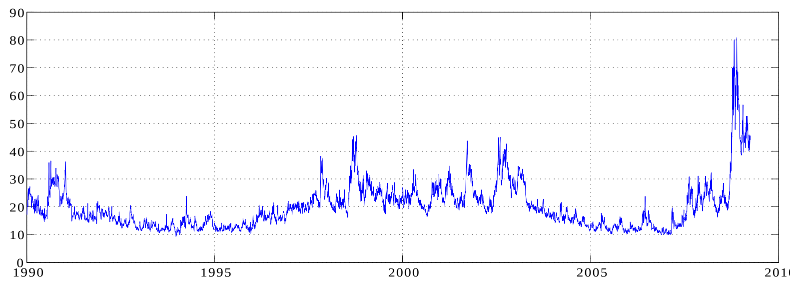

Figure 10: Volatility Index (VIX). This figure plots the close index values of the VIX. Sample period: January 2, 1990 to March 31, 2009.

6.1. Measuring Market Stress Using Filtered Values of the Jump Intensities

Stress in the marketplace is often measured using a volatility index such as the VIX from the Chicago Board Options Exchange (CBOE), which has often been labeled the "fear gauge." VIX levels above 80% at an annualized rate were recorded at the height of the financial crisis in the Fall of 2008, as shown in Figure 10. This is an unrealistic level from the perspective of a purely-diffusive model, at least if we take the implications of such a level over a few months' horizon, and is most likely indicative that jumps are feared. If the model does not allow for jumps, then the only way that high risk can be translated is through extreme volatility numbers. But what is measured by VIX is a form of total (continuous plus jump) variance. To the extent that the risk that is truly feared is that of jumps, then a measure that captures solely that risk would be advantageous.

Consider instead the Hawkes jump intensities as a measure of market stress. By construction, higher values of $\lambda_{i,t}$ lead to higher probability of jumps in asset i, and should be reflected in higher derivative prices such as options. In a Hawkes model, these intensities are time-varying, and it is likely that they will reflect market conditions at the time. Just like volatilities, jump intensities are latent; unlike volatilities, no market instrument is currently traded which references jump intensity. There are no "jump intensity swaps," for instance. The only solution to infer jump intensities is therefore to filter them out of the observed time series.

To illustrate, we focus on the US and UK markets. In order to filter jump intensities from the time series of asset returns, we insert the parameter estimates of Table 5 in (2.9), and call a jump (hence, $dN_{j,t} = 1$) each daily return in the corresponding time series that is below $-1/\hat{\gamma}_{\text{US}} = -3.07\%$, and $-1/\hat{\gamma}_{\text{UK}} = -2.72\%$, respectively. (Filtering jumps on the basis of large returns is a natural approach, see e.g., Lee and Mykland (2008).) The resulting plots of the estimated time series of $(\lambda_{\text{US},t}, \lambda_{\text{UK},t})$ are in Figure 11. The plots show that the filtered $\lambda_{i,t}$'s are plausible indicators of market stress, increasing in particular in the sample around the crisis periods identified in Tables 2, 3 and 4. Indeed, from the plots we observe that the two most noticeable periods of market stress are the summer of 2002 (Group 11 in the tables; Internet bubble burst) and the fall of 2008 (Group 2 in the tables; Credit crisis). Except here for the estimation of the parameters which is based on

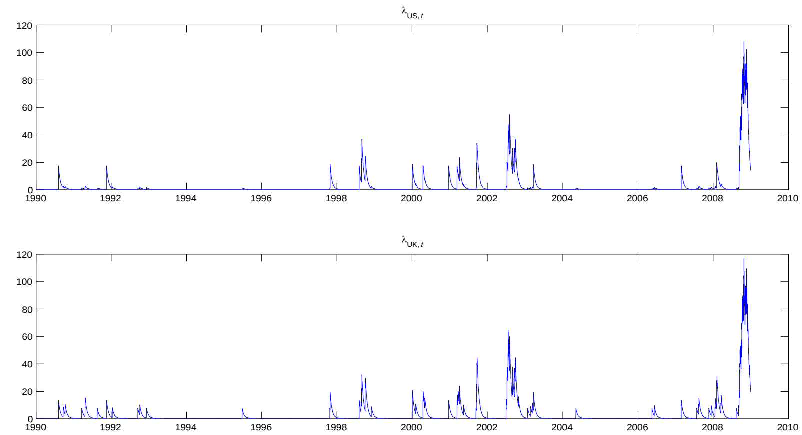

Figure 11: Estimated Time Series of the US and UK Jump Intensities. This figure plots the estimated time series of $\lambda_{\text{US},t}$ and $\lambda_{\text{UK},t}$, filtered from the time series of asset returns. Sample period: January 2, 1990 to March 31, 2009.

the full sample, this filtering approach is in principle not anticipative, i.e., it can be conducted in nearly real time.

6.2. Risk Management and Tail Probabilities Implied by the Model

A second natural application of the model consists in computing various risk measures, which all are functions of the left tail of the asset returns' distribution. From a risk management perspective, one may be interested in the "co-jump" risk of two assets jumping negatively together. For instance, a banking regulator might be concerned with an event where two large financial institutions experience large losses during the same period, perhaps because they were following similar strategies. In this context, reduced-form models are commonly used to set capital requirements, for stress testing, or to generate scenarios in simulations. We have fitted above our model to international equity indices, but the same model could just as easily be fitted to domestic markets, or to different sectors of the economy, or to different stocks.

Our model consists of day-to-day (continuous) variations, but also incorporates jumps. As such, it can conceivably be employed in times of crises, when the intensity of jumps is high, but also in normal times, when the continuous part of the model is the primary driver of asset returns. Among jumps, we argued that Poissonian jump models fail to capture jump clustering to the extent it is present in the data, and that self- and cross-excitation modeling can be a parsimonious way of generating the type of clustering that is observed empirically and that is therefore relevant for risk management.

