Asymmetric excitation and the US bias in Portfolio Choice*

Zhenzhen Fan†, Roger J.A. Laeven‡, Rob van den Goorbergh§

May 25, 2015

Abstract

Literature has found that the large investments in the US cannot be explained by standard portfolio allocation models or diversification motives. In this paper, we explain the overweighting of the US equity in the market portfolio by asymmetric excitation. We employ the mutually exciting jump diffusion with an asymmetric excitation structure to account for the fact that crashes in the US can get reflected quickly in smaller economies but not the other way round. We solve in closed-form the portfolio optimization problem and find that the optimal portfolio is biased towards the US equity market compared to classic portfolio choice models. The US bias comes from the fact that it is more capable of transmitting domestic jumps risks worldwide than other geographical markets, which results in a larger hedging potential in the US equity against state variables. We further show that the welfare loss of the suboptimal strategy that ignores the excitation nature of jumps is substantial. By calibrating the model to historical returns on US, Japanese, and European equity indices, we show that our model is able to reproduce the observed biases in the market portfolio.

Keywords: Portfolio choice; the US bias; International diversification; Asymmetric excitation; Mutually exciting jumps.

JEL classification: G11.

*We are very grateful to Raman Uppal and seminar and conference participants at APG, Tilburg University, the Tinbergen Institute, the University of Amsterdam, the University of Sydney, the 2014 Winter School in Mathematical Finance in Lunteren, the 2014 Insurance: Mathematics and Economics Meeting in Shanghai, and the 2014 Quantitative Methods in Finance Conference in Sydney. This research was funded in part by the Netherlands Organization for Scientific Research (NWO) under grant NWO Vidi 2009 (Laeven).

†University of Amsterdam, Department of Quantitative Economics, the Netherlands. APG Asset Management, the Netherlands. Email: z.fan@uva.nl.

‡University of Amsterdam, Department of Quantitative Economics, the Netherlands. Tilburg University, CentER, the Netherlands. Email: r.j.a.laeven@uva.nl.

§APG Asset Management, the Netherlands. Email: rob.vanden.goorbergh@apg-am.nl

1 Introduction

The potential benefits of international diversification have been known to equity investors for long (see, for example, Solnik (1974)). Nevertheless, the actual equity portfolios held by investors appear to be far from optimally diversified as measured by classic models. The equity home bias, for instance, is a well-recognized pattern of under-diversification. It refers to the empirical finding that investors overinvest in domestic equities relative to the theoretically optimal investment portfolio.[1] Since the seminal paper by French and Poterba (1991), there has been extensive research on the measurement and explanation of home bias. Information asymmetry and familiarity are commonly offered as potential explanations for the equity home bias.[2]

However, the equity home bias is only part of the under-diversification puzzle. Taking the perspective of a world investor, free from home bias, we find that even the international market portfolio is not optimally diversified according to the classic asset allocation theory. Figure 1 plots the dynamics of the market portfolio on the right and on the left the Merton mean-variance portfolio (see Merton (1969)) of US, Japan and European equities from 2007 to the end of 2012 with expected returns estimated over the full sample and covariances estimated from an expanding window. We see that the market portfolio is consistently over-weighting the US equity and under-weighting the other two.

[1] Home bias is also found at the local level. For example, Coval and Moskowitz (1999) and Sialm et al. (2013) find that US investors over-weight local assets that are geographically close.

[2] See, e.g., Epstein and Miao (2003), Uppal and Wang (2003), Bekaert and Wang (2009), Boyle et al. (2012).

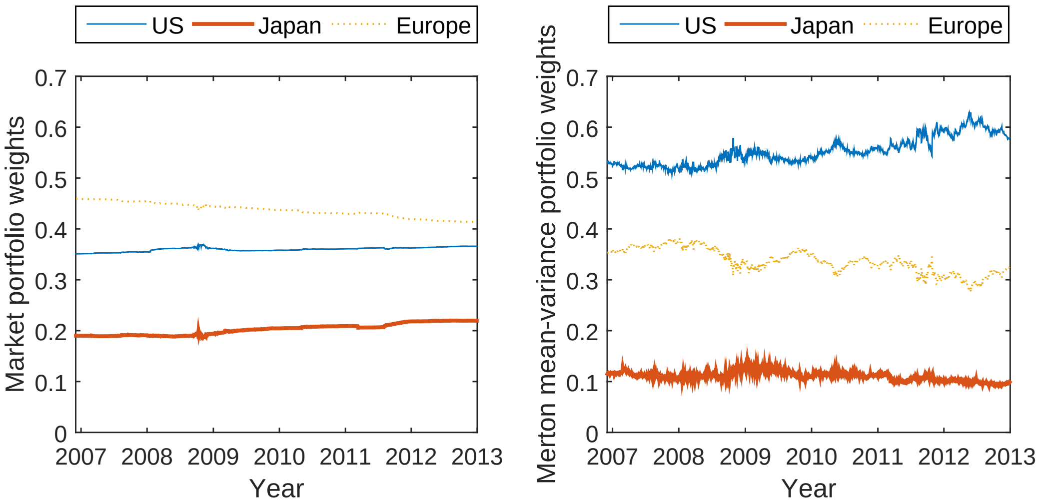

Figure 1: Market equity portfolio weights (rights panel) and the Merton mean-variance (left panel) equity portfolio weights on US, Japanese and European equity markets from the beginning of 2007 to the end of 2012. The market portfolio weights are calculated by dividing the market values (US dollar denominated) of MSCI US, Japan and Europe by their sum at each time point. The Merton mean-variance portfolio is computed using excess log returns of MSCI indices over local risk free rates. US 3 month TBill rates, Japan base discount rates, UK 3 month Libor rates are used as proxies for the local risk free rates. Expected excess returns are estimated using the *total returns* data from January 1970 to December 2012. The variance is estimated using daily *price index* from January 1972 to December 2012 with an expanding window. Weights are normalized to add up to 1.

Acknowledging the fact that expected returns cannot be estimated consistently in samples of finite length (Merton 1980), we can equivalently ask the question of what expected returns would explain the market weights, under the assumption that investors are mean-variance optimizers and put their wealth in the equity markets of US, Japan and Europe. Table 1 compares the empirical expected excess returns with the implied expected excess returns for different sample periods. Following French and Poterba (1991), the implied expected excess returns are calculated such that mean-variance investors choose to hold the market equity portfolio. Denote the portfolio weights on the (currency-hedged) risky assets by $\boldsymbol{w}$ and the weights within the equity portfolio by $\bar{\boldsymbol{w}}$. For a Merton mean variance investor, the equity portfolio composition is irrespective of the risk aversion coefficient and is given by

$$\bar{\boldsymbol{w}} = \frac{\boldsymbol{w}}{\boldsymbol{w}'\boldsymbol{\iota}} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\boldsymbol{\iota}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}},$$

where $\boldsymbol{\Sigma}$ is the estimated covariance matrix of excess log returns of MSCI indices of the corresponding sample, $\boldsymbol{\iota}$ a vector of ones. The implied expected excess log return is the $\boldsymbol{\mu}$ that delivers the observed market equity portfolio weights adjusted for the half squared volatility. We normalize the implied expected excess log returns such that either the implied Japanese return (JA as reference) or the implied European return (EU as reference) is the same as its historical estimate. In the full sample estimate, the implied expected excess log return for US is over 500 basis points higher than its empirical value when the Japanese equity is used as reference, and is 200 basis points higher when the European equity is used as reference. If we exclude the turbulent period of the global financial crisis and terminate the sample as of the end of 2006, we find the US implied return to be 2000, 200 basis points higher using the Japanese and the European equity as reference, respectively. Although the expected excess returns measured as the sample mean are not precise, it is unlikely that the US equity can deliver such a high expected excess return consistently based on the historical data. In other words, the risk return trade off of the US equity is not good enough to attract so much investment as it does.

