Expected Shortfall and Portfolio Management in Contagious Markets

Alice Buccioli[a], Thomas Kokholm[b], and Marco Nicolosi[c]

[a]Aarhus BSS, Aarhus University, Department of Economics and Business Economics, Denmark
email: abuccioli@econ.au.dk

[b]Aarhus BSS, Aarhus University, Department of Economics and Business Economics, Denmark
Corresponding author, email: thko@econ.au.dk

[c]University of Perugia, Department of Economics, Italy
email: marco.nicolosi@unipg.it

Abstract

We study the impact of market contagion on portfolio management. To model possible recurrence in the arrival of extreme events, we equip classic Poisson jumps with long memory via past-weighted randomization of the likelihood of their occurrences (Hawkes processes). Within this framework, we tackle the problem of optimal portfolio selection in terms of Expected Shortfall (ES). We use the generalized method of moments to estimate the model on three US stock indexes, representing three major sectors of the US economy. The moment conditions of the model are computed efficiently in closed form applying a novel technique. Given parameter estimates we maximize (at a monthly frequency in the period 2001-2016) the expected return subject to a constraint on ES of a portfolio consisting of the three US sector indexes. We find that the weights of the optimal portfolio are significantly adjusted when the level of contagion is high. Finally, we perform an extensive out-of-sample back-test of the model's ability to measure ES and find that the Hawkes jump-diffusion model outperforms two traditional models that are commonly implemented.

JEL classification: C58; G11; G32.

Keywords: Hawkes process; Contagion; Expected shortfall; Back-testing; Portfolio management.

Contents

1 Introduction

The cascade of negative shocks to financial markets experienced in 2008-2009 points to the interdependence of risks embedded in asset returns. A prominent example is given by the collapse of Lehman Brothers in September 2008, which spread to markets worldwide in the form of contagious shocks. Following these events, an increasing body of literature has been dedicated to the empirical documentation of the contagious nature of financial markets, see for example Longin and Solnik (2001); Bae et al. (2003); Aït-Sahalia et al. (2015). In particular for portfolio management, neglecting to take this contagious behavior into account will lead to an underestimation of the risk of realizing severe losses on the constituents of the portfolio over a short horizon. For a well-diversified portfolio, a single big negative movement in one of the assets will most likely not impact the performance severely. However, if the negative shock is followed by multiple negative shocks to other assets in the portfolio, the loss incurred may be substantial. Both for proper risk management and for portfolio allocation, it is paramount to use a model which realistically features this interconnected jump behavior of financial assets. To address the possible recurrence of jump events, we consider a multidimensional jump-diffusion framework where the jump processes are equipped with long memory via past-weighted randomization of their intensity. This is attained by modeling the arrival of jumps through a Hawkes (1971) process, instead of the more classic constant intensity Poisson process. In this way, the underlying components of the portfolio are affected by two sources of excitation: (i) "Self-contagion", a self-feeding effect of downside shocks ramping up the intensity of future negative jumps, and (ii) "Cross-contagion", transmission of adverse shocks across different assets.

In a theoretical paper, Aït-Sahalia and Hurd (2016) study the problem of optimal portfolio allocation in models where the available asset prices are subject to mutually exciting jump processes. For the case of a log-utility investor, they are able to solve the problem in closed form. In contrast, we study the problem of an investor who wishes to maximize the expected return of a portfolio subject to a constraint on Expected Shortfall (ES). This optimization problem is relevant for financial institutions as capital requirements imposed by regulation are computed based on a tail risk measure, i.e., Value at Risk (VaR) or ES (Basel, 1996; Basel, 2013). Hence, the equity capital of a financial institution translates into a constraint on ES. Moreover, it is of importance for banks and other major financial institutions to measure their portfolio tail risk adequately, as regulation also requires banks to backtest the performance of their models. If the bank incurs more severe losses over time than predicted

by the model, it runs the risk of being subject to an increased capital requirement. As a last resort the authorities have the mandate to force the bank to use a standard model to measure the portfolio risk. Besides leading to a higher capital requirement, it may also have negative reputational consequences for a bank when this happens.

Utility-based portfolio allocation in markets where assets are allowed to jump has been studied in a number of papers: Liu and Loewenstein (2013) study optimal portfolio choice in a model that allows for regime switching following market crashes. In Kallsen (2000), the optimal allocation of wealth is derived under the assumption that exponential Lévy processes drive the asset prices. Liu et al. (2003) add stochastic volatility to the asset dynamics and let the jump intensities scale with the level of volatility. Das and Uppal (2004); Aït-Sahalia et al. (2009) study optimal portfolio choice in models with simultaneous jumps across the assets. Branger et al. (2009) analyze the optimal portfolio choice of a Constant Relative Risk Aversion (CRRA) investor in a market with two risky assets and with contagious states of the economy, modeled as a Markov chain. Branger et al. (2014) study the optimal portfolio decision of an investor with CRRA in a model with partial information about contagion risk. Our paper contributes by considering Hawkes process driven jumps and the use of ES together with expected return as objective criteria. Portfolio selection based on general spectral risk measures is studied in the papers Adam et al. (2008); Brandtner (2013). In a related paper, Buccioli and Kokholm (2018) study a particular portfolio allocation strategy, Constant Proportion Portfolio Insurance, under the assumption of Hawkes processes driving the dynamics of asset returns.

Figure 1 depicts the log-returns of the three indexes NYSE Arca Natural Gas Index (XNG), the Morgan Stanley Technology Index (MSH), and the NYSE Arca Biotechnology Index (BTK) over the week running from Tuesday, October 21 to Monday, October 27, 2008. On Tuesday, both the XNG and the MSH indexes experienced a big negative shock to their value, which was followed by a very big negative shock to the XNG index of more than 10% on Wednesday, while the indexes MSH and BTK also incurred losses around 4%. On the following two days, the index values remained relatively calm with small losses to all of them by the end of the week. Finally, both the XNG and BTK indexes opened the following week with heavy losses of more than 5%. In a traditional jump diffusive model with constant jump intensity, the probability of these occurrences is essentially zero. This and similar contagion episodes have motivated the research focus to enhance traditional models that are unable to embed this feature of financial assets. Empirical support of contagious assets dynamics is provided

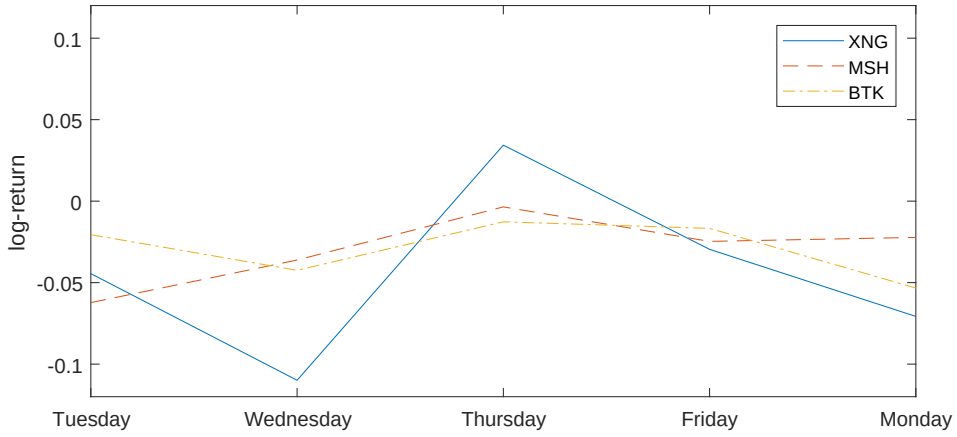

Figure 1: Historical log-returns of the three indexes: XNG, MSH, and BTK over the week from October 21-October 27, 2008.

in Aït-Sahalia et al. (2015), where the contagion is modeled by the introduction of mutually-exciting jumps in the underlying dynamics via Hawkes processes. As opposed to more traditional Lévy-driven models exhibiting constant jump intensity, Hawkes processes embed the risk of jump clustering. Hawkes processes have been applied in a financial context in a number of papers: for credit modeling in Errais et al. (2010); Aït-Sahalia et al. (2014), in Kokholm (2016); Du and Luo (2017) for derivatives modeling, and to model VaR in a univariate setting in Chavez-Demoulin et al. (2005); Herrera and Schipp (2013). The availability of high-frequency data has been the driver of research devoted to the modeling of market price microstructure using Hawkes processes: Bowsher (2007) proposes a model for transaction times and prices at high frequency, Chavez-Demoulin and McGill (2012) apply Hawkes processes to model intraday VaR, and in Bacry et al. (2013) multivariate Hawkes processes are applied in a high frequency setting to model order books. The reader is referred to Hawkes (2018) for a more comprehensive literature review of the applications of Hawkes processes.

Under the framework outlined above, we tackle the problem of optimal portfolio selection in a more realistic model for jump arrivals in asset returns, incorporating self- and cross-excitation. As optimization criteria, we maximize the expected portfolio return subject to a constraint on the ES, as opposed to the "standard" objective functions studied in Aït-Sahalia and Hurd (2016). In general, supervisory authorities advocate the use of ES over VaR for risk management (Basel, 2013), as the former risk measure has better features, e.g. ES is coherent and captures the whole tail of the distribution (Acerbi and Tasche, 2002). We use the generalized method of moments (GMM) to estimate the model on three US stock indexes (XNG, MSH, BTK), representing three major sectors of the US

economy. The moment conditions of the model are computed efficiently in closed form by performing a double-stage version of the methodology of Cuchiero et al. (2012), which allows us to bypass the unobservability of the intensity process. A similar GMM-based estimation technique relying on Taylor aproximations of the moments is developed by Aït-Sahalia et al. (2015).

Given parameter estimates, we optimize (at a monthly frequency in the period 2001-2016) a portfolio consisting of the three indexes mentioned above. We untangle linkages between the level of portfolio concentration implied by the estimated series of optimal portfolio weights and the contagion, as measured by the model, between the three assets. We find that a clear positive relationship exists between the degree of contagion inside a portfolio and the level of wealth concentration among its components. Specifically, in the presence of highly contagious assets, minimal loss allocation calls for investment into the asset displaying the least individual and contagion risk. Moreover, we characterize the dynamic behavior of the optimal portfolio weights, by linking the investors' reactions to changes in the jump intensities. Additionally, we document that the optimal portfolio corresponding to the Hawkes process driven model leads to lower realized losses above the 95% quantile compared to the realized losses associated with the optimal portfolios of two benchmark models commonly implemented, i.e., multivariate versions of the Black and Scholes (1973) model and the Merton (1976) jump-diffusion model. Finally, we analyze the Hawkes process driven model from a risk management perspective and perform an extensive out-of-sample back-test of the portfolio VaR and ES implied by the estimated models. We find that the Hawkes jump-diffusion model outperforms the two benchmark models. Specifically, following the traffic-light back-testing procedure (Basel, 1996; Basel, 2013), we document that in comparison to the benchmark models, the Hawkes process driven model leads to less loss violations belonging to the "yellow" and "red" zones. Hence, in the more distressed periods of time the Hawkes driven model provides higher and more precise estimates of the tail risk measures. Additionally, in the calm periods the Hawkes driven model leads to tail risk measures of the same magnitude as the benchmark models. This is important as banks would be reluctant to use risk models that overestimate the risk in calm periods as this would also imply higher capital requirement.