To illustrate, we start with the univariate case, and contrast the left tails' magnitudes that

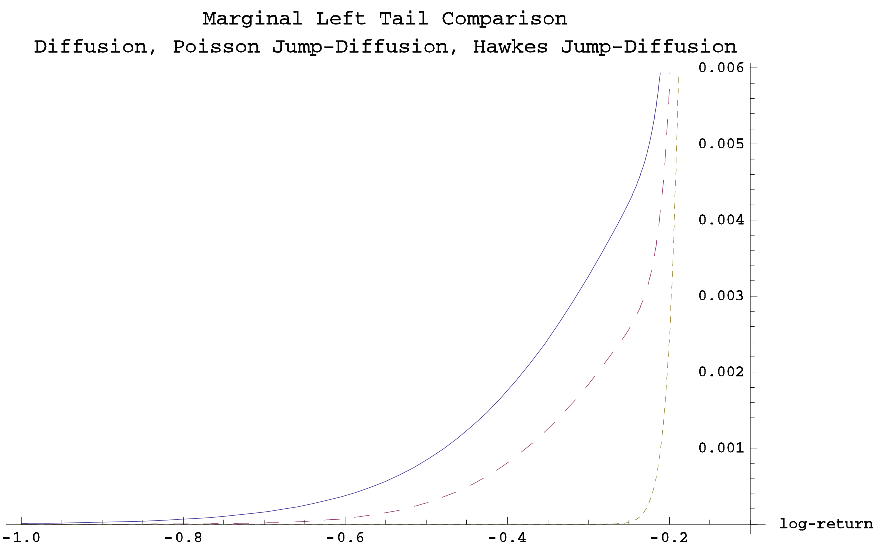

Figure 12: Left Tail Comparison of Univariate Models. This figure plots the log-returns distribution in a pure diffusion model (short dashes), a Poisson jump-diffusion (long dashes) and a self-exciting Hawkes jump-diffusion (solid curve).

follow from a purely diffusive model with no jumps, a Poisson jump-diffusion model, and a Hawkes jump-diffusion for returns measured over the typical VaR calculation horizon of $\Delta = 10$ days. In the univariate case, the only form of excitation that distinguishes Hawkes jumps from Poisson jumps is self-excitation. As we saw above from the empirical parameter estimates, most of the excitation happens over a matter of days, so it is quite plausible to expect that the left tails of the returns distribution will be affected over the 10-day VaR horizon. We show in Figure 12 the respective left tails from these three models.[12] With realistic volatility values, a diffusion-only model will not generate a tail with substantial risk. Incorporating a Poisson jump component into the model gives rise to a larger left tail (the model with Poissonian constant intensities is calibrated to the long-term average intensities already estimated). However, it is very unlikely under Poisson jumps with these parameter values to give rise to more than a single jump over a 10-day horizon, which limits the extent to which the investor or regulator should be concerned once proper capital reserves have been set aside to cover a one-jump event. On the other hand, Hawkes jumps make it possible, and even likely if the self-excitation parameter β is sufficiently high, to record more than one jump in close succession. As the figure shows, this gives rise to a longer left tail with a visible inflexion at the level where a second jump occurs in the same time period.

We now move to the perhaps more relevant situation where one faces the risk embedded in multiple assets together, focusing for illustration purposes on the two-asset case. Consider again a

[12] Because jumps are relatively rare and inherently unpredictable, any model will produce tails that will necessarily only be valid on average and inevitably lead to an underestimation of the risk when a jump actually happens. In no way, is VaR interpretable as a worst-case or maximum tolerable loss once jumps are taken into account. This is an issue with the use and interpretation of VaR, not with the model driving the asset returns per se.

typical VaR calculation over a time horizon of $\Delta = 10$ days, either from the perspective of a regulator concerned with the probability of joint individual losses L_1 and L_2 in two firms under scrutiny, $\mathbb{P}(\Delta X_1 \leq -L_1, \Delta X_2 \leq -L_2)$, or that of a portfolio manager concerned with losses exceeding a level L in a portfolio invested in the two assets in proportions ω_1 and ω_2, $\mathbb{P}(\omega_1 \Delta X_1 + \omega_2 \Delta X_2 \leq -L)$.

In both situations, the relevant calculation involves the joint distribution of returns or prices, and specifically their left tails. And, for sufficiently large loss amounts or percentages, it is effectively the jump component of the model that drives the result. We will therefore start by considering the joint probability of obtaining multiple jumps, n_1 for asset 1 and n_2 for asset 2, in a time interval of the length Δ. We will compare the results obtained from a classical Poisson jump process to those obtained from a Hawkes mutually exciting jump process. For this purpose, we consider a bivariate Poisson jump process (N_1, N_2) with arrival rates (λ_1, λ_2) and possibly correlated jumps. We construct such a process (N_1, N_2) as $N_1 = \tilde{N}_1 + N$, $N_2 = \tilde{N}_2 + N$ where $\tilde{N}_1$, $\tilde{N}_2$ and N are independent Poisson processes with respective parameters $\tilde{\lambda}_1 = \lambda_1 - \phi$, $\tilde{\lambda}_2 = \lambda_2 - \phi$ and ϕ. The marginal distributions of N_1 and N_2 are Poisson with parameters λ_1 and λ_2, which are the mean and variance of the two marginal jump processes, and $\text{cov}(N_1, N_2) = \phi$. The corresponding correlation coefficient must be restricted to the range $[0, \min((\lambda_1/\lambda_2)^{1/2}, (\lambda_2/\lambda_1)^{1/2})]$. The joint distribution of the number of jumps (N_1, N_2) is then given by

$$\mathbb{P}(N_{1,\Delta} = n_1, N_{2,\Delta} = n_2) = e^{-(\lambda_1+\lambda_2-\phi)\Delta} \sum_{i=0}^{\min(n_1,n_2)} \frac{(\lambda_1 - \phi)^{n_1-i} (\lambda_2 - \phi)^{n_2-i} \phi^i \Delta^{n_1+n_2-i}}{(n_1 - i)! (n_2 - i)! i!}. \quad (6.1)$$

The key difference between the mutually exciting model and the Poisson model is that in the Poissonian model the fact that the first asset jumped, even with positive covariation ϕ, only predicts a small increase in the probability that the second asset will jump too in the same time frame,

$$\mathbb{P}(N_{2,\Delta} = n_2 | N_{1,\Delta} = n_1) = \frac{\mathbb{P}(N_{1,\Delta} = n_1, N_{2,\Delta} = n_2)}{\mathbb{P}(N_{1,\Delta} = n_1)}.$$

By contrast, in the mutually exciting model, the fact that the first asset jumps leads to a large increase in the probability of the second asset jumping, by a factor β_{21}, which we recall from the previous empirical analysis is estimated at high values. For instance, the first asset jumping could easily lead to a ten-fold increase in the probability that the second asset will also jump in short order. In addition to the cross-excitation β_{21}, we allow for the presence of self-excitation in the form of β_{11} and β_{22}. In order to examine the effect of excitation asymmetry between the two assets, we rule out cross-excitation in the reverse direction, $\beta_{12} = 0$, which is generally approximately compatible with the empirical findings.