Empirical vs implied expected excess log returns (% per annum)

	Full sample			As of the end of 2006		
	Empirical	Implied (JA as reference)	Implied (EU as reference)	Empirical	Implied (JA as reference)	Implied (EU as reference)
US	4.12	9.53	6.05	4.55	24.65	6.03
JA	2.45	2.45	1.12	4.46	4.46	0.34
EU	4.16	6.61	4.16	3.67	16.26	3.67

Table 1: Empirical and implied expected excess returns denoted in % per annum. Empirical expected excess returns are estimated using MSCI index returns over local risk free rates. Implied expected excess returns are computed based on the market values of MSCI US, Japan and Europe, and are solutions of $\bar{\boldsymbol{w}} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\boldsymbol{\iota}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}$, with $\bar{\boldsymbol{w}}$ the observed market equity portfolio weights, $\boldsymbol{\iota}$ a vector of ones, $\boldsymbol{\Sigma}$ the estimated covariance matrix of excess log returns of MSCI indices using daily returns. Empirical expected excess log returns are estimated using the MSCI *total returns* index over local risk free rates, for which US 3 month TBill rates, Japan base discount rates, UK 3 month Libor rates are used as proxies. Both samples start from January 1972. The full sample ends December 2012, and the sub sample terminates as of the end of 2006.

To be consistent with the composition of the market portfolio, investors should demand a higher expected return in the US market. However, Figure 1 and Table 1 show that the question that why an equity market remains larger (smaller) than others cannot be answered by differences in Sharpe ratios alone. Forbes (2010) finds that the foreign investment in the US cannot be explained by standard portfolio allocation models and diversification motives and thus puts forward the question:

> "Why are foreigners willing to invest an average of well over 5 billion every day in the United States – especially given low returns relative to comparable investments in other countries...?"

It brings us to an equally interesting phenomenon, the "US bias", i.e., the extent to which the market portfolio over-weights the US equity compared to classic asset allocation models. It is a well-documented fact that investors hold biased equity portfolios not only towards the home equities but also towards some other equities. Kang et al. (1997) study the foreign ownership in Japanese firms and show that investors hold foreign portfolios tilted towards large firms with good accounting performance rather than those with better Sharpe ratios. Chan et al. (2005) find that markets that are more developed and larger in market capitalization attract foreign investors. Ferreira and Matos (2008) study the preference of institutional investors worldwide and conclude that institutional investors prefer firms that are cross-listed in the US and constituents of the Morgan Stanley Capital International World Index. Bekaert and Wang (2009) compare country actual equity holdings to a theoretical optimal allocation given by the CAPM framework and find that investors significantly over-invest in the US and under-invest in Japan. Forbes (2010) states that both size and liquidity contribute to the attractiveness of US financial markets on top of the risk-return tradeoff. Diyarbakirlioglu (2011) studies the mutual fund holdings and finds that investors' foreign portfolio tend to be concentrated in large stock markets and well-developed economies.

While some controversy remains whether investors universally bias their equity portfolio towards the US equity, another strand of the literature suggests that the US plays a special role in the international financial market. For instance, King and Wadhwani (1990) investigate high-frequency returns for US, Japan and UK and find that when New York opens, there is a jump in the London price reflecting the information contained in the New York opening price. Eun and Shim (1989) employ a vector autoregression system and find that innovations in the US are rapidly transmitted to other markets whereas no single foreign market can significantly explain US market movements. Similarly, Hamao, Masulis, and Ng (1990) find significant volatility spillover effects from New York to London and Tokyo but no price volatility spillover effects to New York are observed. Bollerslev, Tauchen, and Zhou (2009) find that the US plays a leading role in terms of variance premia. In an interesting paper, Rapach et al. (2013) show that lagged US equity returns significantly predict returns in numerous non-US countries, while lagged non-US returns display limited predictive ability with respect to US returns. They state that

> "...the lead-lag relationships are an important feature of international stock return predictability, with the United States generally playing a leading role...our results call for an international asset pricing model that explicitly incorporates the leading role of the United States."

Inspired by the empirical results from both strands of the literature, we propose an asset pricing model that explicitly takes into account the leading feature of the US equity. We model the lead-lag relation using asymmetric jump excitation, which enables a price plunge in the US equity to get reflected in future prices of foreign equities but not the other way around. As a result, our model suggests that investors should take a larger stake in the US market than classic model predictions, since the US equity has a larger hedging potential against economic state variables than equities in the other markets, thereby generating a US bias as has been documented in the literature.

Specifically, we model a contagious financial market with mutually exciting jumps to account for excess comovement during economic downturns led by the US market.[3] Different from Lévy type models that are widely applied in the literature, mutually exciting jumps are both cross sectionally and serially dependent, meaning that a large price movement that happens in the domestic market today increases the probability of experiencing further price jumps in the same market in the future as well as the probability of experiencing price jumps in other regional markets. There are two important indicators that measure the cross section excitation capability of an equity market – how much a domestic price crash can get reflected in future foreign equity prices, which measures the capability of exciting other markets; and how a foreign price crash can affect future domestic equity prices, which measures the inclination of getting excited by other markets. Empirical evidence mentioned earlier suggests that these two measures are typically not equal. Consistent with the empirical findings, we allow for asymmetric jump excitation structure. The leading role of the US equity market is characterized by having a large cross section excitor as the source jump component and a small cross section excitor as the target jump component, indicating on one hand its capability of spreading domestic jump risks worldwide and on the other its resistance to foreign equity risk spillover.

Apart from allowing for jump propagation, we deviate from the standard asset allocation literature in two aspects. (1) Instead of using representative assets of every regional market, we assume that each local market is made up of a great many individual assets which are exposed to regional risk factors as well as idiosyncratic risks. While the regional risk factors are systematic and cannot be diversified away, idiosyncratic risks can be eliminated by holding a well-diversified portfolio. (2) We adopt the factor investing perspective and focus on allocation to risk factors rather than assets. Inspired by Ang (2014), who remarks that "factors are to assets what nutrients are to food; factor risks are the driving force behind risk premiums", we derive optimal portfolio exposure to risk factors instead of optimal portfolio weights on each individual asset. In this way, thousands of assets reduce to only a few manageable risk factors.

These specifications allow us to solve in closed-form the portfolio optimization problem with multiple regions and a large number of assets that are exposed to mutually exciting jump risks, systematic Brownian risks, and idiosyncratic Brownian risks. The optimal portfolio exploits the diversification ben-

[3] Mutually exciting jump process is a multivariate version of the Hawkes process, which was originally developed by Hawkes (1971a), Hawkes (1971b), and are introduced to modeling the dynamics of asset returns by Aït-Sahalia, Cacho-Diaz, and Laeven (2014), to portfolio optimization problems by Aït-Sahalia and Hurd (2012), and to modeling credit default by Aït-Sahalia, Laeven, and Pelizzon (2014).

efits among independent risk factors, and at the same time exploits the hedging potential within the dependence structure among risk factors and state variables. As a result, the optimal portfolio in this high-dimensional contagious market: (1) is sufficiently diversified, in the sense that it consists of a large number of individual assets to diversify away idiosyncratic risks; (2) is biased towards the US equity as compared to classic portfolio predictions. Intuitively, the reason of the bias is that the US leads international equity returns by being the major driving forth of the economic states and therefore the assets there are better hedges against the state variables.