The remainder of the paper is organized as follows: Section 2 presents the modeling setup. In Section 3 the optimization problem is described and various case studies are considered. The estimation procedure is detailed in Section 4 and implemented on a historical data set containing time series of

prices on three major US sector indexes. In Section 5, the estimated model is used to construct time series of the optimal weights in a portfolio consisting of the three indexes. With the weights at hand, we study the relation between the level of contagion and the concentration in the portfolio. In Section 6, the model is back-tested on historical data and its performance compared to the performance of more traditional models. Finally, Section 7 concludes.

2 Modeling contagion

This section introduces the models utilized in the analysis of the paper. Section 2.1 presents the contagion modeling framework, while Section 2.2 introduces multivariate versions of two classic models. The two latter models will be used in the analysis in order to compare the performance of the Hawkes driven model to more standard models.

2.1 The n-HJD model

To address possible recurrence in the arrival of jump-related events, we consider a multidimensional jump-diffusion framework where classic Poisson jumps are equipped with long memory via past-weighted randomization of their intensity. In this way, the underlying components of the portfolio are affected by two sources of excitation: (i) "Self-contagion", a self-feeding effect of big movements ramping up the intensity of future jumps; (ii) "Cross-contagion", transmission of shocks across different assets.

This is attained by modeling the arrival of jumps through a Hawkes (1971) process. Technically, for a given filtered probability space $\left(\Omega, \mathcal{F}^H, \mathbb{P}, \left(\mathcal{F}_t^H\right)\right)$ verifying the usual hypotheses, an $n-$dimensional Hawkes process $H_t = \left(H_t^{(1)}, \ldots, H_t^{(n)}\right)^\top$ is a *point process* with stochastic intensity, i.e.,

$$\begin{aligned}
\mathbb{P}\left(H_{t+\Delta t}^{(i)} - H_t^{(i)} = 1 \mid \mathcal{F}_t^H\right) &= \lambda_t^{(i)} \Delta t + o(\Delta t) \\
\mathbb{P}\left(H_{t+\Delta t}^{(i)} - H_t^{(i)} > 1 \mid \mathcal{F}_t^H\right) &= o(\Delta t)
\end{aligned}$$

for each i. The intensity process, $\lambda_t = \left(\lambda_t^{(1)}, \ldots, \lambda_t^{(n)}\right)^\top$, defines the expected number of jumps per unit of time and is the solution of

$$\lambda_t = \bar{\lambda} + \int_{-\infty}^{t} \kappa(t-u)\, dH_u\,, \tag{1}$$

where $\bar{\lambda} \in \mathbb{R}^n$, κ is a $n \times n$ matrix-valued kernel and the integral is understood in the sense of vector

stochastic integration. Notably, the dynamics of λ is determined by the paths of H, hereby introducing a feedback effect of jumps on the jump intensities. We choose an exponential kernel matrix $\kappa(x) = \left(\kappa^{(i,j)}(x)\right)_{ij}$ whose components are exponential decay functions of the form

$$\kappa^{(i,j)}(x) = \epsilon^{(i,j)} e^{-\delta^{(i)} x},$$

with $\epsilon^{(i,j)}$ and $\delta^{(i)}$ being positive constants for all $i, j = 1, \ldots, n$. This choice ensures the Markovianity of the Hawkes process coupled with its intensity, $\mathbb{H} := (H, \lambda)$, and of the intensity process λ alone. The intensity dynamics in (1) can be reformulated in a mean-reverting form as follows

$$d\lambda_t = \delta \circ \left(\bar{\lambda} - \lambda_t\right) dt + \epsilon \, dH_t, \tag{2}$$

where $\delta = \left(\delta^{(1)}, \ldots, \delta^{(n)}\right)^\top$, $\epsilon = \left(\epsilon^{(i,j)}\right)_{ij}$, and $\circ$ denotes the element-wise multiplication. We further assume that the intensity process is stationary, which translates into the exponential kernel satisfying the *stability condition* (Hawkes (1971); Hawkes and Oakes (1974))

$$\delta^{(i)} > \sum_{j=1}^{n} \epsilon^{(i,j)}, \qquad i = 1, \ldots, n. \tag{3}$$

In this setting, let S be an n-dimensional vector of stock prices. Under $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$, with $\mathcal{F}^H \subset \mathcal{F}$, the following dynamics are assumed for the log-price $X = \log(S)$,

$$dX_t = \mu \, dt + \sigma \, dW_t + dJ_t^H, \tag{4}$$

where $\mu \in \mathbb{R}^n$, $\sigma = \operatorname{diag}\left(\sigma^{(1)}, \ldots, \sigma^{(n)}\right)$, W is an n-dimensional Brownian motion with correlation matrix $\rho = \left(\rho^{(ij)}\right)_{ij}$, and $J^H \in \mathbb{R}^n$ is a purely jumping vector process of the form $J_t^{H,(i)} = \sum_{j=1}^{H_t^{(i)}} \mathtt{J}_j^{(i)}$, with $\mathtt{J}^{(i)} = \left\{\mathtt{J}_j^{(i)}\right\}_{j \geq 1}$ being a sequence of mutually independent random variables, and independent of $H^{(i)}$, with common law $f_{\mathtt{J}}^{(i)} \sim \mathcal{N}(m_J^{(i)}, s_J^{(i)})$. Furthermore, the vector of Brownian motions W and the vector of marked Hawkes processes J^H are assumed to be mutually independent. This model is labeled n-HJD for short. Remarkably, this model belongs to the class of affine processes (Errais et al. (2010); Kokholm (2016)), which allows for any estimation technique building on the knowledge of the characteristic function or, even more interestingly, of the moments (Cuchiero et al., 2012).

2.2 Benchmark models

Multivariate extensions of the Black and Scholes (1973) model and the Merton (1976) jump-diffusion model are embedded in the n-HJD framework. More precisely, in the multivariate Black and Scholes (1973) model the log-price process X is given by

$$dX_t = \mu\, dt + \sigma\, dW_t\,, \tag{5}$$

where $\mu \in \mathbb{R}^n$, $\sigma = \operatorname{diag}\left(\sigma^{(1)}, \ldots, \sigma^{(n)}\right)$, and W is an n-dimensional Brownian motion with correlation matrix $\rho = \left(\rho^{(ij)}\right)_{ij}$. This model is labeled n-N. In the multivariate Merton (1976) jump-diffusion model, an independent jump component is added to the Black-Scholes model, and the log-price process X is given by

$$dX_t = \mu\, dt + \sigma\, dW_t + dJ_t^N\,, \tag{6}$$

where $J^N \in \mathbb{R}^n$ is a purely jumping vector process of the form $J_t^{N,(i)} = \sum_{j=1}^{N_t^{(i)}} \mathtt{J}_j^{(i)}$, with N being an n-dimensional Poisson process with (fixed) intensity $\lambda \in \mathbb{R}^n$ and $\mathtt{J}^{(i)} = \left\{\mathtt{J}_j^{(i)}\right\}_{j\geq 1}$ a sequence of mutually independent random variables, and independent of $N^{(i)}$, with common law $f_{\mathtt{J}}^{(i)} \sim \mathcal{N}(m_J^{(i)}, s_J^{(i)})$. Furthermore, the vector of Brownian motions W and the vector of compound Poisson processes J^N are assumed to be mutually independent. This model is labeled n-CPJD.

3 Portfolio management in contagious markets

For the purpose of portfolio selection, it is paramount to capture the contagious nature of financial markets. For a well-diversified portfolio, a single big negative movement in one of the assets will most likely not impact the performance severely. However, if the negative shock is followed by multiple negative shocks to other assets in the portfolio, the loss incurred may be substantial. In Section 3.1, we present the general framework to maximize expected return over some time horizon subject to an ES constraint. This optimization problem is relevant for financial institutions as capital requirements imposed by regulation are computed based on a tail risk measure, i.e., VaR or ES. We also formulate the closely related problem of minimizing ES subject to a constraint on expected return. Section 3.2 is devoted to analyzing three cases with varying degrees of contagion across a market of 1 risk-free and 10 risky assets. We illustrate how the optimal portfolio weights depend heavily on the

contagion among the assets.

3.1 ES-based optimization problems

Consider a market of n risky assets $S = \left(S^{(1)}, \ldots, S^{(n)}\right)^\top$ and the risk-free bank account B. Let V_t denote the time-t value of a portfolio consisting of a position in b_t units of the risk-free asset B and $a_t = \left(a_t^{(1)}, \ldots, a_t^{(n)}\right)^\top$ shares of the stocks S,

$$V_t = a_t^\top S_t + b_t B_t.$$

We consider a time period τ and assume that the portfolio composition stays fixed over this time frame. We define the relative loss of the portfolio over the time horizon τ as $L_{t,t+\tau} := -\frac{V_{t+\tau}-V_t}{V_t}$, and denote by $R_{t,t+\tau}$ the expected rate of return of the portfolio during a τ-period.

The ES at level α of the loss $L_{t,t+\tau}$ is defined as

$$\mathrm{ES}_{\alpha,t} = \mathbb{E}[L_{t,t+\tau} | L_{t,t+\tau} > \mathrm{VaR}_{\alpha,t}]\,, \tag{7}$$

where $\mathrm{VaR}_{\alpha,t}$ is the VaR at confidence level α of the losses between time t and $t+\tau$, namely the quantile of order α of the portfolio's loss distribution.

We consider the following two optimization problems:

1. At each time t, select the optimal portfolio allocation vector $\left(a_t^*, b_t^*\right)$ that maximizes the expected rate of return over a time horizon τ subject to a constraint $\mathcal{C}$ on the ES, i.e.

$$\left(a_t^*, b_t^*\right) = \underset{(a,b)\in\mathcal{C}}{\arg\max}\, R_{t,t+\tau}\left(a,b\right)\,, \tag{P1}$$

   where

$$\mathcal{C} := \left\{(a,b) : \mathrm{ES}_{\alpha,t}\left(L_{t,t+\tau}(a,b)\right) \le e\right\}, \quad e \in \mathbb{R}.$$

2. At each time t, select the optimal portfolio allocation vector $\left(a_t^*, b_t^*\right)$ that minimizes the ES of the loss over a time horizon τ subject to a constraint $\mathcal{C}$ on the expected portfolio return, i.e.

$$\left(a_t^*, b_t^*\right) = \underset{(a,b)\in\mathcal{C}}{\arg\min}\left\{\mathrm{ES}_{\alpha,t}\left(L_{t,t+\tau}\left(a,b\right)\right)\right\}\,, \tag{P2}$$

where

$$\mathcal{C} := \left\{(a,b) : R_{t,t+\tau}(a,b) \geq r\right\}.$$

The two optimization problems can be solved by the risk-minimizing technique of Rockafellar and Uryasev (2000) and the follow-up work by Krokhmal et al. (2002). The two problems are equivalent in the sense that they give rise to the same efficient frontier (Krokhmal et al., 2002).[1] The general approach embodied in Rockafellar and Uryasev (2000) relies on the knowledge (either in analytic or simulation-based form) of the probability density function driving the risk of the portfolio. Accordingly, the algorithm calls for efficient calculation of the probability density function of the log-prices. This is a delicate task in the n-HJD modeling framework, particularly when increasing the dimension of the system. In our study, we rely on the simulation-based version of the optimization technique of Rockafellar and Uryasev (2000). The efficiency of this technique can be notably improved when combined with the exact simulation algorithm for Hawkes processes developed by Dassios and Zhao (2013).