From this perspective, cross-excitation introduces systemic risk since it raises significantly the probability that multiple assets will jump over the same relatively short time period. To examine the magnitude of the excitation effect empirically, we compare contour plots showing the returns distributions over 10-day horizons obtained from the model in the three situations where: (i) there are no jumps, all the dependence is due to the common factor variation in the Brownian risk; (ii) the jumps are Poissonian, independent; (iii) the jumps are Poissonian, correlated with coefficient

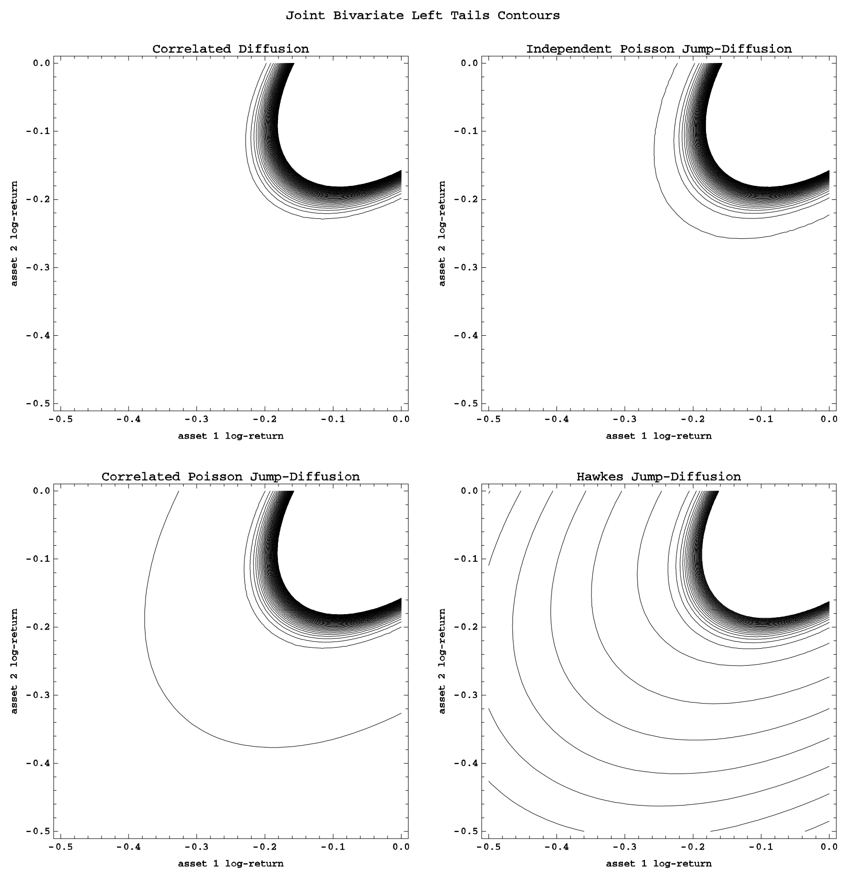

Figure 13: Left Tail Comparison of Bivariate Models. This figure plots the left tail contours of the bivariate log-returns distribution in a correlated diffusion, an independent Poisson jump-diffusion, a correlated Poisson jump-diffusion and a mutually exciting Hawkes jump-diffusion.

ϕ; (iv) the jumps are mutually exciting as in the model we have employed here. In the mutually exciting case, as a first approximation over small time horizons Δ, we can assess the effect of the mutual excitation in the form of an increase of the Poissonian jump intensity from λ_1 to $\lambda_1 + \max(0, n_1 - 1)\beta_{11}$ and from λ_2 to $\lambda_2 + \max(0, n_2 - 1)\beta_{22} + \max(0, n_1)\beta_{21}$, which discards the impact of the mean reversion over small Δ.

The results are reported in Figure 13 in the form of contour plots of the joint left tails of the resulting distribution of returns $(\Delta X_1, \Delta X_2)$. As in the univariate case, tails decay very quickly in the Brownian-only model, even with a common factor structure generating a correlation of 0.5 between the two assets' returns. In the two Poissonian jump models, the probability of observing successive jumps in the same small time period decreases rapidly with the number of jumps. It is very unlikely that more than one jump will be recorded under such a model. Introducing positive correlation (assumed to be 0.5) between the two Poisson jumps increases slightly the probability of observing jumps in the two assets together relative to the independent model but does not fundamentally alter the patterns of probabilities. By contrast, the introduction of mutual excitation radically increases the probability that multiple jumps will be recorded in the same time period compared to the two Poissonian models. The asymmetry between self-excitation and cross-excitation manifests itself for instance in the higher probability of observing $n_2 = 2$ jumps when there have been $n_1 = 1$ jumps than the reverse. As a result, the contour plots are slightly asymmetrical.

6.3. Optimal Portfolio Choice in Closed-Form

Using an extension to the present setting of the method of Aït-Sahalia et al. (2009), it is possible to solve in closed-form for the optimal portfolio of a log-utility investor who faces mutually exciting jump risk across assets. Specifically, consider the asset return dynamics

$$\frac{dS_{0,t}}{S_{0,t}} = rdt, \quad \frac{dS_{i,t}}{S_{i,t-}} = (r + R_i)\, dt + \sigma_i dW_{i,t} + Z_{i,t} dN_{i,t}, \qquad i = 1, \ldots, m$$

where $S_{i,t}$ are the asset prices, $r \geq 0$ is a constant rate of interest, R_i are expected excess rates of return, and $N_t = [N_{1,t}, \ldots, N_{m,t}]'$ is an m-dimensional mutually exciting process with mean-reverting intensity (2.9). The Brownian motions, random jump sizes and jump processes are assumed to be mutually independent.

Across a potentially large number m of assets, it makes sense to impose a factor structure on the variance-covariance matrix of Brownian risk Σ, with elements $\Sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j$, as in the Arbitrage Pricing Theory. The investor's problem at time t, then, is to pick consumption and portfolio weights, $\{C_s, \omega_s\}_{t \leq s}$, to maximize the expected utility,

$$V(Y_t, t) = \max_{\{C_s, \omega_s;\ t \leq s\}} \mathbb{E}_t \left[\int_t^\infty e^{-\varrho s} U(C_s) ds \right],$$

subject to the wealth dynamics $dY_t = -C_t dt + \sum_{i=0}^m \omega_{i,t} \frac{dS_{i,t}}{S_{i,t-}} Y_t$. Here, U is the investor's utility function and ϱ the subjective discount rate.

With this structure, the optimal portfolio weights are solvable in closed-form in some specific situations, for example, when the investor has logarithmic utility: $U(x) = \log(x)$ and the jump sizes are equally distributed. The optimal portfolio weights are now time-varying, with the investor reacting to changes in the intensity of the jumps. And, even though this is a log-utility investor, the solution is not myopic! This is in contrast to what is obtained in the classical Merton optimal solution for a log-utility investor, Brownian-driven (or even Brownian plus Poisson jumps), where log-utility leads to a myopic solution.