Generally speaking, incorporating jump risks in our model brings three effects to the traditional asset allocation where equities are assumed to be driven by Brownian motions alone. First, as discovered by Das and Uppal (2004), Aït-Sahalia et al. (2009), given the same expected return, the investment in risky assets is smaller for an investor who accounts for jumps. This is due to the fact that, when a jump occurs, the wealth can drop significantly before the investor has a chance to adjust the portfolio as he/she would when faced with Brownian risks. As a result, the investor prefers a smaller leverage to stay on the safe side. Second, compared with the constant jump intensity case, jump excitation increases the demand of risky assets. When a jump occurs, the state variables and the equity prices move in opposite directions. To reduce the uncertainty in state variables, the investor should increase the exposure to risky assets in order to exploit the hedging potential in the jump components. This effect is first seen in Liu, Longstaff, and Pan (2003), who show in a univariate model, that jumps in volatility increase the optimal portfolio weight on the risky asset. Hoever, the implications of stochastic jump intensities in a multivariate setting are still not well explored, possibly due to the difficulty in formulating a flexible yet tractable model which yields analytical solution for the optimal asset allocation. We extend the existing non-Poissonian jump diffusion asset allocation literature to a multivariate setting using mutually exciting jumps, which gives rise to the third effect, namely, the US bias, meaning that, compared to the Merton mean variance portfolio, an investor over-invests in the US market whose jump component is more capable of exciting the jump components in other markets but is less prone to be excited by the other jump components. We find that the US bias arises when jump excitation is asymmetric, in which case jump components have heterogeneous hedging potential against state variables. The investor thus tilts the portfolio towards the US equity for a more effective hedging. Ignoring price discontinuities or the excitation nature of jumps will result in substantial welfare losses. The first two effects are known to the literature while the last effect is a novel finding in this paper.

We apply our model to historical prices of MSCI US, Japan and Europe. We estimate the parameters of our model using daily return data. We show that neither Poisson jumps nor self exciting jumps are able to reproduce the pattern of US bias in the market portfolio. Only when jumps are mutually exciting with an asymmetric excitation matrix does the optimal portfolio exhibit the US bias. The portfolio prediction generated by the mutually exciting jump diffusion model closely resembles the risk profile of the market portfolio.

Our paper belongs to the asset allocation literature pioneered by Merton (1969). The past two decades have seen lots of efforts to address the portfolio optimization problem in richer stochastic environments. To list a few, Wachter (2002) and Chacko and Viceira (2005) solve in closed form the consumption/portfolio problem in a diffusion market with mean-reverting state variables; Liu (2007) solves the asset allocation problem for general diffusion return processes; Das and Uppal (2004), Aït-Sahalia et al. (2009) study the portfolio implications of systemic jumps under constant investment opportunities; Liu et al. (2003) look at the portfolio optimization problem when both price and volatility can jump; Jin and Zhang (2012) consider the asset allocation for general Lévy processes.

To our knowledge, an analytical characterization of the US bias using the lead-lag relationships in international returns cannot be easily replicated by other existing portfolio choice models in the literature. Both jump propagation and a multivariate setting are essential ingredients that lead to the US bias. The excitation asymmetry property is well defined only in a multivariate model. Sophisticated portfolio choice models that admit closed form solutions often times focus on univariate settings with a single stock in the market.[4] Moreover, even in a truly multivariate framework, the asymmetric feature cannot be replaced by stochastic volatility (see, e.g., Buraschi et al. (2010)) nor regime-switching models. The linear correlation, as a measure of dependence used in such models, is a symmetric and simultaneous relation, from which there cannot be a lead-lag or asymmetric relation between two equity markets. Ang and Bekaert (2002) propose a regime-switching model to account for the fact that correlations between international equity returns are higher during bear markets than during bull markets. While regime-switching models are able to account for excess linear dependence during economic downturns, the dependence structure of international equities can be nonlinear and asymmetric.[5] More importantly, Rapach et al. (2013) show that US shocks are only fully reflected in non-US equity prices with a lag. Therefore the dependence structure of international equities is beyond conditional linear correlations. The nonlinearity, asymmetry and lead-lag properties distinguish our asymmetric excitation model from stochastic volatility and regime-switching models. Of course, one could expect to generate a bias towards a certain equity market in a stochastic environment that is sophisticated enough. We believe that the mutually exciting jump diffusion model is a natural, realistic and parsimonious way that gives rise to this interesting effect.

The remainder of this paper is organized as follows. Section 2 postulates a model of asset prices that generates lead-lag relations in international equity returns featuring mutually exciting jumps. We solve

[4]For instance, see the stochastic volatility model solved by Liu (2007), stochastic volatility with jumps model proposed in Liu and Pan (2003), the double jump model of Liu et al. (2003), and Branger et al. (2008) who extend Liu and Pan (2003) to allow for multiple jumps in volatility but stay within the single stock framework.

[5]See Ang and Chen (2002), Christoffersen, Jacobs, and Ornthanalai (2012), for empirical evidence.

for the optimal portfolio using the martingale method in a market with a large number of individual assets. We show that the market is asymptotically complete, in the sense that the optimal portfolio path can be closely tracked by investing in a large basket of individual stocks to diversify away idiosyncratic risks. In Section 3, we study the property of the optimal exposure to jump risks using comparative statics analysis, and show that the optimal portfolio exhibits the US bias. In Section 4, we quantify the certainty equivalent loss in terms of annualized returns if the investor were to ignore jump excitation. Section 5 reports the calibration results and numerical findings. Section 6 concludes.

2 Optimal asset allocation in a contagious financial market

In this section, we propose a model of asset prices that generates lead-lag relations in international returns. This is achieved by extending the pure diffusion processes of asset returns to include both cross sectionally and serially dependent jump components, namely, mutually exciting jumps. We specify a general contagious financial market with mutually exciting jumps in Section 2.1. In Section 2.2, we discuss the general features of the optimal portfolio weights in this market without really solving for the optimal portfolio weights. We impose additional structure on equity risk premiums in Section 2.3 which enables us to solve the portfolio optimization problem in closed form using the martingale approach in Section 2.4. In Section 2.5, we discuss the market completeness of the financial market.

2.1 A general model

We work in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ that satisfies the usual conditions. Let $\boldsymbol{N}_t = (N_{1,t}, \ldots, N_{n,t})'$ be mutually exciting jumps with intensities $\lambda_{i,t}, i = 1, \ldots, n$, following the dynamics

$$d\lambda_{i,t} = \alpha_i(\lambda_{i,\infty} - \lambda_{i,t})dt + \sum_{j=1}^{n} \beta_{ij} dN_{j,t}, \quad \alpha_i, \beta_{ij}, \lambda_{i,\infty} \geq 0, \quad i, j = 1, \ldots, n. \tag{1}$$

The occurrence of a jump in component j at time t, i.e., $dN_{j,t} = 1$, not only raises the intensity of jump component j, $\lambda_{j,t}$, by a non-negative amount β_{jj}, but also increases the intensities of other jump components, $\lambda_{i,t}, i \neq j$, by a non-negative amount β_{ij}. After being excited, the intensity of each jump component $\lambda_{i,t}$ mean reverts to the steady state, $\lambda_{i,\infty}$, at an exponential decaying rate α_i, until it gets excited by a next jump occurrence.

In the remainder, we call $\boldsymbol{\beta}$, defined as

$$\boldsymbol{\beta} := (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n) = \begin{pmatrix} \beta_{11} & \ldots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \ldots & \beta_{nn} \end{pmatrix},$$

the *excitation matrix*; β_{ii} is called the *self-excitor* of jump component i; β_{ij} is called the *cross section excitor* of jump component j, in which jump component j is called the *source jump component*, and the jump component i is called the *target jump component*.

The unconditional expectation of the jump intensity is given by $E[\boldsymbol{\lambda}_t] = (\boldsymbol{I} - \boldsymbol{\beta}./(\boldsymbol{\alpha\iota}'))^{-1}\boldsymbol{\lambda}_\infty$. The intensity processes can be made stationary by imposing $(\boldsymbol{I} - \boldsymbol{\beta}./(\boldsymbol{\alpha\iota}'))^{-1} > 0$. Here, $\boldsymbol{I}$ is an n by n identity matrix; $\boldsymbol{\alpha}, \boldsymbol{\lambda}_\infty$ are vectors of $\alpha_i, \lambda_{i,\infty}, i = 1, \ldots, n$, respectively; $\boldsymbol{\iota}$ is a column vector of all ones. We adopt the convention of denoting vectors and matrices using boldface characters to distinguish them from scalars. We use $\circ$ to denote element-wise multiplication of matrices and ./ to denote element-wise division. We use "," for column breaks and ";" for row breaks in a matrix.