3.2 Case studies

In this section, we take a market of 10 risky assets and a single risk-free asset as given and consider an investor who wishes to maximize the expected return of a portfolio over a horizon τ equal to one month and subject to a constraint on the ES at the 95% level equal to 6%.[2] Via 1 million simulations, we analyze three case studies with increasing degrees of contagion across the assets:

(a) *No cross contagion*: As base case we take a symmetric market with asset dynamics as in equation (4) and parameters, $\sigma^{(i)} = \sigma = 0.15$, $\rho^{(i,j)} = 0, i \neq j$, $\delta^{(i)} = 120$, $\bar{\lambda}^{(i)} = 2$, $\epsilon^{(i,i)} = 60$, $\epsilon^{(i,j)} = 0$ for all $i, j = 1, \ldots, 10$ and $i \neq j$. The jumps in log-prices share the same mean $m_J = -0.05$ and standard deviation $\sigma_J = 0.05$. The jump arrival intensities are initialized at the long-run mean with $\lambda_0^{(i)} = \frac{\delta^{(i)}\bar{\lambda}^{(i)}}{\delta^{(i)} - \sum_{j=1}^{10}\epsilon^{(i,j)}}$ and the drift $\mu^{(i)} = 0.05 - \lambda_0^{(i)} m_J$. The specific choice of parameters is inspired by those reported in Aït-Sahalia et al. (2015). The risk-free rate is assumed equal to zero.

(b) *One primary asset*: The same parameters as in the base case but with $\epsilon^{(i,1)} = 15, i > 1$ and $\epsilon^{(i,j)} = 0, j \neq i$ and $i, j > 1$.

[1] The efficient frontier is defined as the selection of portfolios that for varying levels of ES produce the highest expected returns.

[2] Minimization of the ES subject to a constraint on the expected return yields the same optimal portfolio for appropriate level of required expected return.

(c) *One independent asset and three sectors linked internally*: The same parameters as in the base case but with $\epsilon^{(i,1)} = 0, i > 1, \epsilon^{(i,j)} = 20, j \neq i$ and $j,i = 2,3,4$, $\epsilon^{(i,j)} = 15, j \neq i$ and $j,i = 5,6,7$, $\epsilon^{(i,j)} = 10, j \neq i$ and $j,i = 8,9,10$. The remaining cross contagion parameters are set equal to 0.

Figure 2 depicts the loss distributions associated with the optimal portfolio in the three cases. The contagion primarily impacts the left tail of the loss distributions, while the right tails are approximately the same, due to the identical constraint on ES in all three scenarios. Moreover, from Table 1 we observe how the expected return of the optimal portfolio is decreasing in the overall level of contagion in the market. Focusing only on the allocation to the risk-free asset, we notice from Figure 3 that the weight goes from highly negative in case (a) to highly positive in case (c). This observation explains the return pattern reported in Table 1. Due to the constraint on ES, when the level of risk in the market is high (as measured by the contagion parameters), more wealth is allocated to the risk-free asset, resulting in a smaller expected return. In the symmetric market with no cross contagion, Figure 3 reveals that the optimal portfolio is composed of a leveraged position in the 10 risky assets with equal weight. The part of the wealth allocated to the risk-free bond is directly related to the constraint on the ES. Setting the ES higher will shift investments in the risk-free bond to the risky assets and vice versa. In case (b) where the negative spillover from jumps is single-way from asset 1 to the rest of the market, we see that the optimal portfolio involves a short position in asset 1 and long positions in the other 9 risky assets. The short position in asset 1 diversifies the risk of the portfolio, since a negative jump to asset 1 will increase the risk of additional shocks to the other 9 risky assets. Hence, in this case the investor gains on the short position but with increased probability of realizing a loss on the long positions. In the final case, we see that the investor is long in all the assets. The most cross contagious assets, 2-4, are held in the smallest portions in the optimal portfolio with increasing sizes of the positions for assets 5-7 and 8-10. The biggest investment is allocated to the independent risky asset, 1. The examples illustrate that it is fundamental to capture the contagious links between the assets, when selecting the weights in the optimal portfolio.

Table 1: Expected returns at a horizon of one month and with constrained 95%-ES equal to 6% in the three cases (a)-(c).

Case	(a)	(b)	(c)
Expected Return	90 bp	84 bp	64 bp

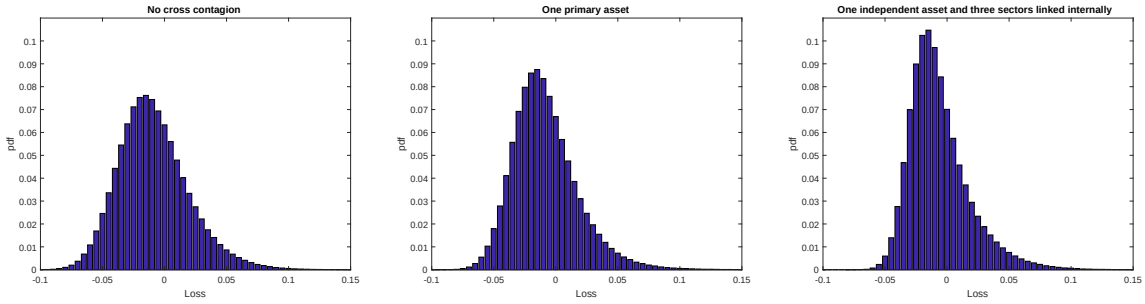

Figure 2: The loss distributions of the $ES_{95\%}$ constrained and expected return maximizing portfolios in the three separate cases and at a horizon of one month. The distributions are the output of 1 million simulations.

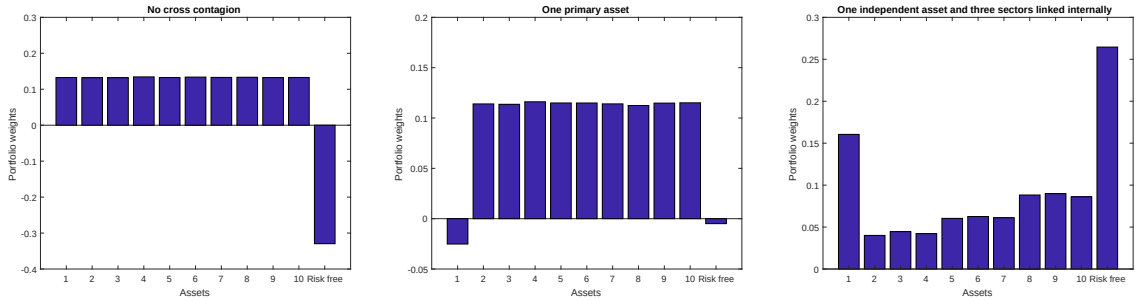

Figure 3: The optimal portfolio weights for varying degrees of cross contagion across the assets.

4 Model estimation

This Section is divided into two parts: first, in Section 4.1 the general approach to estimate the n-HJD model is presented, while Section 4.2 presents the data and the estimation of the three-dimensional Hawkes driven model.

4.1 Estimation procedure of the n-HJD model

For the estimation of the n-HJD model, we use GMM, where the moments are computed in closed-form using a double-stage version of the methodology of Cuchiero et al. (2012). The technique provides an efficient way to compute polynomial moments up to order k in closed-form for *polynomial processes*, that is, Markov processes whose infinitesimal generator maps polynomials of degree k into polynomials of degree less or equal to k. We apply this procedure to the process of log-returns which are stationary when the model condition (3) holds. Specifically, let $Y_{t,t+\tau} := X_{t+\tau} - X_t$

be the vector of log-returns. Then,

$$Y_{t,t+\tau} = Y^D_{t,t+\tau} + Y^J_{t,t+\tau}\,, \tag{8}$$

where $Y^D_{t,t+\tau} \sim \mathcal{N}(\mu\tau, \sigma^2\tau)$ and $Y^J_{t,t+\tau} = \int_t^{t+\tau} dJ^H_u$. The process $\left(Y^J_{t,t+s}, \lambda_{t+s}\right)$ proves to be a polynomial process adapted to the filtration $(\mathcal{F}_{t+s})$, which implies that the expected value of any polynomial of $\left(Y^J_{t,t+s}, \lambda_{t+s}\right)$, conditioned at time t, is a polynomial in s and in the values of the process at time t, $\left(Y^J_{t,t}, \lambda_t\right)$.[3] The coefficients of this polynomial can be computed by exponentiation of a matrix A, which is deduced from the infinitesimal generator of $\left(Y^J, \lambda\right)$, denoted by $\mathcal{A}$, as follows

$$\mathcal{A}e_i = \sum_{j=1}^{K} A_{ij} e_j\,,$$

where $\{e_1, \ldots, e_K\}$ is a basis of the linear space of all polynomials of degree less than or equal to k. In the first stage, we apply the method of Cuchiero et al. (2012) to $\left(Y^J_{t,t+\tau}, \lambda_{t+\tau}\right)$, and we find that, for any polynomial function f,

$$\mathbb{E}\left[f\left(Y^J_{t,t+\tau}, \lambda_{t+\tau}\right) \mid \mathcal{F}_t\right] = \left(\phi_1, \ldots, \phi_K\right) e^{\tau A} \left(e_1, \ldots, e_K\right)^\top{}_{|\lambda_t}\,, \tag{9}$$

where $\phi := \left(\phi_1, \ldots, \phi_K\right)^\top$ is the vector of the coordinates of f in the basis $\{e_1, \ldots, e_K\}$, i.e. $f = \sum_{j=1}^K \phi_j e_j$. This means that all the (mixed) moments of order k are polynomial functions in λ_t. As the intensity process is not directly observable, we apply the rule of iterated expectations

$$\mathbb{E}\left[f\left(Y^J_{t,t+\tau}, \lambda_{t+\tau}\right)\right] = \mathbb{E}\left[\mathbb{E}\left[f\left(Y^J_{t,t+\tau}, \lambda_{t+\tau}\right) \mid \mathcal{F}_t\right]\right]\,,$$

where the inner expectation is given by (9) and the outer expectation is computed again through the technique of Cuchiero et al. (2012) (second stage), being the result of the inner expectation a polynomial function of λ_t. This results in an efficient algorithm for the calculation of all the (mixed) moments of the jump process $Y^J_{t,t+\tau}$, which are polynomials in the initial value of the intensity of the log-return processes, λ_0. Finally, the moments of $Y_{t,t+\tau}$ can be easily deduced from the binomial formula. A similar GMM-based estimation technique is developed by Aït-Sahalia et al. (2015), where the moments are calculated through iterated evaluations of the infinitesimal generator. In the bivariate case $n = 2$, they provide explicit expressions of the moments up to the second order. Higher-order

[3] Note that $Y^J_{t,t} = 0$.

moments are approximated as Taylor expansions in τ truncated after the second order term. Our methodology distinguishes itself from Aït-Sahalia et al. (2015) as all moments of $Y_{t,t+\tau}$ are known analytically, up to a matrix exponential, and do not require any approximation in order to be evaluated efficiently. In our implementation, when $n = 3$, the procedure proves to be efficient and operational also for high-order moments.

4.2 Data description and 3-HJD GMM estimation

The model estimation and the empirical analysis in this paper are based on a sample of three major US sector indexes.[4] We are interested in capturing self- and mutual contagion between different US industry sectors. Accordingly, we consider three non-overlapping stock indexes that track large companies in three leading US sectors: the commodity, the technology, and the biotechnology industries. As representative index of the commodity sector, we consider the XNG index, which measures the stock performance of a basket of large companies that are active in the US natural gas industry. To represent the technology sector, we select the MSH index, which measures the performance of a cross section of highly capitalized US companies operating in the broad technology sector. Finally, we choose the BTK index as market indicator of the US biotechnology industry. The data sample consists of a 15-year time series of daily closing prices of the three indexes from January 2, 2001 to April 29, 2016. The historical quotes are collected from the OptionMetrics database. We estimate the 3-HJD model using GMM on the standardized time series of (daily) log-returns. We consider standardized data in order to put moments of different orders on the same scale of magnitude.