7. Conclusions

We proposed a natural reduced-form model for asset returns that is able to capture jump clustering in time and across assets, or financial contagion. The model is coined the *Hawkes jump-diffusion* *model*. Unlike other models based on Hawkes jump processes, our model incorporates the standard elements of drift and volatility in addition to these jumps, thereby being most naturally thought of as a generalization of the Poisson jump-diffusion model familiar to financial economists. We developed an estimation procedure for the model based on GMM. The estimation procedure relies entirely on closed-form moment functions, hence is amenable to a multivariate setting and is numerically tractable. Monte Carlo evidence shows the population parameters of the data generating process can be recovered with a sufficient degree of precision for practical applications. We implemented the estimation procedure on international equity data, studying the patterns of jump excitation among six world markets. Our empirical results indicate that both the US market as well as the other markets strongly self-excite. Furthermore, we find that US jumps tend to get reflected quickly in most other markets, while we find that statistical evidence for the reverse transmission is much less pronounced.

The model is amenable to a few applications, some of which we briefly outlined: measuring market stress using filtered values of the jump intensities, using the model to generate the tail probabilities relevant in risk management context, and in particular the risk of observing jumps in multiple assets (or firms) in the same time period, and the calculation of optimal portfolio weights. Let us finally indicate possible future work: one could incorporate in the analysis the information implied from derivatives data, or optimal portfolio returns obtained by solving the corresponding consumption-portfolio selection problem, in addition to the time series observations of the underlying asset returns.

References

Aït-Sahalia, Y. (2004): "Disentangling diffusion from jumps," *Journal of Financial Economics*, 74, 487–528.

Aït-Sahalia, Y., J. Cacho-Diaz, and T. R. Hurd (2009): "Portfolio choice with jumps: A closed-form solution," *The Annals of Applied Probability*, 19, 556–584.

Ang, A. and G. Bekaert (2002): "International asset allocation with regime shifts," *Review of Financial Studies*, 15, 1137–1187.

Ang, A. and J. Chen (2002): "Asymmetric correlations of equity portfolios," *Journal of Financial Economics*, 63, 443–494.

Azizpour, S. and K. Giesecke (2008): "Self-exciting corporate defaults: Contagion vs. frailty," Tech. rep., Stanford University.

Bae, K.-H., G. A. Karolyi, and R. M. Stulz (2003): "A new approach to measuring financial contagion," *Review of Financial Studies*, 16, 717–763.

Becker, K. G., J. E. Finnerty, and M. Gupta (1990): "The intertemporal relation between the U.S. and Japanese stock markets," *The Journal of Finance*, 45, 1297–1306.

Bekaert, G. and C. R. Harvey (1995): "Time-varying world market integration," *The Journal of Finance*, 50, 403–444.

——— (2000): "Foreign speculators and emerging equity markets," *The Journal of Finance*, 55, 565–613.

Bowsher, C. G. (2007): "Modelling security market events in continuous time: Intensity based, multivariate point process models," *Journal of Econometrics*, 141, 876–912.

Brillinger, D. R. (1988): "Some Statistical Methods for Random Process Data from Seismology and Neurophysiology," *Annals of Statistics*, 16, 1–54.

Calvo, G. and E. Mendoza (2000): "Rational contagion and the globalization of securities markets," *Journal of International Economics*, 51, 79–113.

Caramazza, F., L. Ricci, and R. Salgado (2004): "International financial contagion in currency crises," *Journal of International Money and Finance*, 23, 51–70.

Chang, R. and A. Velasco (2001): "A model of financial crises in emerging markets," *Quarterly Journal of Economics*, 116, 489–517.

Crane, R. and D. Sornette (2008): "Robust dynamic classes revealed by measuring the response function of a social system," *Proceedings of the National Academy of Sciences*, 105, 15649–15653.

Das, S. and R. Uppal (2004): "Systemic risk and international portfolio choice," *The Journal of Finance*, 59, 2809–2834.

Dornbusch, R., Y. Park, and S. Claessens (2000): "Contagion: Understanding how it spreads," *The World Bank Research Observer*, 15, 177–197.

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

Dungey, M. and B. Gonzalez-Hermosillo (2005): "Empirical modeling of contagion: A review of methodologies," *Quantitative Finance*, 5, 9–24.

Dungey, M. and V. Martin (2007): "Unravelling financial market linkages during crises," *Journal of Applied Econometrics*, 22, 89–119.

Eichengreen, B., A. Rose, and C. Wyplosz (1996): "Contagious currency crises: First tests," *Journal of Economics*, 98, 463–484.

Engle, R. F. (1982): "Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation," *Econometrica*, 50, 987–1008.

Errais, E., K. Giesecke, and L. R. Goldberg (2010): "Affine Point Processes and Portfolio Credit Risk," *SIAM Journal on Financial Mathematics, forthcoming.*

FELLER, W. (1951): "Two singular diffusion problems," *Annals of Mathematics*, 54, 173–182.

FORBES, K. J. AND R. RIGOBON (2002): "No contagion, only interdependence: Measuring stock market comovements," *The Journal of Finance*, 57, 2223–2261.

FRY, R., V. L. MARTIN, AND C. TANG (2009): "A new class of tests of contagion with applications to real estate markets," *Journal of Business and Economic Statistics, forthcoming.*

GERLACH, S. AND F. SMETS (1995): "Contagious speculative attacks," *European Journal of Political Economy*, 11, 5–63.

GLICK, R. AND A. ROSE (1999): "Why are currency crises regional?" *Journal of International Money and Finance*, 18, 603–617.

GOETZMANN, W. N., L. LI, AND K. G. ROUWENHORST (2005): "Long-term global market correlations," *Journal of Business*, 78, 1–38.

HAMAO, Y., R. W. MASULIS, AND V. NG (1990): "Correlations in price changes and volatility across international stock markets," *Review of Financial Studies*, 3, 281–307.

HANSEN, L. P. (1982): "Large sample properties of generalized method of moments estimators," *Econometrica*, 50, 1029–1054.

HAWKES, A. G. (1971a): "Spectra of some self-exciting and mutually exciting point processes," *Biometrika*, 58, 83–90.

——— (1971b): "Point spectra of some mutually exciting point processes," *Journal of the Royal Statistical Society, Series B*, 33, 438–443.

HAWKES, A. G. AND D. OAKES (1974): "A cluster representation of a self-exciting process," *Journal of Applied Probability*, 11, 493–503.

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

KAMINSKY, G. AND C. REINHART (1998): "Financial crises in Asia and Latin America: Then and now," *American Economic Review*, 88, 444–448.

——— (2000): "On crises, contagion and confusion," *Journal of International Economics*, 51, 145–168.

KARR, A. F. (1991): *Point processes and their statistical inference*, New York: Marcel Dekker.

KNIGHT, J. L. AND S. E. SATCHELL (1998): "GARCH Processes - Some Exact Results, Some Difficulties and a Suggested Remedy," in *Forecasting Volatility in the Financial Markets*, ed. by J. L. Knight and S. E. Satchell, Oxford: Butterworth Heineman, 321–346.