Let there be a risk-free asset S_t^0, generating an instantaneous risk free return r_t,

$$S_t^0 = S_0^0 \exp\Big(\int_0^t r_s ds\Big), \quad S_0^0 > 0,\ r_t \geq 0,\ t \in [0, T].$$

We assume that each S_t^0-deflated security price process is in the space $\mathcal{H}^2$ containing any progressively measurable and square integrable semi-martingale process $S_{i,t}, i = 1, \ldots, m$, following

$$\begin{cases} \frac{dS_{i,t}}{S_{i,t^-}} = \mu_{i,t}dt + \sum_{j=1}^m \sigma_{i,j,t} dW_{j,t}^\circ + \sum_{l=1}^n d_{i,l} z_{l,t} dN_{l,t}, \\ d\lambda_{l,t} = \alpha_l(\lambda_{l,\infty} - \lambda_{l,t})dt + \sum_{j=1}^n \beta_{lj} dN_{j,t}. \end{cases} \tag{2}$$

Here, $\mu_{i,t} > 0$ is the (state-dependent) excess returns of asset i from the diffusion component; $\sigma_{i,j,t} > 0$ is the (state-dependent) exposure of asset i to the Brownian risk W_j°. $\boldsymbol{W}_t^\circ = (W_{i,t}^\circ, \ldots, W_{m,t}^\circ)'$ is a vector of standard and independent Brownian motions. We use S_{i,t^-} to denote the left-limit of $S_{i,t}$. The $(i,j)^{\text{th}}$ entry of the instantaneous covariance matrix $\boldsymbol{\Sigma}_t$ given by

$$\boldsymbol{\Sigma}_t[i,j] = \sum_{k=1}^m \sigma_{i,k,t}\sigma_{j,k,t}. \tag{3}$$

The exposure of asset i to jump component l is denoted by a constant $d_{i,l}$. The amplitude of jump component l is denoted by i.i.d. random variables $z_{l,t}, t \in [0, T]$, which determine the percentage change in the asset price caused by an occurrence in jump component N_l at time t. The jump amplitudes are assumed to be independent of all risk factors.

The mutually exciting jump diffusion model postulated in (2) is able to produce important stylized facts of asset returns. For example, the asset returns exhibit jump clustering as a result of the time series excitation, and systemic jumps as a result of the cross section excitation. The model generates lead-lag and asymmetric relations in international equity returns. Unlike dependence generated by (stochastic) covariance, which is simultaneous and symmetric in the sense that $\mathrm{Cov}_t(X_1, X_2) = \mathrm{Cov}_t(X_2, X_1)$, contagion allows for lagged dependence. The dependence structure can be further made asymmetric by setting $\beta_{ij} \neq \beta_{ji}$ to indicate that some jump components have a larger potential to excite other jump components. Equities with these jump components tend to lead international equity returns, since a price plunge there can get reflected in future prices of other equities.

The model also generates excess comovement during market turmoil. During tranquil periods, international asset returns are correlated through the instantaneous covariance $\mathbf{\Sigma}_t$, with uncommon large price movements taking place occasionally. In periods of financial crises, initiated by the first few downside jumps, jump intensities build up and give rise to clustered subsequent jumps in the initial market as well as the other markets across the world, creating nonlinear excess tail dependence in economic downturns.

2.2 Optimal asset allocation

We consider an expected utility investor with power utility $u(x) = \frac{1}{1-\gamma}x^{1-\gamma}, \gamma > 0$. The investor is given a non-stochastic initial endowment $x_0 > 0$ to invest in the risk-free and risky assets. The investor neither consumes nor receives any intermediate income. Assume that the investor can rebalance the portfolio in continuous-time without incurring any transaction costs. The objective is to maximize the expected utility over terminal wealth X_T through optimal continuous time trading. Denote the portfolio weights (percentage of wealth) on the risky assets at time t by $\boldsymbol{w}_t = (w_{1,t}; \ldots; w_{m,t}), 0 \leq t \leq T$, assumed to be adapted cáglád processes, bounded in $\mathcal{L}^2$.[6] We do not impose leverage restrictions, so the position on the risk-free asset at time t is given by $w_{0,t} = 1 - \sum_{i=1}^{m} w_{i,t}$, which can be a negative amount. The S_t^0-deflated wealth process X_t is self-financing:

$$\begin{aligned} \frac{dX_t}{X_{t^-}} &= \sum_{i=1}^{m} w_{i,t} \frac{dS_{i,t}}{S_{i,t^-}} \\ &= \boldsymbol{w}_t' \boldsymbol{\mu}_t dt + \boldsymbol{w}_t' \mathbf{\Sigma}_t^{1/2} d\boldsymbol{W}_t^{\circ} + \sum_{l=1}^{n} \boldsymbol{w}_t' \boldsymbol{d}_l z_{l,t} dN_{l,t}. \end{aligned} \tag{4}$$

Here, $\boldsymbol{\mu}_t, \boldsymbol{d}_l$ are vectors containing $\mu_{i,t}$ and $d_{i,l}$ introduced in Equation (2).

[6] Since portfolio weights cannot anticipate jumps, they are $\mathcal{F}_{t-}$ measurable and left continuous (cf. Aït-Sahalia and Hurd (2012)).

The asset allocation problem is formulated as

$$\sup_{\{\boldsymbol{w}_t, 0\leq t\leq T\}} E[u(X_T)|\mathcal{F}_0]. \tag{5}$$

To solve the asset allocation problem, we employ initially the stochastic control theory. Define the indirect utility function J at time t as

$$J(t,x,\boldsymbol{\lambda}) = \sup_{\{\boldsymbol{w}_t\}} E_t\Big[\frac{X_T^{1-\gamma}}{1-\gamma}\Big],$$

where the expectation is conditional on the information available at time t.[7] Bellman's optimality principle requires that

$$0 = \sup_{\{\boldsymbol{w}_t\}} \mathcal{A}J(t,x,\boldsymbol{\lambda}),$$

where $\mathcal{A}$ denotes the infinitesimal generator operator. The Hamilton-Jacobi-Bellman (HJB) equation reads (we omit the arguments $t, x, \boldsymbol{\lambda}$ of function J when no confusion is caused)

$$\begin{aligned} 0 = \sup_{\boldsymbol{w}_t}\Big\{ & J_t + \boldsymbol{w}_t'\boldsymbol{\mu}_t J_x x + \frac{1}{2}\boldsymbol{w}_t'\boldsymbol{\Sigma}_t\boldsymbol{w}_t J_{xx}x^2 + \sum_{l=1}^n \alpha_l(\lambda_{l,\infty}-\lambda_l)J_{\lambda_l} \\ & + \sum_{l=1}^n \lambda_l E[J(t, x(1+\boldsymbol{w}_t'\boldsymbol{d}_l z_{l,t}), \boldsymbol{\lambda}+\boldsymbol{\beta}_l) - J]\Big\}. \end{aligned} \tag{6}$$

We use J_t, J_x, J_{λ_l} to denote the partial derivatives of $J(t,x,\boldsymbol{\lambda})$ with respect to t, x, λ_l and similarly for the higher order derivatives. The expectation is taken over the jump amplitude distribution of $z_{l,t}$. $\boldsymbol{\beta}_l$ denotes the excitor vector when N_l is the source jump component, which is the l^{th} column of the excitation matrix, $\boldsymbol{\beta}_l = (\beta_{1l};\ldots;\beta_{nl})$.