We adopt a two-stage estimation procedure: first, we match the model-implied expressions for the mean and the standard deviation of the standardized log-return processes to 0 and 1, respectively. This delivers parametric formulas for the parameter vectors μ and s_J

$$\begin{aligned}
\mu^{(i)} &= -\mu_\lambda^{(i)} m_J^{(i)} \\
s_J^{(i)} &= \sqrt{\left(1-2\mu^{(i)}\tau^2\mu_\lambda^{(i)}m_J^{(i)}-{m_J^{(i)}}^2\mu_\lambda^{(i)}\tau-{\sigma^{(i)}}^2\tau-{\mu^{(i)}}^2\tau^2\right)/\mu_\lambda^{(i)}\tau}, \qquad i = 1,2,3\,,
\end{aligned}$$

where $\mu_\lambda^{(i)} := \lim_{t\to+\infty}\mathbb{E}\left[\lambda_t^{(i)}\right] = \frac{\delta^{(i)}\bar{\lambda}^{(i)}}{\delta^{(i)}-\sum_{j=1}^n \epsilon^{(i,j)}}$. Then, in a second stage, we estimate the remaining pa-

[4] The portfolio selection procedure detailed in Section 3.1 is feasible in real time for a much higher number of assets. However, we restrict the analysis to three assets for two reasons: first, the concept of financial contagion makes most sense at a sector level. Second, the model estimation becomes heavy with more assets due to the number of parameters involved. For the same reason, we also constrain the model from containing stochastic volatility.

rameters using GMM, and we obtain estimates for μ and s_J accordingly. Since jump intensities are latent processes and due to the high dimensionality of the model, we need to impose some parameter restrictions to facilitate the estimation and minimize problems of parameter identification. To reduce the dimension of the parameter set, we set $\rho^{(i,j)} = \bar{\rho}\, c_{ij}$, where c_{ij} is the empirical correlation between log-returns of asset i and asset j, and we estimate $\bar{\rho}$ so that $\rho^{(i,j)} \in [-1,1]$ for any i, j. To reduce problems of identification, we set $\bar{\lambda}^{(i)} = \kappa \mu_\lambda^{(i)}$, which combined with the stability condition (3) leads to $\kappa \in (0,1)$ and the following parametrization for the speed of mean reversion in the intensity dynamics, $\delta^{(i)} = \frac{\sum_{j=1}^{n} \epsilon^{(i,j)}}{(1-\kappa)}$. Finally, we set $\lambda_0 = \mu_\lambda$, and we assume that the cross-sectional excitation is symmetric, which translates into the matrix ϵ being symmetric. This leaves us with 17 parameters to be estimated using GMM,

$$\Theta = \left\{ m_J^{(1)}, m_J^{(2)}, m_J^{(3)}, \sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}, \mu_\lambda^{(1)}, \mu_\lambda^{(2)}, \mu_\lambda^{(3)}, \epsilon^{(1,1)}, \epsilon^{(2,2)}, \epsilon^{(3,3)}, \epsilon^{(1,2)}, \epsilon^{(1,3)}, \epsilon^{(2,3)}, \bar{\rho}, \kappa \right\}.$$

To identify the parameters of the 3-HJD model, we use $\left[2n + n(\text{lags of autocovariance}) + n + 2n\right]$ moment conditions on the standardized log-returns, with $n = 3$, corresponding to:

- marginal moments of order 3 and 4: $\mathbb{E}\left[{Y_{t,t+\tau}^{(i)}}^{k} \right]$, for $k = 3,4$ and $i = 1,2,3$,
- autocorrelations of lags $l = 1,2,3$: $\mathbb{E}\left[Y_{t,t+\tau}^{(i)} Y_{t+l\tau,t+(l+1)\tau}^{(i)} \right]$, for $i = 1,2,3$,
- correlations of order 1 and mixed order 1 and 2: $\mathbb{E}\left[{Y_{t,t+\tau}^{(i)}}^{k_1} {Y_{t,t+\tau}^{(j)}}^{k_2} \right]$, for $i, j = 1,2,3$, $i \neq j$, $k_1, k_2 > 0$, and $k_1 + k_2 \leq 3$.

The estimation results are reported in Table 2. The large estimates for the diagonal elements of ϵ reveal a significant level of self-excitation in all the three US industry sectors considered. Also, our estimated values for the off-diagonal elements of ϵ are relatively high, providing evidence that the three industry sectors are contagiously linked. Specifically, we find comparable levels of relatively high contagion between the commodity and the biotechnology sectors and between the biotechnology and the technology sectors. In contrast, a smaller estimate for $\epsilon^{(\text{XNG,MSH})}$ $\left(\text{and } \epsilon^{(\text{MSH,XNG})}\right)$ is indicative of a more moderate inter-connection between the commodity and the technology sectors. A more uniform correlation structure is produced by the continuous part of the model as revealed by

the estimated values for ρ. By reversion of the constraints listed above, the full set of parameters for the 3-HJD model is derived and summarized in Table 3.

Table 2: GMM parameter estimates for the 3-HJD model. Standard errors are in parentheses.

Index	m_J	σ	μ_λ	$\epsilon^{(i,\text{XNG})}$	$\epsilon^{(i,\text{MSH})}$	$\epsilon^{(i,\text{BTK})}$	$\bar{\rho}$	κ
XNG	-0.003	0.251	11.813	65.183	14.328	21.144	0.963	0.082
	(0.000)	(0.001)	(0.036)	(0.089)	(0.024)	(0.025)	(0.024)	(0.015)
MSH	-0.017	0.263	6.946	14.328	75.198	21.328	0.963	0.082
	(0.001)	(0.001)	(0.023)	(0.024)	(0.053)	(0.056)	(0.024)	(0.015)
BTK	-0.007	0.294	4.644	21.144	21.328	78.624	0.963	0.082
	(0.001)	(0.002)	(0.020)	(0.025)	(0.056)	(0.064)	(0.024)	(0.015)

Table 3: Estimated 3-HJD parameters.

Index	μ	σ	δ	$\bar{\lambda}$	$\epsilon^{(i,\text{XNG})}$	$\epsilon^{(i,\text{MSH})}$	$\epsilon^{(i,\text{BTK})}$	$\rho^{(i,\text{MSH})}$	$\rho^{(i,\text{BTK})}$	m_J	s_J	λ_0
XNG	0.080	0.251	109.67	0.972	65.183	14.328	21.144	0.500	0.421	-0.003	0.052	11.813
MSH	0.153	0.263	120.79	0.571	14.328	75.198	21.328		0.641	-0.017	0.025	6.946
BTK	0.145	0.294	131.95	0.382	21.144	21.328	78.624			-0.007	0.037	4.644

5 Optimal portfolio selection: empirical analysis

In this section we study the composition of the optimal portfolio over time on the data introduced in Section 4.2. First, in Section 5.1 we infer the paths of the jump intensities of the three assets based on observed log-returns. Next, with the intensity time series at hand we study, in Section 5.2, the optimization problems (P1)-(P2) and construct time series of the optimal weights of a portfolio consisting of the three indexes. For the optimization problem (P2) we put no constraint on the expected return, since we wish to study the portfolio with minimal ES. In Section 5.3 we analyze the impact of contagion on portfolio choice. In particular, we focus on the relationship between the level of contagion and the concentration in the optimal portfolios studied in Section 5.2. Finally, in Section 5.4 we compare the losses implied by the portfolio corresponding to the 3-HJD model with the losses associated with the benchmark models introduced in Section 2.2.

5.1 Inferring the time series of jump intensities

The 3-HJD model is estimated on the full data set as described in Section 4, and the paths of the intensities are backed out from observed movements in the daily asset log-returns using equation

(2). We detect jumps in the assets on days when the absolute value of the log-returns is greater than a number of standard deviations of the diffusive component, i.e.

$$|Y_{t,t+\tau}| > c\sigma\sqrt{\tau}\,, \tag{10}$$

where $c \in \mathbb{R}_{>0}$ is a cut-off parameter. For a given c, let $n_c = \left(n_c^{(1)},\ldots,n_c^{(n)}\right)$ and $m_c = \left(m_c^{(1)},\ldots,m_c^{(n)}\right)$ denote the vectors containing the total number of detected asset jumps normalized by the length (in years) of the sample period T and the corresponding mean of the absolute value of the jump sizes, respectively. Let $\bar{n} = \left(\bar{n}^{(1)},\ldots,\bar{n}^{(n)}\right)$ denote the model-implied expected number of jumps per year and let $\bar{m} = \left(\bar{m}^{(1)},\ldots,\bar{m}^{(n)}\right)$ be the model-implied mean vector of the absolute values of the jump sizes. The optimal cut-off level, c^*, is found such that

$$c^* = \arg\min_c \varphi(c)\,, \tag{11}$$

where

$$\varphi(c) = \sqrt{\sum_{i=1}^{n}\left(\frac{n_c^{(i)}}{m_c^{(i)}} - \frac{\bar{n}^{(i)}}{\bar{m}^{(i)}}\right)^2}\,. \tag{12}$$

Applied to the full set of observed daily log-returns, the minimal distance in (11) is obtained when $c = 2.17$ (see Figure 4).

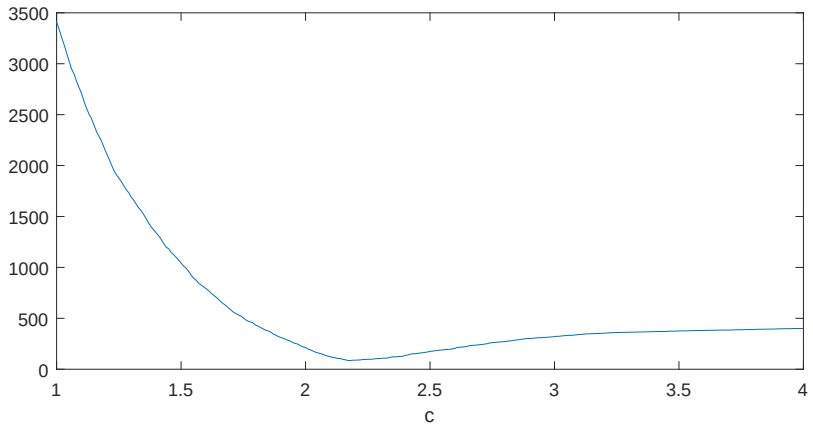

Figure 4: Plot of φ, defined in (12), as a function of the cut-off parameter c.

Jump detection on the basis of large realized returns standardized by a measure of the variation due to the continuous part of the process is a natural criterion, see e.g. Lee and Mykland (2008).

Similar threshold-based jump detection schemes, pioneered by Mancini (2001, 2009), are standard in the literature on jump regression, see e.g. Li et al. (2017) and the references therein. In order to limit the error induced by discrete sampling, we assume that the jump occurs halfway in the interval between the daily sampling points. Following this procedure gives us time series of the Hawkes process H_t, which coupled with equation (2) and the parameters estimated allows us to construct the time series of the asset jump intensities depicted in Figure 5. The Figure clearly illustrates the increased jump risk following the dot-com crash of 2000-2002 and the financial crisis 2008-2009, where the intensities spike to high levels for a sustained period of time.