Kodres, L. and M. Pritsker (2002): "A rational expectations model of financial contagion," *The Journal of Finance*, 57, 769–799.

Kou, S. G. (2002): "A jump-diffusion model for option pricing," *Management Science*, 48, 1086–1101.

Krugman, P. (1979): "A model of balance of payments crises," *Journal of Money, Credit, and Banking*, 11, 311–325.

Lee, S. and P. A. Mykland (2008): "Jumps in financial markets: A new nonparametric test and jump dynamics," *Review of Financial Studies*, 21, 2535–2563.

Longin, F. and B. Solnik (2001): "Extreme correlation of international equity markets," *The Journal of Finance*, 56, 649–676.

Maheu, J. M. and T. H. McCurdy (2004): "News arrival, jump dynamics and volatility components for individual stock returns," *The Journal of Finance*, 59, 755–793.

Merton, R. C. (1976): "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics*, 3, 125–144.

Nikitin, M. and R. Smith (2008): "Information acquisition, coordination, and fundamentals in a financial crisis," *Journal of Banking and Finance*, 32, 907–914.

Oakes, D. (1975): "The Markovian self-exciting process," *Journal of Applied Probability*, 12, 69–77.

Obstfeld, M. (1996): "Models of currency crises with self-fulfilling features," *European Economic Review*, 40, 1037–1047.

Ogata, Y. and H. Akaike (1982): "On linear intensity models for mixed doubly stochastic Poisson and self-exciting point processes," *Journal of the Royal Statistical Society, Series B*, 44, 102–107.

Pavlova, A. and R. Rigobon (2008): "The role of portfolio constraints in the international propagation of shocks," *Review of Economic Studies*, 75, 1215–1256.

Rigobon, R. (2003): "On the measurement of the international propagation of shocks: Are they stable?" *Journal of International Economics*, 61, 261–283.

Rijckeghem, C. V. and B. Weder (2001): "Sources of contagion: Is it finance or trade?" *Journal of International Economics*, 54, 293–308.

Salmon, M. and W. W. Tham (2007): "Time Deformation and the Yield Curve," Tech. rep., Warwick Business School.

Yu, J. (2004): "Empirical characteristic function estimation and its applications," *Econometric Reviews*, 23, 93–123.

Appendix: Explicit Formulae for the Moment Conditions

A. Univariate Case

Let $m = 1$ and assume that

$$\begin{cases} dX_t = \mu dt + \sqrt{V_t} dW_t^X + Z_t dN_t \\ dV_t = \kappa(\theta - V_t)dt + \eta\sqrt{V_t} dW_t^V \\ d\lambda_t = \alpha\left(\lambda_\infty - \lambda_t\right)dt + \beta dN_t \end{cases} \tag{A.1}$$

with $\mathbb{E}\left[dW_t^X dW_t^V\right] =: \rho^V dt$. The vector of Brownian motions W, the jump size Z and the jump process N are assumed to be mutually independent. The corresponding integral equation for λ_t reads:

$$\lambda_t = \lambda_\infty + \int_{-\infty}^{t} \beta e^{-\alpha(t-s)} dN_s. \tag{A.2}$$

A.1. Moment Conditions: Univariate Self-Exciting Jumps

Write

$$\lambda := \mathbb{E}\left[\lambda_t\right], \qquad V_N(\tau) := \mathbb{E}\left[dN_{t+\tau} dN_t\right] / (dt)^2 - \lambda^2. \tag{A.3}$$

Then $\mathbb{E}\left[dX_t\right]/dt = \mu + \lambda\mathbb{E}\left[Z\right]$, and

$$\frac{\mathbb{E}[dX_{t+\tau} dX_t]}{(dt)^2} - \left(\frac{\mathbb{E}[dX_t]}{dt}\right)^2 = \left(\mathbb{E}\left[Z\right]\right)^2 V_N(\tau), \qquad \tau > 0; \tag{A.4}$$

with $\lambda = \lambda_\infty \alpha/(\alpha - \beta)$ and $V_N(\tau) = \frac{\beta\lambda(2\alpha-\beta)}{2(\alpha-\beta)} e^{-(\alpha-\beta)\tau}$ for $\tau > 0$. Furthermore,

$$\begin{aligned} \mathbb{E}\left[\left(dX_t - \mathbb{E}\left[dX_t\right]\right)^2\right]/dt &= \theta + \lambda\mathbb{E}\left[Z^2\right] \\ \mathbb{E}\left[\left(dX_t - \mathbb{E}\left[dX_t\right]\right)^3\right]/dt &= \lambda\mathbb{E}\left[Z^3\right] \\ \mathbb{E}\left[\left(dX_t - \mathbb{E}\left[dX_t\right]\right)^4\right]/dt &= \lambda\mathbb{E}\left[Z^4\right]. \end{aligned}$$

A.2. Interval-Based Moment Conditions: Univariate Self-Exciting Jumps

Let $0 \le s_1 < s_2 < s_3 < s_4$ and let $\Delta_1 := s_2 - s_1$, $\Delta_2 := s_4 - s_3$ and $\tau := s_3 - s_1$. Write $\lambda := \mathbb{E}\left[\lambda_t\right]$, and

$$\begin{aligned} f_N(\tau, \Delta_1, \Delta_2) &:= \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dN_u \int_{s_1}^{s_2} dN_t\right]}{\Delta_1\Delta_2} - \lambda^2, \\ g_{N,Z}(\Delta_1) &:= \frac{\mathbb{E}\left[\left(\int_{s_1}^{s_2} Z_t dN_t\right)^2\right]}{\Delta_1} - \left(\mathbb{E}[Z]\right)^2 \lambda^2 \Delta_1. \end{aligned}$$

Then $\mathbb{E}\left[\int_{s_1}^{s_2} dX_t\right]/\Delta_1 = \mu + \lambda\mathbb{E}\left[Z\right]$, and

$$\frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_u \int_{s_1}^{s_2} dX_t\right]}{\Delta_1\Delta_2} - \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_u\right]\mathbb{E}\left[\int_{s_1}^{s_2} dX_t\right]}{\Delta_1\Delta_2} = \left(\mathbb{E}\left[Z\right]\right)^2 f_N(\tau, \Delta_1, \Delta_2);$$

with

$$f_N(\tau,\Delta_1,\Delta_2) = \frac{\beta\lambda(2\alpha-\beta)}{2\Delta_1\Delta_2(\alpha-\beta)^3}\left(e^{-(\alpha-\beta)(\tau-\Delta_1)} - e^{-(\alpha-\beta)(\tau-\Delta_1+\Delta_2)} - e^{-(\alpha-\beta)\tau} + e^{-(\alpha-\beta)(\tau+\Delta_2)}\right).$$