It is known that the indirect utility function for a power utility investor can be written as

$$J(t,x,\boldsymbol{\lambda}) = \frac{x^{1-\gamma}}{1-\gamma}f(t,\boldsymbol{\lambda}),$$

where $f(t,\boldsymbol{\lambda})$ is a deterministic function of time t and the value of state variables $\boldsymbol{\lambda}$.[8] We substitute for this functional form in the HJB Equation (6), and solve the first order condition with respect to $\boldsymbol{w}_t$ to

[7] We sometimes denote the conditional expectation $E[\cdot|\mathcal{F}_t]$ as $E_t[\cdot]$. We use the two notations interchangeably.

[8] Since

$$J(t,cx,\boldsymbol{\lambda}) = \sup E_{t,cx,\boldsymbol{\lambda}}\Big[\frac{(cX_T)^{1-\gamma}}{1-\gamma}\Big] = c^{1-\gamma}\sup E_{t,x,\boldsymbol{\lambda}}\Big[\frac{X_T^{1-\gamma}}{1-\gamma}\Big] = c^{1-\gamma}J(t,x,\boldsymbol{\lambda}),$$

we conclude that the value function is homogeneous of degree $1-\gamma$ in the wealth level. Let $c=\frac{1}{x}$. It holds that

$$J(t,1,\boldsymbol{\lambda}) = x^{-(1-\gamma)}J(t,x,\boldsymbol{\lambda}).$$

Rearrange and get

$$J(t,x,\boldsymbol{\lambda}) = \frac{x^{1-\gamma}}{1-\gamma}f(t,\boldsymbol{\lambda}),$$

where

$$f(t,\boldsymbol{\lambda}) = (1-\gamma)J(t,1,\boldsymbol{\lambda}),$$

with terminal condition $f(T,\boldsymbol{\lambda})=1$.

get the following implicit function that characterizes the optimal portfolio weights $\boldsymbol{w}_t^*$. For $0 \le t \le T$, the optimal portfolio weights $\boldsymbol{w}_t^*$ solve

$$\boldsymbol{\mu}_t - \gamma \boldsymbol{\Sigma}_t \boldsymbol{w}_t^* + \sum_{l=1}^{n} \frac{f(t, \boldsymbol{\lambda} + \boldsymbol{\beta}_l)}{f} \lambda_l E[(1 + \boldsymbol{w}_t^{*\prime} \boldsymbol{d}_l z_{t,l})^{-\gamma} \boldsymbol{d}_l z_{t,l}] = 0, \tag{7}$$

with $f(t, \boldsymbol{\lambda})$ satisfying

$$\begin{aligned} 0 =& f_t + (1-\gamma)\boldsymbol{w}_t^{*\prime}\boldsymbol{\mu}_t f - \frac{1}{2}\gamma(1-\gamma)\boldsymbol{w}_t^{*\prime}\boldsymbol{\Sigma}_t \boldsymbol{w}_t^* f + \sum_{l=1}^{n} \alpha_l(\lambda_{l,\infty} - \lambda_l) f_{\lambda_l} \\ &+ \sum_{l=1}^{n} \lambda_l E[(1 + \boldsymbol{w}_t^{*\prime} \boldsymbol{d}_l z_{l,t})^{1-\gamma} f(t, \boldsymbol{\lambda} + \boldsymbol{\beta}_l) - f]. \end{aligned} \tag{8}$$

One can easily verify that the pair $(\boldsymbol{w}_t^*, f(t, \boldsymbol{\lambda}))$ jointly determined by Equations (7) and (8) satisfies the HJB Equation (6) and therefore $\boldsymbol{w}_t^*$ is optimal.

The optimal portfolio weights given in Equation (7) can be decomposed into familiar components:

$$\boldsymbol{w}_t^* = \underbrace{\frac{1}{\gamma}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\mu}_t}_{\text{(I)}} + \frac{1}{\gamma}\boldsymbol{\Sigma}_t^{-1}\Big(\underbrace{\sum_{l=1}^{n}\lambda_l \boldsymbol{M}_{l,t}}_{\text{(II)}} + \underbrace{\sum_{l=1}^{n}\lambda_l \frac{f(t, \boldsymbol{\lambda} + \boldsymbol{\beta}_l) - f}{f}\boldsymbol{M}_{l,t}}_{\text{(III)}}\Big), \tag{9}$$

where

$$\boldsymbol{M}_{l,t} := E[(1 + \boldsymbol{w}_t^{*\prime} \boldsymbol{d}_l z_{t,l})^{-\gamma} \boldsymbol{d}_l z_{t,l}].$$

The optimal portfolio weights consist of a mean-variance demand (I), a myopic buy-and-hold demand (II), and an intertemporal hedging demand (III). The mean-variance demand (I) is given by the mean-variance weights, exploiting diversification benefits of the instantaneous covariance structure.

The myopic buy-and-hold demand (II) arises because the asset prices have discontinuities. As explained by Liu et al. (2003), unlike continuous fluctuations, jumps may occur before the investor has the opportunity to adjust the portfolio. Jump risks, therefore, are similar to "illiquidity risk": the investor has to hold the asset until the jump has occurred. Observe that

$$\boldsymbol{M}_{l,t} \propto \nabla_{\boldsymbol{w}_t} E[u(X_t) - u(X_{t^-}) | N_{l,t} - N_{l,t^-} = 1].$$

$E[u(X_t) - u(X_{t^-}) | N_{l,t} - N_{l,t^-} = 1]$ is the expected utility gain at time t conditional on an occurrence in jump component l at time t . Therefore (II) is the expected marginal utility increase induced by jump component l from investing in one unit of risky assets at time t. The buy-and-hold demand is "myopic" in the sense that it does not take into account the uncertainties of the future jump intensities.

The last term (III) is tailored to account for the fact that the jumps are mutually exciting. Since the asset prices $\boldsymbol{S}_t$ and the state variables $\boldsymbol{\lambda}_t$ are both driven by jumps $\boldsymbol{N}_t$, the risky assets can be used to hedge future realizations of the state variables. Intuitively, the mean-variance demand and the myopic buy-and-hold demand exploit the risk-return trade-off of the risky assets, whereas the intertemporal hedging demand is only concerned with state variable uncertainties.

All three components of the portfolio weights can be time-varying, but for different reasons. The mean-variance demand (I) and myopic buy-and-hold demand (II) depend on the spot values of the asset return parameters. Hence they change with the spot values instantaneously. The intertemporal hedging demand, on the other hand, depends not only on the spot values of the asset return parameters, but also on how the returns and the state variables evolve within the investment horizon. The information of future outcomes is contained in $f(\cdot)$, which is horizon dependent.

Remark 1. In the most general case, we allow the state variable $\boldsymbol{\lambda}_t$, similar to equities, to follow a jump diffusion process

$$d\boldsymbol{\lambda}_t = \boldsymbol{y}_0(\boldsymbol{\lambda}_t)dt + \boldsymbol{y}_1(\boldsymbol{\lambda}_t)d\boldsymbol{W}_t^\circ + \tilde{\boldsymbol{y}}_1(\boldsymbol{\lambda}_t)d\widetilde{\boldsymbol{W}}_t + \boldsymbol{y}_2 d\boldsymbol{N}_t. \tag{10}$$