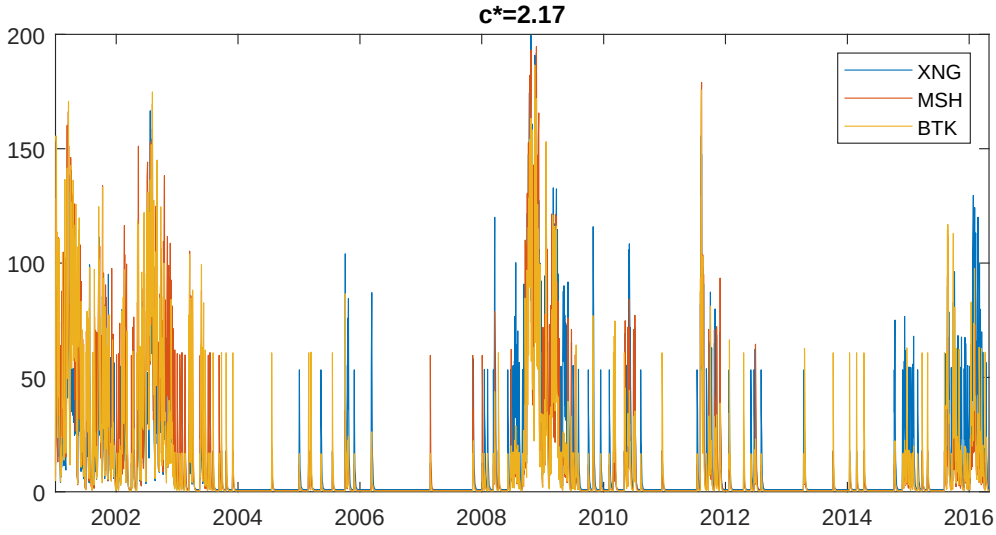

Figure 5: Estimated time series of the three intensities, filtered from the time series of asset returns when $c = 2.17$.

5.2 The dynamic behavior of the optimal portfolio weights

With the intensities at hand and the model estimated, we employ the portfolio optimization detailed in Section 3. Take a tenor grid $\mathcal{T} = \{T_0, T_1, \ldots, T_N\}$ spanning the period from January 2001 to April 2016, with $T_i - T_{i-1} = \tau$ equal to one month, and consider an investor who has the three sector indexes XNG, MSH, and BTK available to invest in. We divide the analysis of the portfolios inferred by the optimization problems (P1) and (P2) into Sections 5.2.1 and 5.2.2, respectively.

5.2.1 Portfolio dynamics implied by (P1)

Over each interval in the tenor grid, the investor wishes to maximize the expected portfolio return subject to a constraint on the ES of the portfolio at the 95% level equal to 20%.[5] The resulting recursive optimization problem entails, at each time step $t \in \mathcal{T}$, the resolution of the sub-problem of risk minimization (P1) given in Section 3.

Figure 6: Time series of weights resulting from (P1), with τ=1 month, α = 95%, and ES constraint $\mathrm{ES}_{\alpha,t}\left(L_{t,t+\tau}\right) \leq 20\%$.

Figure 6 depicts the dynamics of the weights in the portfolio. Two observations are striking: first, in the calm periods with low levels of intensities the XNG index is shortened to gear up the position in the BTK index, while a small fraction is allocated to the MSH index. Inspection of Table 3 reveals that the XNG index has the lowest drift of the three assets. Even if it has lower diffusive volatility and the lowest diffusive correlation with the other two assets, the asset is still shortened in calm periods. The MSH and BTK indexes have drift parameters of the same magnitude, but the BTK index has a lower mean jump size than the MSH index. In the calm periods with low levels of intensities, the risk mainly stems from the diffusive behavior and few but big jumps being realized. Hence, the optimal allocation calls for a levered position in the BTK index while the MSH has a weight slightly greater than zero. Second, in distressed periods the allocations to the BTK and the MSH indexes

[5] Similar to the case studies of Section 3.2, one could also add a risk-free asset to the investment set. However, as this addition does not change the analysis qualitatively, we avoid doing it here.

decrease while the allocation to the XNG index increases. The XNG index has both the smallest self-contagion parameter and the smallest mean jumps size of -30 bp, which is significantly lower than the same parameter of the other two assets. These characteristics impact the weight in the XNG index positively in periods of distress.

5.2.2 Portfolio dynamics implied by (P2)

At each time step in the monthly tenor grid, we solve the problem (P2) with no constraint on the expected return. Figure 7 depicts the dynamics of the weights in the portfolio that minimizes the ES at the 95% level over the next month. In contrast to Figure 6, the XNG index is now allocated the

Figure 7: Time series of weights resulting from (P2), with τ=1 month, α = 95%, and no return constraints.

biggest share of wealth on most of the days. When the expected return is no longer a criterion, the low mean jump size of the XNG index results in a larger share invested in the index. In the calm periods, the MSH and the BTK indexes have roughly the same weight. However, in more distressed periods the BTK index increases in weight while the MSH index decreases by roughly the same amount. When intensities are high, the speed of the mean reversion δ and the mean size of jumps m_J play an important role. Since the BTK index, in comparison to the MSH index, has the highest speed of reversion coupled with the lowest value of the mean jump size, the weight is increased in this index. Following the same reasoning, the weight on the MSH decreases considerably due to its more negative jump

sizes coupled with its relatively slower conversion of the jump intensity to normal levels.

5.3 The concentration of the optimal portfolio

In the sequel, we focus on untangling possible linkages between the risk-minimizing capital allocation and the dynamic dependence across the portfolio's assets. To measure the degree of concentration in a given portfolio, we propose a Concentration Index (CI) defined as

$$\mathrm{CI}(t) = \frac{\left\| \omega_t - \mu_t^{\omega} \right\|_2}{K}, \tag{13}$$

where $\omega_t = \left(\omega_t^{(1)}, \ldots, \omega_t^{(n)}\right)^{\top}$ is the vector of portfolio weights, i.e.,

$$\omega_t^{(i)} = \frac{\alpha_t^{(i)} S_t^{(i)}}{V_t^{(i)}}, \qquad i = 1, \ldots, n,$$

and $\mu_t^{\omega} = \frac{1}{n}\sum_{i=1}^{n} \omega_t^{(i)} = \frac{1}{n}$ is the average weight value. K is a normalization constant that ensures that, for a portfolio with no short positions, the index ranges from 0 (*minimal concentration*) to 1 (*maximal concentration*), i.e.

$$K = \left\| \mathrm{e}_1 - \mu_t^{\omega} \right\|_2 = \sqrt{\frac{n-1}{n}},$$

where e_1 stands for the (first) canonical unit vector. The CI is minimal and equal to 0 when the portfolio comprises equally weighted positions and equal to 1 when all wealth is invested in a single asset. The concentration index can be greater than 1 in case the portfolio involves short positions. Other characterizations of concentration relying on entropy measures are often proposed in the literature (Meucci, 2009; Pola, 2016). However, as these are only valid for portfolios with no short positions, we use the measure (13) in our analysis.

Using the CI, we study the relation between the concentration of the capital and the risk structure of the available assets. In view of the decomposition (8), we split the covariance matrix of the log-returns, C, into two independent components. Suppressing the dependency on τ, which we choose equal to one month, and for $t \in \mathcal{T}$ we have

$$C(t) = C_D + C_J(t), \tag{14}$$

where $C_D = \left(c_D^{(i,j)}\right)_{ij}$, $C_J(t) = \left(c_J^{(i,j)}(t)\right)_{ij}$, for $i, j = 1, \ldots, n$.

Note that for the model proposed, the C_D matrix is constant over time, while the levels of the intensities, being the state variables of the model, impact the $C_J(t)$ matrix. Hence, only changes in the $C_J(t)$ matrix can be the driver of changes in the optimal portfolio composition.

To quantify the heterogeneity of the covariance structure of the jump vector Y^J, we propose the following metric

$$\mathcal{R}_J(t) = \left\| \gamma_J(t) - \mu^{\gamma_J}(t) \right\|_2, \qquad \text{for } t \in \mathcal{T}, \tag{15}$$

where $\gamma_J(t) \in \mathbb{R}^n$, with i-th component $\gamma_J^{(i)}(t) = \sum_{j=1}^n c_J^{(i,j)}(t)$, and $\mu^{\gamma_J}(t) = \frac{1}{n}\sum_{i=1}^n \gamma_J^{(i)}(t)$. We remark that $\gamma_J^{(i)}(t)$ is a measure of the riskyness of asset i as it is the sum of both the variance of the jumps of asset i and the covariances of the jumps of the remaining assets with the jumps of asset i. Hence, equation (15) is a measure of the heterogeneity of the riskyness of the jumps in the assets. To identify and quantify the impact of the $\mathcal{R}_J$ factor on the portfolio concentration, we study the following linear regression model

$$\mathrm{CI}(t) = a + b\mathcal{R}_J(t) + \epsilon(t). \tag{R1}$$

The lower panel in Table 4 reports the result of the regression for the time series of the ES minimizing portfolio composed of the three assets depicted in Figure 7. The R-squared value indicates that the linear model (R1) explains almost 90% of the variability of the concentration index. According to the t-statistic results, all the coefficients are significant. The high R-squared value is also revealed upon inspection of the right panel in Figure 8, where it is evident that the regression tracks the concentration index well. Hence, the suggested measure (15) of the heterogeneity of the jump risk seems to capture most of the information embedded in the covariance matrix C_J relevant for computing the concentration of the risk minimizing portfolio. This finding has important consequences for risk management: in a market scenario where one or some of the assets have high levels of jump intensities and are contagiously linked, the optimal portfolio should be more concentrated in assets with less contagious links in order to diversify the tail risk.

Performing the same exercise on the optimal weights depicted in Figure 6 and corresponding to the optimization problem (P1), we get the regression model reported in the top panel of Table 4 with depicted fit to the CI shown in the left panel of Figure 8. We observe how the R-squared value is now at a lower level and equal to 57.5%. Since the optimal portfolio now targets two objectives, namely

high expected return with limited ES, the risk structure of the assets explains a smaller part of the concentration of the portfolio.

Table 4: Parameter estimates for the linear regression model (R1), where CI is computed using the weights resulting from (P1) (top panel) and from (P2) (lower panel).

Optimization problem	Coefficient	Estimate	SE	t-statistic
(P1)	a	1.522	0.015	98.878
	b	-23.415	1.462	-16.017
	Root Mean Squared Error: 0.185			
	R-squared: 0.575			
	Coefficient	Estimate	SE	t-statistic
(P2)	a	0.161	0.003	49.751
	b	11.834	0.307	38.540
	Root Mean Squared Error: 0.039			
	R-squared: 0.887			

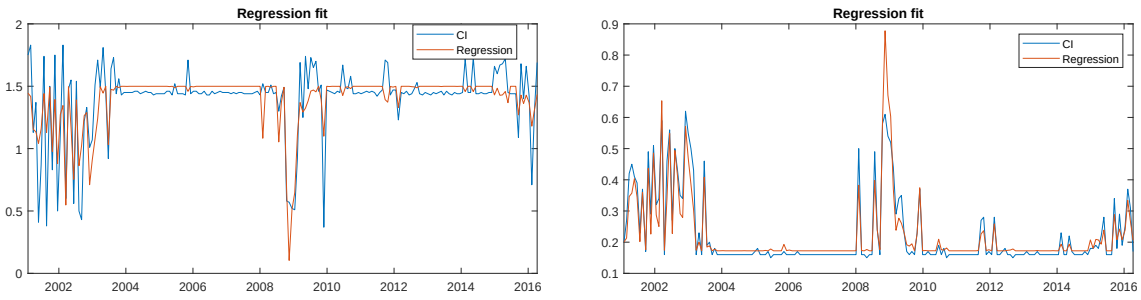

Figure 8: *Left*: Fit of the regression (R1) to the concentration index, when CI is computed using the portfolio weights resulting from (P1). *Right*: Fit of the regression (R1) to the concentration index, when CI is computed using the portfolio weights resulting from (P2).