Furthermore,

$$\frac{\mathbb{E}\left[\left(\int_{s_1}^{s_2} dX_t\right)^2\right]}{\Delta_1} - \frac{\left(\mathbb{E}\left[\int_{s_1}^{s_2} dX_t\right]\right)^2}{\Delta_1} = \theta + g_{N,Z}(\Delta_1);$$

with

$$g_{N,Z}(\Delta_1) = \lambda\mathbb{E}\left[Z^2\right] + (\mathbb{E}[Z])^2\frac{\beta\lambda(2\alpha-\beta)}{(\alpha-\beta)^2}\left(1 + \frac{1}{\Delta_1(\alpha-\beta)}\left(e^{-(\alpha-\beta)\Delta_1} - 1\right)\right).$$

B. Bivariate Case

Let $m = 2$ and assume that

$$\begin{cases} dX_{1,t} = \mu_1 dt + \sqrt{V_{1,t}}dW_{1,t}^X + Z_{1,t}dN_{1,t} \\ dX_{2,t} = \mu_2 dt + \sqrt{V_{2,t}}dW_{2,t}^X + Z_{2,t}dN_{2,t} \\ dV_{1,t} = \kappa(\theta_1 - V_{1,t})dt + \eta_1\sqrt{V_{1,t}}dW_t^V \\ dV_{2,t} = d\left(\frac{\theta_2}{\theta_1}\right)V_{1,t} \\ d\lambda_{1,t} = \alpha_1\left(\lambda_{1,\infty} - \lambda_{1,t}\right)dt + \beta_{11}dN_{1,t} + \beta_{12}dN_{2,t} \\ d\lambda_{2,t} = \alpha_2\left(\lambda_{2,\infty} - \lambda_{2,t}\right)dt + \beta_{21}dN_{1,t} + \beta_{22}dN_{2,t} \end{cases} \tag{B.1}$$

with $\mathbb{E}[dW_{1,t}^X dW_{2,t}^X] =: \rho dt$ and $\mathbb{E}[dW_{i,t}^X dW_t^V] =: \rho_i^V dt$, $i = 1,2$. The vector of Brownian motions W, the vector of jump sizes Z and the vector of jump processes N are assumed to be mutually independent. The corresponding integral equation for $\lambda_{i,t}$ reads:

$$\lambda_{i,t} = \lambda_{\infty,i} + \int_{-\infty}^t \beta_{i,1}e^{-\alpha_i(t-s)}dN_{1,s} + \int_{-\infty}^t \beta_{i,2}e^{-\alpha_i(t-s)}dN_{2,s}, \qquad i = 1,2. \tag{B.2}$$

B.1. Moment Conditions: Bivariate Mutually Exciting Jumps

Write

$$\lambda_i := \mathbb{E}\left[\lambda_{i,t}\right],\ i = 1,2, \quad V_N(\tau) := \begin{pmatrix} \frac{\mathbb{E}[dN_{1,t+\tau}dN_{1,t}]}{(dt)^2} - \lambda_1^2 & \frac{\mathbb{E}[dN_{1,t+\tau}dN_{2,t}]}{(dt)^2} - \lambda_1\lambda_2 \\ \frac{\mathbb{E}[dN_{2,t+\tau}dN_{1,t}]}{(dt)^2} - \lambda_1\lambda_2 & \frac{\mathbb{E}[dN_{2,t+\tau}dN_{2,t}]}{(dt)^2} - \lambda_2^2 \end{pmatrix}. \tag{B.3}$$

Then $\mathbb{E}\left[dX_{i,t}\right]/dt = \mu_i + \lambda_i\mathbb{E}\left[Z_i\right]$, for $i = 1,2$, and

$$\begin{aligned} &\begin{pmatrix} \frac{\mathbb{E}[dX_{1,t+\tau}dX_{1,t}]}{(dt)^2} - \left(\frac{\mathbb{E}[dX_{1,t}]}{dt}\right)^2 & \frac{\mathbb{E}[dX_{1,t+\tau}dX_{2,t}]}{(dt)^2} - \frac{\mathbb{E}[dX_{1,t}]}{dt}\frac{\mathbb{E}[dX_{2,t}]}{dt} \\ \frac{\mathbb{E}[dX_{2,t+\tau}dX_{1,t}]}{(dt)^2} - \frac{\mathbb{E}[dX_{1,t}]}{dt}\frac{\mathbb{E}[dX_{2,t}]}{dt} & \frac{\mathbb{E}[dX_{2,t+\tau}dX_{2,t}]}{(dt)^2} - \left(\frac{\mathbb{E}[dX_{2,t}]}{dt}\right)^2 \end{pmatrix} \\ &= \begin{pmatrix} (\mathbb{E}[Z_1])^2 V_{1,1,N}(\tau) & \mathbb{E}[Z_1]\mathbb{E}[Z_2]V_{1,2,N}(\tau) \\ \mathbb{E}[Z_1]\mathbb{E}[Z_2]V_{2,1,N}(\tau) & (\mathbb{E}[Z_2])^2 V_{2,2,N}(\tau) \end{pmatrix}, \qquad \tau > 0; \end{aligned} \tag{B.4}$$

with

$$\lambda_1 = \frac{\lambda_{1,\infty}\alpha_1(\alpha_2-\beta_{22})+\lambda_{2,\infty}\alpha_2\beta_{12}}{(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})-\beta_{12}\beta_{21}},$$
$$\lambda_2 = \frac{\lambda_{2,\infty}\alpha_2(\alpha_1-\beta_{11})+\lambda_{1,\infty}\alpha_1\beta_{21}}{(\alpha_1-\beta_{11})(\alpha_2-\beta_{22})-\beta_{12}\beta_{21}};$$