Here, $\widetilde{\boldsymbol{W}}_t$ is a $m \times 1$ vector of Brownian motions that are independent of $\boldsymbol{W}_t^\circ$; $\boldsymbol{y}_0(\boldsymbol{\lambda}_t)$ is an $n \times 1$ vector of drift terms; $\boldsymbol{y}_1(\boldsymbol{\lambda}_t)$ and $\tilde{\boldsymbol{y}}_1(\boldsymbol{\lambda}_t)$ are both $n \times m$ matrices; $\boldsymbol{y}_2$ is an $n \times n$ matrix of constants. The optimal portfolio weights in (9) would thus include a fourth volatility hedging component (IV), $\frac{1}{\gamma}\boldsymbol{y}_1' \frac{\nabla_{\boldsymbol{\lambda}} f}{f}$, in order to use the risky assets to hedge the common Brownian risks $\boldsymbol{W}_t^\circ$ in the state variables. In this case, the model nests the stochastic volatility model of Liu (2007), the Poisson jump diffusion model in Das and Uppal (2004) and Aït-Sahalia et al. (2009), the contagion model in Aït-Sahalia and Hurd (2012), the univariate double jump model in Liu et al. (2003) and Branger et al. (2008). Although jump-diffusion-driven state variables will not create additional difficulties in the analysis, we focus on the more parsimonious mutually exciting jump diffusion model for simplicity. In case of the mutually exciting jumps, it holds that

$$\boldsymbol{y}_0 = \boldsymbol{\alpha} \circ (\boldsymbol{\lambda}_\infty - \boldsymbol{\lambda}_t), \quad \boldsymbol{y}_1 = \tilde{\boldsymbol{y}}_1 = \boldsymbol{0}, \quad \boldsymbol{y}_2 = \boldsymbol{\beta}.$$

The purpose of this paper is to evaluate the impact of excitation asymmetry on the optimal portfolio choice, rather than to develop a general multivariate jump diffusion model to nest existing models in the literature.

2.3 Proportional risk premium

In general, the function $f(\cdot)$ in Equation (7) does not admit an analytical expression. In order to fully solve the asset allocation problem, we impose additional structure on the equity risk premium. Inspired by Aït-Sahalia et al. (2014), we propose a parsimonious model to focus on jump propagation through time and across different geographic markets. Let there be n regions with m_i assets in region i. Let $N_1, \ldots, N_n$ represent regional jump components to capture large price drops in equity indices.

Assume that each asset is only exposed to the jump risk of its own region but not to those of the other regions. Equivalently, the jump exposure $d_{i,l}$ in Equation (2) takes the form

$$d_{i,l} = \begin{cases} 1, & \text{if } i = l, \\ 0, & \text{if } i \neq l. \end{cases}$$

Even though jump components in the peripheral markets do not influence domestic asset prices directly, jump risks are systemic in the sense that they mutually excite. The jump intensities $\boldsymbol{\lambda}_t$ follow

$$d\lambda_{i,t} = \alpha(\lambda_{i,\infty} - \lambda_{i,t})dt + \sum_{j=1}^{n} \beta_{ij} dN_{j,t}, \tag{11}$$

where for simplicity we assume that all jump intensities share the same mean-reversion rate α, as in Aït-Sahalia et al. (2014).

We model the "normal" (day-to-day) covariance among regions by *correlated* Brownian motions $\boldsymbol{W}_t = (W_{1,t}, \ldots, W_{n,t})'$ given by

$$\boldsymbol{W}_t = \boldsymbol{L}\boldsymbol{W}_t^{\circ}.$$

Here, $\boldsymbol{L}\boldsymbol{L}'$ is a correlation matrix with ones on the diagonal and correlation coefficients off-diagonal. Besides systematic Brownian risks, assets are also subject to idiosyncratic fluctuations, captured by standard and independent Brownian motions $\boldsymbol{Z}_t = (Z_{i,t}^k),\ i = 1, \ldots, n,\ k = 1, \ldots, m_i$, which are independent of the regional Brownian risks $W_{i,t}, i = 1, \ldots, n$.

Following, among others, French and Poterba (1991), we further assume that the representative investor hedges 100% of the exchange rate risk using, say, forward exchange rate contracts. The hedged return is given by

$$R_t^{\text{hedged}} = r_t + (R_t^{\text{local}} - r_t^{\text{local}}),$$

where r_t is the risk free rate of a reference country and r_t^{local} is the local risk free rate. In other words, the hedged excess log returns are computed as local returns denominated in local currency over local risk free rates. Taking the risk free bond $S_t^0 = S_0^0 \exp\left(\int_0^t r_s ds\right)$ from a reference currency as numeraire, we normalize all price processes as S_t^0-deflated prices henceforth. According to the Numeraire Invariance Theorem (see, for example, Duffie (2010)), such normalization should place essentially no economic effects.

To focus on the effect of jump propagation on the optimal asset allocation, we assume, for simplicity, that the expected return and volatility of equity prices are state independent. Denote individual asset identities by superscripts and region identities denoted by subscripts. We suppose that the *currency-hedged* deflated price of a risky asset k from region i is in the space $\mathcal{S}^2 \in \mathcal{H}^2$ containing $S^k_{i,t}, i = 1, \ldots, n, k = 1, \ldots, m_i, t \in [0, T]$, which follows

$$\frac{dS^k_{i,t}}{S^k_{i,t^-}} = \nu^k_i dZ^k_{i,t} + \sigma^k_i (dW_{i,t} + \eta_i dt) + z^k_i (dN_{i,t} - (1 + \kappa_i)\lambda_{i,t} dt), \tag{12}$$

with constants $\sigma^k_i, \eta_i, \nu^k_i, \kappa_i \geq 0$ for all i, k. Within a given region, the price of any individual asset k is driven by both region specific systematic risks, $W_{i,t}, N_{i,t}$ as well as idiosyncratic risks, $Z^k_{i,t}$.

Following Cox and Ross (1976) and Liu and Pan (2003), we assume that the jump amplitudes z^k_i, $i = 1, \ldots, n$, $k = 1, \ldots, m_i$, are constant. In addition, we restrict that $-1 < z^k_i \leq 0$ to rule out probability of ruin and to indicate that jumps are unfavorable events. Conditioning on an occurrence in jump component N_i, each asset k in region i drops by a deterministic amount. This assumption simplifies the analysis, allowing us to focus on the impact of adverse rare events and the contagious nature of such events in a straightforward way.

Upon comparing Equation with Equation (2), the drift term of asset k from region i is a linear function of the state variable $\lambda_{i,t}$,

$$\mu^k_{i,t} = \sigma^k_i \eta_i - z^k_i (1 + \kappa_i)\lambda_{i,t}.$$

The covariance matrix $\mathbf{\Sigma}$ is constant over time and has the structure

$$\mathbf{\Sigma}[p, p] = \left(\sigma^p_i\right)^2 + \left(\nu^p_i\right)^2, \quad \mathbf{\Sigma}_t[p, q] = \rho_{i,j} \sigma^p_i \sigma^p_j,$$

where i, j are the regional markets to which asset p and asset q belongs, respectively. $\rho_{i,j}$ is the $[i, j]^{\text{th}}$ entry of the correlation matrix $\boldsymbol{LL}'$.

Equation (2.3) is the dynamics specified under the physical measure P. If the market is free of arbitrage, there exists an equivalent martingale measure Q, under which the expected excess return of any S^0_t-deflated risky asset is zero, i.e., $E^Q\left[\frac{dS^k_{i,t}}{S^k_{i,t^-}}\right] = 0$. We start with specifying a pricing kernel process that uniquely prices the three sources of risks: the idiosyncratic Brownian risks, the systematic Brownian risks, and the jump risks, and then show in Section 3.3 that the market is complete in the sense that any random payoff that is consistent with this pricing kernel can be replicated by investing in the available assets in the market.