5.4 Comparison of losses implied by the optimal portfolios

We end this section with a comparison of the biggest losses realized on monthly investment horizons over the period 2001-2016. For the comparison, we estimate the two benchmark models described in Section 2.2. The estimated model parameters are reported in Table 5. Figure 9 depicts the one-month losses exceeding the $\text{VaR}_{95\%}$ over the full sample period of the portfolios implied by the optimal allocation corresponding to the three different models and for the two optimization prob-

lems (P1) and (P2). The losses are arranged in descending order. We observe that for either of the optimization problems (and the associated optimal portfolios) and for essentially all the data points the portfolio corresponding to the Hawkes driven model gives rise to the smallest losses. Hence, the figure indicates that when the Hawkes process driven model is used to deduce the portfolio, the actual realized losses above the 95% quantile are smaller in magnitude than those of the two benchmark models. We believe that this result serves as a convincing argument for the use of a model that incorporates contagion effects.

Table 5: Estimated parameters of the two benchmark models 3-N and 3-CPJD. For the estimation of the 3-CPJD model, we use as initial values the corresponding parameter estimates for the 3-HJD model reported in Table 3. For the intensities, we use as the initial values the long-run mean value $\lim_{t\to\infty}\mathbb{E}\left[\lambda_t^{(i)}\right]$ of the 3-HJD model.

Model	Index: i	μ	σ	$\rho^{(i,\mathrm{MSH})}$	$\rho^{(i,\mathrm{BTK})}$	m_J	s_J	λ
3-N								
	XNG	0.051	0.308	0.519	0.437			
	MSH	0.034	0.275		0.665			
	BTK	0.110	0.305					
3-CPJD								
	XNG	0.052	0.212	0.609	0.512	-0.003	0.052	20.952
	MSH	0.033	0.273		0.780	-0.017	0.025	4e-05
	BTK	0.108	0.295			-0.007	0.037	3.070

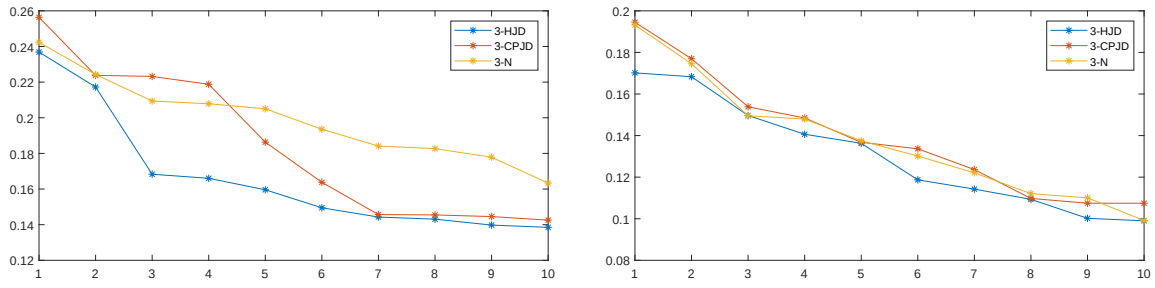

Figure 9: *Left*: Comparison of losses implied by the portfolios' maximizing returns subject to an ES level below 20% (see P1). *Right*: Comparison of losses implied by the ES-minimizing portfolios (see P2). The plot reports one-month losses exceeding the portfolio $\mathrm{VaR}_{95\%}$ implied by the three models and sorted in descending order.

6 Back-testing

This section is devoted to study the n-HJD model from a risk management perspective. In particular, we back-test the performance of the n-HJD model for the computation of tail risk measures. We consider the most widely used quantile-based measures of risk, that is the ES and the VaR, and we compare the model predictions of extreme losses with those obtained using the two benchmark models described in Section 2.2. The self- and cross-contagions modeled by incorporating Hawkes processes in the log-price dynamics, as described in Section 2.1, are expected to provide a more conservative risk measure with better performances in periods of financial crisis. We implement the back-test on the portfolios implied by (P1) and on equally weighted portfolios. The latter choice is included in order to have a clear separation of the risk management performance, from the problem of optimal portfolio allocation.

6.1 Back-testing ES and VaR

To back-test the ES, we follow Acerbi and Székely (2014). Consider a discrete set of dates $t = 1,\dots,T$, where t is measured in units of τ, and at these time points we observe the realized portfolio loss $L_{t,t+\tau}$ between t and $t+\tau$. For instance, if τ is one day, then we have the historical series of the daily portfolio losses. For each $t = 1,\dots,T$, we compute the ES at confidence level α, $\mathrm{ES}_{\alpha,t}$, according to the definition given in Equation (7). Then we define the indicator variable, I, as

$$I_{t+\tau} = \begin{cases} 1 & L_{t,t+\tau} > \mathrm{VaR}_{\alpha,t} \\ 0 & L_{t,t+\tau} \leq \mathrm{VaR}_{\alpha,t} \end{cases}.$$

The indicator $I_{t+\tau}$ is 1 when there is a VaR violation, that is when the realized loss between t and $t+\tau$ is higher than the corresponding VaR computed at time t. The statistics proposed in Acerbi and Székely (2014) to back-test the ES over T observations is

$$Z = \frac{1}{T}\sum_{t=1}^{T} \frac{-L_{t,t+\tau} I_{t+\tau}}{(1-\alpha)\mathrm{ES}_{\alpha,t}} + 1. \tag{16}$$

The null hypothesis assumes that the model predictive distribution of $L_{t,t+\tau}$, P_t (conditionally on the information at time t) correctly predicts the real unknown loss distribution, F_t. Hence

$$H_0 : P_t = F_t\,, \quad \forall\, t.$$

The alternative hypothesis is

$$\begin{aligned} H_1 : \qquad & \mathrm{ES}^F_{\alpha,t} \geq \mathrm{ES}_{\alpha,t} \quad \forall\, t, \text{and strictly greater for some } t \\ & \mathrm{VaR}^F_{\alpha,t} \geq \mathrm{VaR}_{\alpha,t}, \end{aligned}$$

where $\mathrm{ES}^F_{\alpha,t}$ and $\mathrm{VaR}^F_{\alpha,t}$ are the risk measures according to the real unknown distribution F_t. Acerbi and Székely (2014) show that

$$\mathbb{E}_{H_0}[Z] = 0, \quad \text{and} \quad \mathbb{E}_{H_1}[Z] < 0.$$

The above results do not require independence of losses. However, independence of the arrival of tail events, as in losses exceeding the VaR, is assumed. Note that testing H_1 against H_0 is a test in the direction of risk underestimation.

To test H_0 against H_1, first one has to compute Z using the realized loss $L_{t,t+\tau}$, the model implied $\mathrm{ES}_{\alpha,t}$, and the realized violation indicator $I_{t+\tau}$, at all observation points. Then the distribution of Z under H_0 is needed. This is obtained by simulating M independent portfolio losses $L^i_{t,t+\tau}$ from the model, for each $t = 1,\ldots,T$ and $i = 1,\ldots,M$. In each simulation i, Z^i is computed as

$$Z^i = \frac{1}{T}\sum_{t=1}^{T} \frac{-L^i_{t,t+\tau} I^i_{t+\tau}}{(1-\alpha)\mathrm{ES}_{\alpha,t}} + 1, \tag{17}$$

and the tail probability is computed as

$$p\text{-value} = \frac{1}{M}\sum_{i=1}^{M} \mathbb{1}_{\{Z^i < Z\}}.$$

Then the null hypothesis is rejected at level p if

$$p\text{-value} < p.$$

To back-test the VaR, we follow the standard violation-based methodology (see e.g. McNeil et al.

(2015)). By definition of the quantile,

$$\mathbb{E}[I_{t+\tau} \mid \mathcal{F}_t] = 1-\alpha\,,$$

$\{I_{t+\tau}\}_{t=1,\dots,T}$ forms a sequence of i.i.d. Bernoulli random variables with parameter $1-\alpha$ (Christoffersen, 1998). It follows that $\sum_{t=1}^{T} I_{t+\tau}$ is binomially distributed with parameters T and 1-α. To test this hypothesis, we apply the binomial test with statistics

$$Z_{\text{VaR}} = \frac{\sum_{t=1}^{T} I_{t+\tau} - T(1-\alpha)}{\sqrt{T\alpha(1-\alpha)}} \tag{18}$$

and reject the hypothesis at level p if $Z_{\text{VaR}} > n_{1-p}$, where n_{1-p} is the quantile of the standard normal distribution at level $1-p$. Alternative tests include the Markov test of Christoffersen (1998) and the duration-based test of Christoffersen and Pelletier (2004). For a detailed overview of existing methods for back-testing VaR estimates, we refer to Berkowitz et al. (2011).

6.2 Performance based on the expexted return maximizing portfolios

For the back-test, we consider 15 years of daily historical series of prices in the period from January 2, 2001 to April 29, 2016. At each point in time t, we estimate the model using historical data from the previous 3 years. Hence, at each time point t this yields a set of parameters starting from t_0 = January 2, 2004. As we are estimating the model on a shorter window of time (relative to the full sample), we fix the entries of the contagion matrix ϵ to the values estimated based on the full sample and reported in Table 3. Moreover, we follow the procedure described in Section 5.2 and back out the jump intensities from the time series of realized log-returns with the inferred cut-off parameter $c = 2.17$ and by the use of the time-t prevailing $\sigma^{(i)}$ estimates. In line with the regulation imposed on banks, we back-test the models on daily losses realized on the expected return maximizing portfolios with the $ES_{95\%}$ constrained to be less than 5% and on a horizon of 1 day. Hence, at any $t \geq t_0$, we record the realized 1-day portfolio loss, and we calculate the model-implied ES and VaR for a confidence level equal to $\alpha = 95\%$. Specifically, the risk-measure estimates are computed using simulated log-prices over a 1-day horizon. Then the back-test is performed on a rolling window of one year (i.e. $T = 250$ trading days) from January 2, 2005 to the end of the observation period.