and

$$\begin{gathered}
V_{1,1,N}(\tau) = \Big((1+e^{r\,\tau})\ \Big(2\,\beta_{11}\,\beta_{12}\,\beta_{21}\,\alpha_1\,\lambda_1 + 2\,\beta_{11}\,\alpha_1\,(\beta_{22}-\alpha_2)\,(-\beta_{22}+\alpha_1+\alpha_2)\,\lambda_1 \\
+\beta_{12}^3\,\beta_{21}\,\lambda_2 - \beta_{12}^2\,\big(\beta_{21}^2\,\lambda_1 + (-(\beta_{22}\,\alpha_1) + \beta_{11}\,(\beta_{22}-\alpha_2) + \alpha_2\,(\alpha_1+\alpha_2))\,\lambda_2\big)\Big)\,r \\
+2\,e^{\frac{r\,\tau}{2}}\ \big(\beta_{11}^2\,(-(\beta_{12}\,\beta_{21}) + (\beta_{22}-\alpha_2)\,(\beta_{11}+\beta_{22}-3\,\alpha_1-\alpha_2))\ \ \lambda_1\,r\,\cosh\left(\tfrac{r\,\tau}{2}\right) \\
+(\beta_{11}+\beta_{22}-\alpha_1-\alpha_2)\ \Big(2\,\beta_{11}\,(\beta_{12}\,\beta_{21}\,(2\,\beta_{22}+\alpha_1-2\,\alpha_2) \\
+\alpha_1\,(\beta_{22}-\alpha_2)\,(\beta_{22}+\alpha_1-\alpha_2))\ \ \lambda_1 - \beta_{11}^2\,(\beta_{12}\,\beta_{21} + (\beta_{22}-\alpha_2)\,(\beta_{22}+3\,\alpha_1-\alpha_2))\,\lambda_1 \\
+\beta_{11}^3\,(\beta_{22}-\alpha_2)\,\lambda_1 - \beta_{11}\,\beta_{12}^2\,(\beta_{22}-\alpha_2)\,\lambda_2 \\
+\beta_{12}\,\big(-(\beta_{21}\,(3\,\beta_{12}\,\beta_{21}+4\,\alpha_1\,(\beta_{22}-\alpha_2))\,\lambda_1\ ) \\
+\beta_{12}\,\big(\beta_{12}\,\beta_{21}+\beta_{22}\,\alpha_1-\alpha_1\,\alpha_2+\alpha_2^2\big)\,\lambda_2\big)\Big)\ \ \sinh\left(\tfrac{r\,\tau}{2}\right)\big)\Big) \\
/\Big(4\,e^{\frac{(-\beta_{11}-\beta_{22}+\alpha_1+\alpha_2+r)\,\tau}{2}}\ \big(-(\beta_{12}\,\beta_{21}) + (\beta_{11}-\alpha_1)\,(\beta_{22}-\alpha_2)\big)\,(\beta_{11}+\beta_{22}-\alpha_1-\alpha_2)\,r\Big);
\end{gathered}$$

$$\begin{gathered}
V_{1,2,N}(\tau) = \Big((-1-e^{r\,\tau})\ \Big(\beta_{21}\,(\beta_{12}\,\beta_{21}\,\alpha_1 - \beta_{11}\,(\beta_{12}\,\beta_{21}+2\,\alpha_1\,(\beta_{22}-\alpha_2)))\ \ \lambda_1 \\
+\beta_{12}\,\big((\beta_{22}-2\,\alpha_1)\,(\beta_{12}\,\beta_{21}+\beta_{22}\,\alpha_1) \\
-(\beta_{12}\,\beta_{21}+2\,(\beta_{22}-\alpha_1)\,\alpha_1)\,\alpha_2 + 2\,\alpha_1\,\alpha_2^2 \\
+\beta_{11}\,\big(2\,\beta_{12}\,\beta_{21}-\beta_{22}^2+2\,\beta_{22}\,(2\,\alpha_1+\alpha_2)-2\,\alpha_2\,(2\,\alpha_1+\alpha_2)\big)\big)\,\lambda_2\Big)\,r \\
-2\,e^{\frac{r\,\tau}{2}}\ \big(\beta_{11}^2\,(\beta_{22}-\alpha_2)\,(\beta_{21}\,\lambda_1-2\,\beta_{12}\,\lambda_2)\,r\,\cosh\left(\tfrac{r\,\tau}{2}\right) \\
+\Big(\beta_{11}\,\beta_{21}\,\big(\beta_{12}\,\beta_{21}\,(3\,\beta_{22}+2\,\alpha_1-3\,\alpha_2) \\
+2\,\alpha_1\,(\beta_{22}-\alpha_2)\,(\beta_{22}+\alpha_1-\alpha_2)\ \big)\,\lambda_1 - \beta_{12}\,\beta_{21}^2\,(2\,\beta_{12}\,\beta_{21}-\alpha_1\,(-3\,\beta_{22}+\alpha_1+3\,\alpha_2))\,\lambda_1 \\
-\beta_{11}\,\beta_{12}\,\big(\beta_{22}^3+\beta_{12}\,\beta_{21}\,(\beta_{22}+4\,\alpha_1-\alpha_2)-6\,\alpha_1^2\,\alpha_2+2\,\alpha_2^3+2\,\beta_{22}\,\alpha_1\,(3\,\alpha_1+2\,\alpha_2) \\
-\beta_{22}^2\,(2\,\alpha_1+3\,\alpha_2)\big)\,\lambda_2 + \beta_{12}\,\big(2\,\beta_{12}^2\,\beta_{21}^2 \\
+\beta_{12}\,\beta_{21}\,\big(\beta_{22}^2+\beta_{22}\,(\alpha_1-2\,\alpha_2)+(\alpha_1-\alpha_2)\,(2\,\alpha_1+\alpha_2)\big) \\
+\alpha_1\,\big(\beta_{22}^3+2\,\beta_{22}\,\alpha_1\,(\alpha_1+\alpha_2)-2\,(\alpha_1-\alpha_2)\,\alpha_2\,(\alpha_1+\alpha_2)-\beta_{22}^2\,(\alpha_1+3\,\alpha_2)\big)\big)\,\lambda_2 \\
+\beta_{11}^3\,(\beta_{22}-\alpha_2)\,(\beta_{21}\,\lambda_1-2\,\beta_{12}\,\lambda_2) \\
+\beta_{11}^2\,\big(-(\beta_{21}\,(\beta_{12}\,\beta_{21}+(\beta_{22}-\alpha_2)\,(\beta_{22}+3\,\alpha_1-\alpha_2)\ )\ \lambda_1) \\
+\beta_{12}\,\big(2\,\beta_{12}\,\beta_{21}-\beta_{22}^2-6\,\alpha_1\,\alpha_2+2\,\beta_{22}\,(3\,\alpha_1+\alpha_2)\big)\,\lambda_2\big)\ \Big)\,\sinh\left(\tfrac{r\,\tau}{2}\right)\big)\Big) \\
/\Big(4\,e^{\frac{(-\beta_{11}-\beta_{22}+\alpha_1+\alpha_2+r)\,\tau}{2}}\ \big(-(\beta_{12}\,\beta_{21}) + (\beta_{11}-\alpha_1)\,(\beta_{22}-\alpha_2)\big)\,(\beta_{11}+\beta_{22}-\alpha_1-\alpha_2)\,r\Big);
\end{gathered}$$

where

$$r := \sqrt{\beta_{11}^2 + 4\beta_{12}\beta_{21} - 2\beta_{11}\left(\beta_{22}+\alpha_1-\alpha_2\right) + \left(\beta_{22}+\alpha_1-\alpha_2\right)^2},$$

and $\tau > 0$. The expressions for $V_{2,2,N}(\tau)$ and $V_{2,1,N}(\tau)$ are obtained from these by interchanging suffices. Furthermore,

$$\begin{aligned}
\frac{\mathbb{E}\left[(dX_{i,t})^2\right]}{dt} - \frac{(\mathbb{E}[dX_{i,t}])^2}{dt} &= \theta_i + \lambda_i \mathbb{E}\left[Z_i^2\right], \qquad i = 1,2; \\
\frac{\mathbb{E}[dX_{1,t}dX_{2,t}]}{dt} - \frac{\mathbb{E}[dX_{1,t}]\mathbb{E}[dX_{2,t}]}{dt} &= \rho\sqrt{\theta_1\theta_2}.
\end{aligned} \tag{B.5}$$