Define the systematic Brownian risk premium vector $\boldsymbol{\eta} = (\eta_1; \ldots; \eta_n)$ and the jump risk premium vector $\boldsymbol{\kappa} = (\kappa_1; \ldots; \kappa_n)$. Following Liu and Pan (2003), consider a pricing kernel process π_t given by

$$\frac{d\pi_t}{\pi_{t^-}} = -\boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}d\boldsymbol{W}_t + \sum_{i=1}^{n} \kappa_i(dN_{i,t} - \lambda_{i,t}dt), \quad \pi_0 = 1. \tag{13}$$

It is clear from Equation (13) that π is a local martingale. If π is actually a martingale, one can verify according to the Lenglart-Girsanov Theorem that π is the Radon-Nikodym derivative that changes the measure P to a risk neutral measure Q, under which asset prices evolve according to

$$\frac{dS_{i,t}^k}{S_{i,t^-}^k} = \nu_i^k dZ_{i,t}^{k,Q} + \sigma_i^k dW_{i,t}^Q + z_i^k(dN_{i,t}^Q - (1+\kappa_i)\lambda_{i,t}dt), \tag{14}$$

where $W_{i,t}^Q, Z_{i,t}^{k,Q}$ are standard Brownian motions under Q. The jump process $N_{i,t}^Q$ has intensity $(1+\kappa_i)\lambda_{i,t}$ under Q, while the jump amplitude z_i^k remains unchanged. Consequently, $S_{i,t}^k$ is a local martingale under risk measure Q.

Upon comparing the asset dynamics under P given by Equation (2.3) and those under Q given by Equation (14), we can obtain an intuitive understanding of how the three types of risks are priced. First, similar to Merton (1976), the idiosyncratic risk $\boldsymbol{Z}_t$ is assumed to be perfectly diversifiable. As a result, the market portfolio is free of idiosyncratic Brownian risks and the market price of idiosyncratic Brownian risk $\boldsymbol{Z}_t$ is zero. Only systematic risks are priced.

Second, note that the Brownian risk $\Delta W_{i,t}$ has constant variance Δ, while the jump risk $\Delta N_{i,t}$ has variance $\lambda_{i,t}\Delta$ (approximately). Loosely speaking, we are assuming that the risk premium is proportional to the "risk" of the risk factors – the Brownian risk is compensated with $\eta_i\Delta$ and the jump risk is compensated with $\kappa_i\lambda_{i,t}\Delta$. Similar jump risk premium specification can also be found in Pan (2002), Liu et al. (2003), and Boswijk et al. (2015). It implies that the expected stock returns are increasing in the jump intensities $\boldsymbol{\lambda}_t$. Intuitively, this risk premium is sensible, since during recessions when there is a high probability of experiencing large price drops, the investor is compensated by a better risk-return tradeoff. This jump premia specification is consistent with the empirical estimation results in Bollerslev and Todorov (2011), who show that most peaks in the equity jump risk premia are associated with events that mark the market turmoil, and also with Santa-Clara and Yan (2010), who find that the equilibrium equity risk premium is a function of the jump intensity. In this way, the jump intensities under the risk neutral measure are larger than the physical jump frequencies, an empirically relevant fact that has been confirmed in the non-parametric estimation by Bollerslev and Todorov (2011).

2.4 Optimal portfolio exposure to risk factors

We present the optimal portfolio results in this section from the perspective of a world representative investor. Problem 1 defines the portfolio optimization problem.

Problem 1. *Suppose there are no arbitrage opportunities in the market introduced in Section 2.3 and all assets are priced according to the pricing kernel given by Equation* (13)*. Let* $\boldsymbol{\theta}_t^Z$ *be a* $\left(\sum_i^n m_j\right) \times 1$ *vectored process, and* $\boldsymbol{\theta}_t^W, \boldsymbol{\theta}_t^N = (\theta_t^{N_i}), i = 1, \ldots, n,$ *be* $n \times 1$ *vectored processes, which are adapted, cáglád, and bounded in* $\mathcal{L}^2$*. Define the portfolio optimization problem for an expected utility investor with power utility,* $u(x) = \frac{x^{1-\gamma}}{1-\gamma}, \gamma > 0$ *as*

$$\sup_{\{\boldsymbol{\theta}_t^Z, \boldsymbol{\theta}_t^W, \boldsymbol{\theta}_t^N\}} E_0\Big[\frac{X_T^{1-\gamma}}{1-\gamma}\Big], \tag{15}$$

subject to the budget constraint:

$$\frac{dX_t}{X_{t-}} = \boldsymbol{\theta}_t^{Z\prime} d\boldsymbol{Z}_t + \boldsymbol{\theta}_t^{W\prime}(d\boldsymbol{W}_t + \boldsymbol{\eta} dt) + \sum_{i=1}^n \big(\exp(\theta_t^{N_i}) - 1\big)(dN_{i,t} - (1+\kappa_i)\lambda_{i,t} dt). \tag{16}$$

We invoke the martingale method developed by Cox and Huang (1989) to solve for the optimal portfolio exposure to risk factors. The main results are stated in the following proposition.

Proposition 1 (Optimal portfolio exposure to risk factors)**.** *Consider Problem 1. The optimal portfolio exposure to risk factors is given by*

$$\begin{cases} \boldsymbol{\theta}_t^{Z*} =: \boldsymbol{\theta}^{Z*} = \boldsymbol{0}, \\ \boldsymbol{\theta}_t^{W*} =: \boldsymbol{\theta}^{W*} = \frac{1}{\gamma}(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta}, \\ \theta_t^{N_i*} = -\frac{1}{\gamma}\log(1+\kappa_i) + \boldsymbol{\beta}_i'\boldsymbol{B}(t), \end{cases} \tag{17}$$

and the indirect utility function at $t = 0$ *is given by*

$$J(0, x_0, \boldsymbol{\lambda}_0) = \frac{(e^{rT}x_0)^{1-\gamma}}{1-\gamma} \exp\big(\gamma(A(0) + \boldsymbol{B}(0)'\boldsymbol{\lambda}_0)\big). \tag{18}$$

Here, $A(t), \boldsymbol{B}(t)$ *satisfy*

$$\begin{cases} \dot{\boldsymbol{B}}(t) = \frac{\gamma-1}{\gamma}\boldsymbol{\kappa} + \alpha\boldsymbol{B}(t) - (\boldsymbol{\kappa}+1)^{\frac{\gamma-1}{\gamma}} \circ e^{\boldsymbol{\beta}'\boldsymbol{B}(t)} + 1, \\ \dot{A}(t) = \frac{\gamma-1}{2\gamma}\boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta} - \alpha\boldsymbol{B}'(t)\boldsymbol{\lambda}_\infty, \end{cases} \tag{19}$$

with $A(T) = 0, \boldsymbol{B}(T) = \boldsymbol{0}$.

Notice that the optimal exposure to Brownian risks (both idiosyncratic and systematic) is time-independent and that the optimal exposure to jump risks is a continuous deterministic process. Therefore the optimal portfolio exposure to all risk factors satisfy the cáglád assumption.

Alternatively, we may also use the stochastic control method outlined in Section 2.2 with risk exposure $(\boldsymbol{\theta}_t^Z, \boldsymbol{\theta}_t^W, \boldsymbol{\theta}_t^N)$ as control variables. The HJB equation is given by

$$0 = \sup_{\{\boldsymbol{\theta}_t^Z, \boldsymbol{\theta}_t^W, \boldsymbol{\theta}_t^N\}} \Big\{ J_t + \Big(\boldsymbol{\theta}_t^{W\prime} \boldsymbol{\eta} - \sum_{i=1}^n \big(\exp(\theta_t^{N_i}) - 1 \big)(1+\kappa_i)\lambda_{i,t} \Big) J_x x + \frac{1}{2} \big(\boldsymbol{\theta}_t^{W\prime} \boldsymbol{L}\boldsymbol{L}' \boldsymbol{\theta}_t^W + \boldsymbol{\theta}_t^{Z\prime} \boldsymbol{\theta}_t^Z \big) J_{xx} x^2 + \sum_{l=1}^n \alpha(\lambda_{l,\infty} - \lambda_l) J_{\lambda_l} + \sum_{l=1}^n \lambda_l \Big(J\big(t, x\exp(\theta_t^{N_l}), \boldsymbol{\lambda} + \boldsymbol{\beta}_l\big) - J \Big) \Big\}. \tag{20}$$

From Proposition 1, we already know that the indirect utility is exponentially affine in jump intensities. Let $J(t,x,\lambda) = \frac{x^{1-\gamma}}{1-\gamma} \exp(A(t) + \boldsymbol{B}(t)\boldsymbol{\lambda}_t)$. Plugging it into the HJB equation and taking first order conditions with respect to $\boldsymbol{\theta}_t^Z, \boldsymbol{\theta}_t^W, \theta_t^{N_i}$, respectively. One can easily show that the optimal risk exposure coincides with $\boldsymbol{\theta}^{Z*}, \boldsymbol{\theta}^{W*}, \theta_t^{N_i*}$ given by Equation (17).