According to the Basel traffic light levels (Basel, 1996; Basel, 2013), we consider two different lev-

els of confidence, 95% and 99.99%, which correspond to $p = 5\%$ and $p = 0.01\%$, respectively.[6] If Z is above the 5-th percentile of the simulated Z distribution, then the ES risk measure is in the *green zone*, and the model is not rejected. For a model in the green zone, no regulatory actions are taken. When the observed statistics is between the 0.01-th percentile and the 5-th percentile, then the model could be rejected at the 95% confidence level and the risk measure is in the *yellow zone*. For a model in the yellow zone, regulatory actions would likely lead to an increase in capital requirement as well as additional back-testing documentation, e.g., for risk measures computed at different quantiles α. Finally, when the statistics is below the 0.01-percentile, then the risk measure is in the *red zone*, and the model can be rejected at the 99.99% confidence level. When a model falls into the red zone, the general regulatory response is to assume that the model is flawed. Capital requirements are increased, and the bank is required to improve the model predictions immediately. As a last resort, the bank can be forced to replace its internal risk model with a standard model. The supervisory framework acknowledges that there may be special circumstances in which even an accurate model will enter the red zone. For example, in (Basel, 1996) on page 11 it says: "*...there will on very rare occasions be a valid reason why an accurate model will produce so many exceptions. In particular, when financial markets are subjected to a major regime shift, many volatilities and correlations can be expected to shift as well, perhaps substantially...such a regime shift could generate a number of exceptions in a short period of time. In essence, however, these exceptions would all be occurring for the same reason, and therefore the appropriate supervisory reaction might not be the same as if there were ten exceptions, but each from a separate incident.*"

Figure 10 compares the Z-statistics defined in equation (16) for the Hawkes process driven model (dotted line), the jump-diffusion Merton model (dashed line), and the Black-Scholes model (continuous line). The three different colors indicate whether the risk measure is in the green, the yellow, or the red zone.[7] The Figure clearly reveals that the Hawkes driven model performs better than the benchmark models. Only in the beginning of 2009, corresponding to the period where markets were in severe turmoil, is the Hawkes model in the red zone. This is in contrast to the two other models that enter the red zone for a longer period around the financial crisis and, additionally, also fall in the red category by the end of the sample.

[6] The Basel traffic light framework is based on VaR back-testing. However, the definitions of the green, yellow, and red zones can be readily extended to the ES measure. We perform a formal back-test of the models based on both the ES and the VaR measures.

[7] Figure 14 in Appendix A depicts separately for the three models the Z-statistics together with the 5-th and 0.01-th percentiles of the simulated distributions of Z.

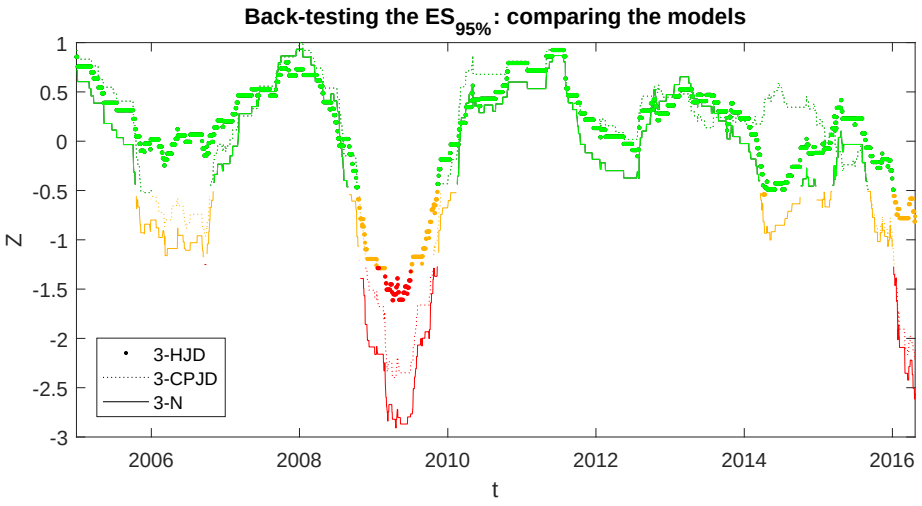

Figure 10: The Z-statistics based on the expected return maximizing portfolios subject to a constraint on the daily ES equal to 5% corresponding to the Hawkes process driven model (dotted line), the jump-diffusion Merton model (dashed line), and the Black-Scholes model (continuous line). The three colors indicate whether the risk measure is in the green, the yellow, or the red zone.

The $\text{VaR}_{95\%}$ back-testing is based on the standard normal quantiles $n_{95\%} = 1.6449$ and $n_{99.99\%} = 3.7190$. Figure 11 shows the statistics Z_{VaR} computed over the previous $T = 250$ days for the three models considered along with the quantiles at the levels 95% and 99.99%, defined by the Basel traffic light levels. Similar to the figure depicting the Z-statistics for the ES, we see that the two benchmark models stay in the red zone for longer periods throughout the sample. Moreover, the benchmark models also enter the yellow zone in periods where the Hawkes driven model stays in the green zone.

6.3 Performance based on the equally weighted portfolios

To have a clear separation of the back-testing of the model from the problem of optimal portfolio selection, we also consider an equally weighted portfolio of the three sector indexes (XNG, MSH, BTK), and backtest the models following the procedure described in Sections 6.1-6.2

Figure 12 shows the realized portfolio loss and the ES estimates calculated using the three different models considered, 3-HJD, 3-CPJD, and 3-N. Two observations are striking: first, the ES implied by the Hawkes driven model reacts immediately to the arrival of new information in the form of the detected jumps in the asset processes. This is in contrast to the other two models where the estimated

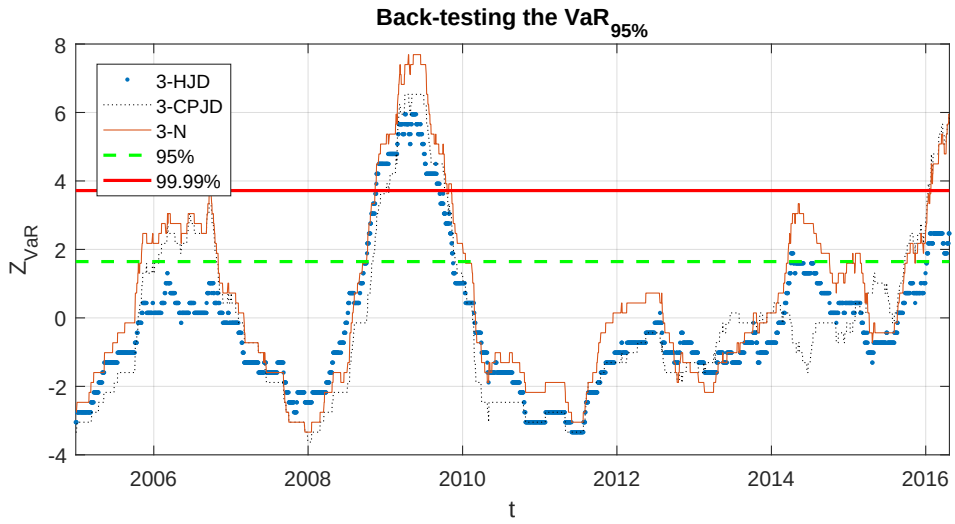

Figure 11: Z_{VaR} statistics computed over the previous 250 days for the Hawkes model (3-HJD, continuous line), the Merton model (3-CPJD,dashed line,) and the Black-Scholes model (3-N,dotted line).

ES slowly adapts to new information. Second, in the calm periods where no jumps are realized, the risk measures of all three models are of the same magnitude.

Big movements in assets prices translate into spikes in the jump intensities which lead to the spikes in the ES of the 3-HJD model observed in Figure 12. It may be problematic for the bank within a short period of time to either readjust the portfolio to achieve a targeted level of ES or to increase its capital buffer in order to match the capital requirement implied by the ES. Hence, for proper risk management, and in line with the counter-cyclical capital buffer regulation implemented after the financial crisis of 2008 (Basel, 2010), the bank should operate with a surplus in the capital during calm periods in order to meet the increase in capital requirement in distressed periods.

A back-test of the three models based on the losses associated with the equally weighted portfolio leads to the same conclusion as in Section 6.2. Figure 13 plots the Z-statistics of the three models and confirms that the Hawkes process driven model outperforms the benchmark models. In fact, the two benchmark models enter the red zone at an even earlier stage before the financial crisis.

In the non-crisis periods, the Z-statistics of all three models follow each other. Hence, the Hawkes driven model does not return higher risk measures *at all the dates*. Only in the periods following "big" movements in prices does the Hawkes driven model produce more conservative risk measures. This is important since the bank has two conflicting incentives: first, it has a regulatory incentive to use a

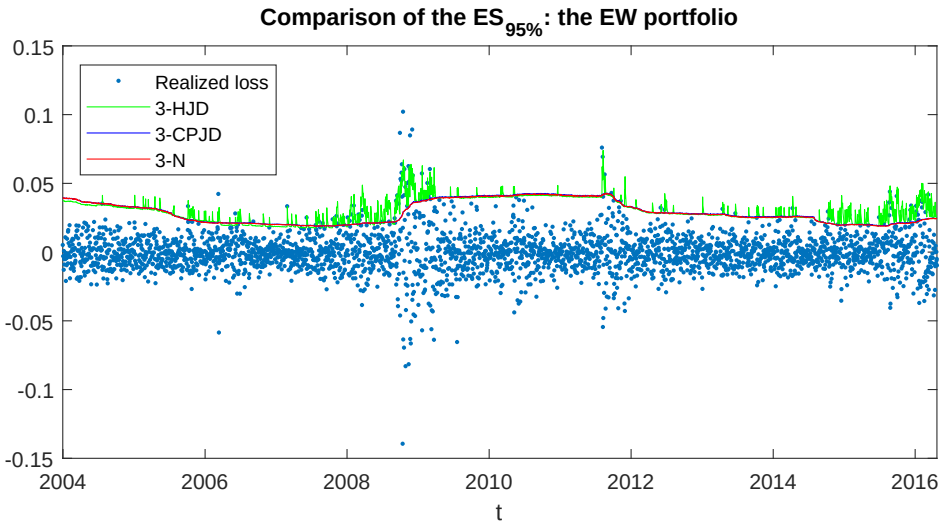

Figure 12: Comparison of the ES for the Hawkes model (3-HJD, continuous line), the Merton model (3-CPJD, dashed line), and the Black-Scholes model (3-N, dotted line). The figure also shows the realized portfolio loss (pointed line).

model that is being labeled in the "green zone" in all market scenarios. Second, it has an incentive to keep the capital requirements at a minimum level as putting aside too much capital is costly (O'Hara and Shaw, 1990; Diamond and Rajan, 2000; Baker and Wurgler, 2015).[8]

7 Conclusion

To model contagion among assets, we consider a multidimensional jump-diffusion framework where classic Poisson jumps are equipped with long memory via past-weighted randomization of their intensity (Hawkes processes). Under this framework, we tackle the problem of expected portfolio return maximization subject to a constraint on the ES. We use the generalized method of moments to estimate the model on three US stock indexes (XNG, MSH, BTK), representing three major sectors of the US economy. Given parameter estimates, we optimize (at a monthly frequency in the period 2001-2016) the portfolio consisting of the three indexes mentioned above. We find that there exists a clear positive relationship between the degree of contagion inside a portfolio and the level of wealth concentration among its components. Moreover, we document that the optimal portfolio corresponding to the Hawkes process driven model leads to lower realized losses above the 95% quantile

[8]It should be noted that the debate is not settled on this claim, see for instance Miles et al. (2013). In any case, it seems to be the common belief in the banking industry that capital is expensive.