B.2. Interval-Based Moment Conditions: Bivariate Mutually Exciting Jumps

Let $0 \leq s_1 < s_2 < s_3 < s_4$ and let $\Delta_1 := s_2 - s_1$, $\Delta_2 := s_4 - s_3$ and $\tau := s_3 - s_1$. Write

$$\lambda_i := \mathbb{E}\left[\lambda_{i,t}\right], \; i = 1, 2, \quad V_N(\tau) := \begin{pmatrix} \frac{\mathbb{E}[dN_{1,t+\tau} dN_{1,t}]}{(dt)^2} - \lambda_1^2 & \frac{\mathbb{E}[dN_{1,t+\tau} dN_{2,t}]}{(dt)^2} - \lambda_1 \lambda_2 \\ \frac{\mathbb{E}[dN_{2,t+\tau} dN_{1,t}]}{(dt)^2} - \lambda_1 \lambda_2 & \frac{\mathbb{E}[dN_{2,t+\tau} dN_{2,t}]}{(dt)^2} - \lambda_2^2 \end{pmatrix}. \tag{B.6}$$

Then $\mathbb{E}\left[\int_{s_1}^{s_2} dX_{i,t}\right] / \Delta_1 = \mu_i + \lambda_i \mathbb{E}\left[Z_i\right]$, for $i = 1, 2$, and

$$\begin{pmatrix} \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_{1,u} \int_{s_1}^{s_2} dX_{1,t}\right]}{\Delta_1 \Delta_2} - \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_{1,u}\right]}{\Delta_2} \frac{\mathbb{E}\left[\int_{s_1}^{s_2} dX_{1,t}\right]}{\Delta_1} & \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_{1,u} \int_{s_1}^{s_2} dX_{2,t}\right]}{\Delta_1 \Delta_2} - \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_{1,u}\right]}{\Delta_2} \frac{\mathbb{E}\left[\int_{s_1}^{s_2} dX_{2,t}\right]}{\Delta_1} \\ \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_{2,u} \int_{s_1}^{s_2} dX_{1,t}\right]}{\Delta_1 \Delta_2} - \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_{2,u}\right]}{\Delta_2} \frac{\mathbb{E}\left[\int_{s_1}^{s_2} dX_{1,t}\right]}{\Delta_1} & \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_{2,u} \int_{s_1}^{s_2} dX_{2,t}\right]}{\Delta_1 \Delta_2} - \frac{\mathbb{E}\left[\int_{s_3}^{s_4} dX_{2,u}\right]}{\Delta_2} \frac{\mathbb{E}\left[\int_{s_1}^{s_2} dX_{2,t}\right]}{\Delta_1} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{(\mathbb{E}[Z_1])^2 I^{\tau}_{1,1,N}(\tau,\Delta_1,\Delta_2)}{\Delta_1 \Delta_2} & \frac{\mathbb{E}[Z_1]\mathbb{E}[Z_2] I^{\tau}_{1,2,N}(\tau,\Delta_1,\Delta_2)}{\Delta_1 \Delta_2} \\ \frac{\mathbb{E}[Z_1]\mathbb{E}[Z_2] I^{\tau}_{2,1,N}(\tau,\Delta_1,\Delta_2)}{\Delta_1 \Delta_2} & \frac{(\mathbb{E}[Z_2])^2 I^{\tau}_{2,2,N}(\tau,\Delta_1,\Delta_2)}{\Delta_1 \Delta_2} \end{pmatrix}; \tag{B.7}$$

with

$$I^{\tau}_{1,1,N}(\tau, \Delta_1, \Delta_2) := \int_{u=s_3}^{s_4} \int_{t=s_1}^{s_2} V_{1,1,N}(u-t) du dt,$$
$$I^{\tau}_{1,2,N}(\tau, \Delta_1, \Delta_2) := \int_{u=s_3}^{s_4} \int_{t=s_1}^{s_2} V_{1,2,N}(u-t) du dt.$$

The expressions for $I^{\tau}_{2,2,N}$ and $I^{\tau}_{2,1,N}$ are obtained from these by replacing $V_{1,1,N}$ and $V_{1,2,N}$ by $V_{2,2,N}$ and $V_{2,1,N}$, respectively. Explicit expressions for $V_{i,j,N}(\tau)$ have been derived earlier; see Section B.1. Furthermore,

$$\frac{\mathbb{E}\left[\left(\int_{s_1}^{s_2} dX_{i,t}\right)^2\right]}{\Delta_1} - \frac{\left(\mathbb{E}\left[\int_{s_1}^{s_2} dX_{i,t}\right]\right)^2}{\Delta_1} = \theta_i + \lambda_i \mathbb{E}\left[Z_i^2\right] + \frac{2(\mathbb{E}[Z_i])^2 I_{i,i,N}(\Delta_1)}{\Delta_1}, \qquad i = 1, 2; \tag{B.8}$$

$$\frac{\mathbb{E}\left[\int_{s_1}^{s_2} dX_{1,t} \int_{s_1}^{s_2} dX_{2,u}\right]}{\Delta_1} - \frac{\mathbb{E}\left[\int_{s_1}^{s_2} dX_{1,t}\right] \mathbb{E}\left[\int_{s_1}^{s_2} dX_{2,u}\right]}{\Delta_1} = \rho\sqrt{\theta_1 \theta_2} + \frac{\mathbb{E}[Z_1]\mathbb{E}[Z_2]\left(I_{1,2,N}(\Delta_1) + I_{2,1,N}(\Delta_1)\right)}{\Delta_1}; \tag{B.9}$$

with

$$I_{i,i,N}(\Delta_1) := \int_{u=s_1}^{s_2} \int_{t=s_1}^{u} V_{i,i,N}(u-t) du dt, \qquad i = 1, 2,$$
$$I_{1,2,N}(\Delta_1) := \int_{u=s_1}^{s_2} \int_{t=s_1}^{u} V_{1,2,N}(u-t) du dt.$$

The expression for $I_{2,1,N}$ is obtained by replacing $V_{1,2,N}$ in the above expression by $V_{2,1,N}$.