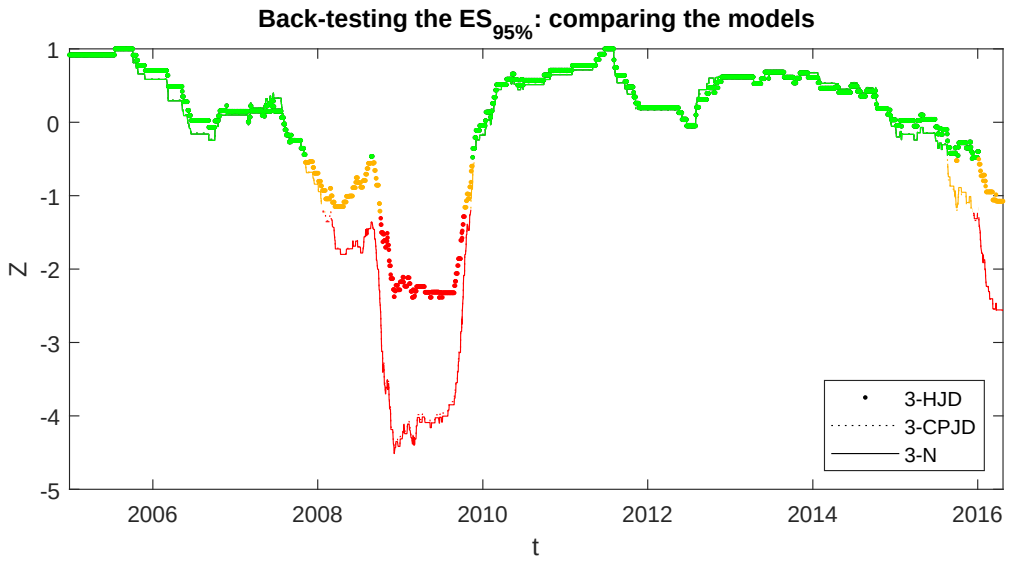

Figure 13: The Z-statistics based on the equally weighted portfolios corresponding to the Hawkes process driven model (dotted line), the jump-diffusion Merton model (dashed line), and the Black-Scholes model (continuous line). The three colors indicate whether the risk measure is in the green, the yellow, or the red zone.

than the realized losses associated with the optimal portfolios of the two benchmark models. Finally, we perform an extensive out-of-sample back-test. We find that the Hawkes jump-diffusion model outperforms two traditional models that are commonly implemented.

Acknowledgments

The authors wish to thank Andrea Barletta, Wolfgang Härdle, Peter Løchte Jørgensen, Cosimo Munari, and Elisa Nicolato for comments. The paper has been presented at the Quantitative Finance Workshop 2018, Rome, at the CENTRAL Workshop, Humboldt University, Berlin, at the 11th International Risk Management Conference 2018, Paris, at the 10th World Congress of the Bachelier Finance Society 2018, Dublin, at the IESEG School of Management finance research seminar series, Paris, at Risk Nordics 2018, Stockholm, and at the Quantitative Methods in Finance 2018, Sydney. Wharton Research Data Services (WRDS) was used in preparing this manuscript. This service and the data available thereon constitute valuable intellectual property and trade secrets of WRDS and/or its third-party suppliers.

References

Acerbi, C. and Székely, B. (2014). Back-testing expected shortfall. *Risk*, 27:76–81.

Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503.

Adam, A., Houkari, M., and Laurent, J.-P. (2008). Spectral risk measures and portfolio selection. *Journal of Banking & Finance*, 32(9):1870–1882.

Aït-Sahalia, Y., Cacho-Diaz, J., and Hurd, T. R. (2009). Portfolio choice with jumps: A closed-form solution. *The Annals of Applied Probability*, 19(2):556–584.

Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3):585–606.

Aït-Sahalia, Y. and Hurd, T. R. (2016). Portfolio choice in markets with contagion. *Journal of Financial Econometrics*, 14(1):1–28.

Aït-Sahalia, Y., Laeven, R. J., and Pelizzon, L. (2014). Mutual excitation in Eurozone sovereign CDS. *Journal of Econometrics*, 183(2):151–167.

Bacry, E., Delattre, S., Hoffmann, M., and Muzy, J.-F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, 13(1):65–77.

Bae, K.-H., Karolyi, G. A., and Stulz, R. M. (2003). A new approach to measuring financial contagion. *The Review of Financial Studies*, 16(3):717–763.

Baker, M. and Wurgler, J. (2015). Do strict capital requirements raise the cost of capital? Bank regulation, capital structure, and the low-risk anomaly. *American Economic Review*, 105(5):315–20.

Basel Committee on Banking Supervision (1996). Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements. Technical report, available at www.bis.org.

Basel Committee on Banking Supervision (2010). Guidance for national authorities operating the countercyclical capital buffer. Technical report, available at www.bis.org.

Basel Committee on Banking Supervision (2013). Fundamental review of the trading book: A revised market risk framework. Technical report, available at www.bis.org.

Berkowitz, J., Christoffersen, P., and Pelletier, D. (2011). Evaluating Value-at-Risk models with desk-level data. *Management Science*, 57(12):2213–2227.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654.

Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2):876–912.

Brandtner, M. (2013). Conditional value-at-risk, spectral risk measures and (non-) diversification in portfolio selection problems–a comparison with mean–variance analysis. *Journal of Banking & Finance*, 37(12):5526–5537.

Branger, N., Kraft, H., and Meinerding, C. (2009). What is the impact of stock market contagion on an investor's portfolio choice? *Insurance: Mathematics and Economics*, 45(1):94–112.

Branger, N., Kraft, H., and Meinerding, C. (2014). Partial information about contagion risk, self-exciting processes and portfolio optimization. *Journal of Economic Dynamics and Control*, 39:18–36.

Buccioli, A. and Kokholm, T. (2018). Constant proportion portfolio insurance strategies in contagious markets. *Quantitative Finance*, 18(2):311–331.

Chavez-Demoulin, V., Davison, A. C., and McNeil, A. J. (2005). Estimating Value-at-Risk: a point process approach. *Quantitative Finance*, 5(2):227–234.

Chavez-Demoulin, V. and McGill, J. (2012). High-frequency financial data modeling using Hawkes processes. *Journal of Banking & Finance*, 36(12):3415–3426.

Christoffersen, P. and Pelletier, D. (2004). Backtesting Value-at-Risk: A duration-based approach. *Journal of Financial Econometrics*, 2(1):84–108.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4):841–862.

Cuchiero, C., Keller-Ressel, M., and Teichmann, J. (2012). Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16(4):711–740.

Das, S. R. and Uppal, R. (2004). Systemic risk and international portfolio choice. *The Journal of Finance*, 59(6):2809–2834.

Dassios, A. and Zhao, H. (2013). Exact simulation of Hawkes process with exponentially decaying intensity. *Electron. Commun. Probab.*, 18(62):1–13.

Diamond, D. W. and Rajan, R. G. (2000). A theory of bank capital. *The Journal of Finance*, 55(6):2431–

2465.

Du, D. and Luo, D. (2017). The pricing of jump propagation: Evidence from spot and options markets. *Management Science*, forthcoming.

Errais, E., Giesecke, K., and Goldberg, L. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665.

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

Hawkes, A. G. (2018). Hawkes processes and their applications to finance: A review. *Quantitative Finance*, 18(2):193–198.

Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11(3):493–503.

Herrera, R. and Schipp, B. (2013). Value at Risk forecasts by extreme value models in a conditional duration framework. *Journal of Empirical Finance*, 23:33–47.

Kallsen, J. (2000). Optimal portfolios for exponential Lévy processes. *Mathematical Methods of Operations Research*, 51(3):357–374.

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

Krokhmal, P., Palmquist, J., and Uryasev, S. (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk*, 4:43–68.

Lee, S. S. and Mykland, P. A. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *The Review of Financial Studies*, 21(6):2535–2563.

Li, J., Todorov, V., and Tauchen, G. (2017). Jump regressions. *Econometrica*, 85(1):173–195.

Liu, H. and Loewenstein, M. (2013). Market crashes, correlated illiquidity, and portfolio choice. *Management Science*, 59(3):715–732.

Liu, J., Longstaff, F. A., and Pan, J. (2003). Dynamic asset allocation with event risk. *The Journal of Finance*, 58(1):231–259.

Longin, F. and Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance*, 56(2):649–676.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian mo-

tion. *Giornale dell'Istituto Italiano degli Attuari*, 64(19-47):44.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2):270–296.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools*. Princeton university press.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144.

Meucci, A. (2009). Managing diversification. *Risk*, 22(5):74–79.

Miles, D., Yang, J., and Marcheggiano, G. (2013). Optimal bank capital. *The Economic Journal*, 123(567):1–37.

O'Hara, M. and Shaw, W. (1990). Deposit insurance and wealth effects: the value of being "too big to fail". *The Journal of Finance*, 45(5):1587–1600.

Pola, G. (2016). On entropy and portfolio diversification. *Journal of Asset Management*, 17(4):218–228.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional Value-at-Risk. *Journal of Risk*, 2:21–42.

A Figures: Back-testing performance of the models based on the expected return maximizing portfolios

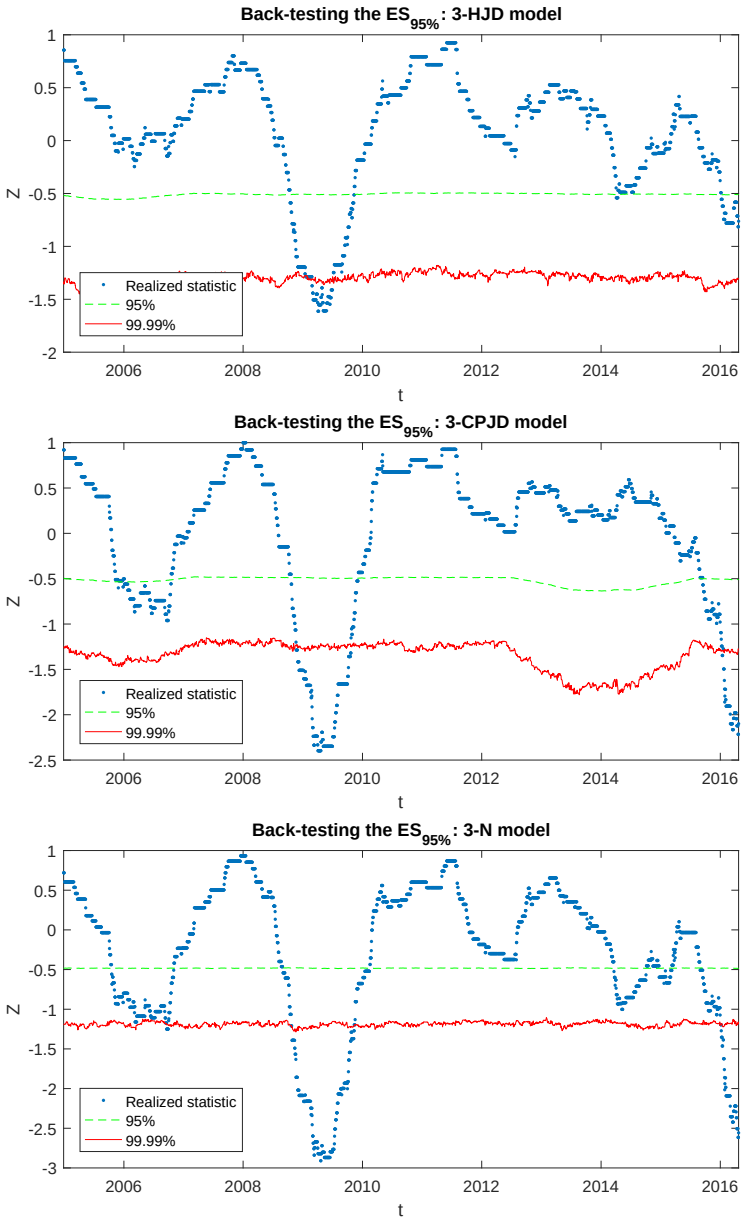

Figure 14: The Z-statistics according to the Hawkes process model (top), the Merton model (middle), and the log-normal model (bottom) together with the 5-th and 0.01-th percentile of the simulated distribution of Z.