We show that the Merton mean-variance portfolio can be recovered as a special case. Define regional representative assets as those free of idiosyncratic risks. Let the assets be representative assets from each region and free of jump risks, i.e., $m_i = 1,\ \nu_i^k = z_i^k = 0, \forall k, i$. The representative assets therefore follow

$$\frac{dS_{i,t}}{S_{i,t}} = \sigma_i (dW_{i,t} + \eta_i dt), \quad i = 1, \dots, n.$$

The portfolio choice problem thus becomes the following.

Problem 2. *Consider the setting in Problem 1 with the following budget constraint:*

$$\frac{dX_t}{X_{t-}} = \boldsymbol{\theta}_t^{W\prime} (d\boldsymbol{W}_t + \boldsymbol{\eta} dt). \tag{21}$$

The next lemma states that when we consider representative assets only driven by Brownian motions, the optimal portfolio exposure to risk factors coincide with that of the Merton mean-variance portfolio.

Lemma 1 (The Merton mean-variance portfolio)**.** *Consider Problem 2. The solution is the Merton mean-variance portfolio and the optimal wealth follows*

$$\frac{dX^*_{Merton}}{X^*_{Merton}} = \boldsymbol{\theta}_{Merton}^{W\prime} (d\boldsymbol{W}_t + \boldsymbol{\eta} dt),$$

with the optimal exposure to Brownian risks given by

$$\boldsymbol{\theta}_{Merton}^{W*} = \boldsymbol{\theta}^{W*} = \frac{1}{\gamma} (\boldsymbol{L}\boldsymbol{L}')^{-1} \boldsymbol{\eta}.$$

There are three features of the optimal portfolio exposure we would like to point out. First, the exposure to risk factors at time t, $(\boldsymbol{\theta}^{Z*}, \boldsymbol{\theta}^{W*}, \boldsymbol{\theta}_t^{N*})$, is independent of the wealth level x_t and the realization of the state variables $\boldsymbol{\lambda}_t$. Even if mutually exciting jumps give rise to stochastic investment opportunities, under the assumption of proportional risk premia, the optimal portfolio composition does not vary with the investor's wealth or realizations of the state variables that indicate economic cycles. In other words, there is no market timing of the portfolio strategy. The independence of the wealth level is a result of the wealth homogeneity property of the power utility. The reason that the optimal risk exposure is independent of the realization of the state variables $\boldsymbol{\lambda}_t$ stems from the assumption that the jump risk premium $\kappa_i \lambda_{i,t}$ is a multiple of $\lambda_{i,t}$. In general, one may expect that when the current jump intensities $\boldsymbol{\lambda}_t$ are high, the optimal portfolio exposure to jump risks should be low to stay away from the high probability of a price plunge. In our model, the investor is rewarded proportionally to jump intensities. When the probability of jump occurrences is high, the risk premium is also high to the extent that the demand for the jump risk is independent of the jump intensity.

The second property of the optimal risk exposure is that, although state independent, the optimal jump risk exposure is horizon dependent. In the special case where $\boldsymbol{\beta} = 0$, the jumps are not mutually exciting and the investment opportunities are constant. In this case, there are no hedging incentives, hence no horizon dependence in the jump risk exposure. When $\boldsymbol{\beta} \neq 0$, the investment opportunities are stochastic, giving rise to incentives to hedge against changes in the investment opportunities. Observe that $\boldsymbol{B}(t)$ in Equation (18) measures the sensitivity of the log indirect utility function to the values of the state variables, i.e.,

$$\boldsymbol{B}(t) = \frac{1}{\gamma} \nabla_{\boldsymbol{\lambda}_t} \log J(t, x_t, \boldsymbol{\lambda}_t).$$

The longer the investment horizon, the further away $\boldsymbol{B}(t)$ is from zero, implying a larger impact of state variables on indirect utility, which in turn leads to a stronger motivation for the investor to hedge against the changes in the state variables.

The third property of the optimal portfolio is that the optimal portfolio has no exposure to idiosyncratic risks, i.e., $\boldsymbol{\theta}^{Z*} = \mathbf{0}$. Naturally, since the exposure to idiosyncratic risks are not compensated by any risk premium, the investor stays away from these risk factors. In practice, it means that the investor should invest in a large basket of assets in every region to diversify away the idiosyncratic risks as much as possible.

2.5 Market completeness

Suppose for now that the idiosyncratic risks $\boldsymbol{Z}_t$ are absent. Each region introduces two risk factors – a systematic Brownian motion $W_{i,t}$ and a jump component $N_{i,t}$. As long as there are two investable assets (which are not linearly dependent) from each region, we have a complete market in the sense that any no arbitrage payoff path in the space $\mathcal{S}^2$ can be replicated. Denote the S_t^0-deflated value of the replicating portfolio at time t by P_t with weight $w_{i,t}^k$ on asset k of region i. It holds that

$$P_t = \sum_{i=1}^{n} \sum_{k=1}^{m_i} w_{i,t}^k S_{i,t}^k, \quad t \in [0.T].$$

The following lemma states the complete market result.

Lemma 2 (Market completeness). *Let there be two non-redundant assets (not linearly dependent) from each region following*

$$\frac{dS_{i,t}^k}{S_{i,t^-}^k} = \nu_i^k dZ_{i,t}^k + \sigma_i^k (dW_{i,t} + \eta_i dt) + z_i^k (dN_{i,t} - (1+\kappa_i)\lambda_{i,t} dt), \quad k = 1, 2, \quad i = 1, \ldots, n,$$

with $\nu_i^k \equiv 0, \forall k, i$. *For any payoff* $\{F_t\} \in \mathcal{S}^2$ *which follows*

$$\frac{dF_t}{F_{t^-}} = \sum_{i=1}^{n} g_i (dW_{i,t} + \eta_i dt) + \sum_{i=1}^{n} h_i \Big(dN_{i,t} - (1+\kappa_i)\lambda_{i,t} dt \Big), \tag{22}$$

there exists a $2n \times 1$ *vector* $\boldsymbol{w}_t$ *containing portfolio weights, with* $w_{i,t}^k$ *being the weight on asset* k *of region* i*, such that the resulting portfolio value* P_t *is equal to* F_t *almost surely, i.e.,*

$$P_t = F_t, \quad a.s. \ \forall t.$$

When assets are exposed to idiosyncratic risks, however, we need more than two assets each region so as to diversify away idiosyncratic risks. Let $\boldsymbol{m} = (m_1, \ldots, m_n)$ be a vector containing m_i as the number of available assets in region i. In fact, the result in Lemma 2 holds when the number of assets in each region goes to infinity. The next proposition formalizes this result.

Proposition 2 (Market completeness in the asymptotic sense). *Let there be* m_i *non-redundant assets (i.e., not linearly dependent) in region* i *following:*

$$\frac{dS_{i,t}^k}{S_{i,t^-}^k} = \nu_i^k dZ_{i,t}^k + \sigma_i^k (dW_{i,t} + \eta_i dt) + z_i^k (dN_{i,t} - (1+\kappa_i)\lambda_{i,t} dt),$$

with $\nu_i^k \geq 0, \forall k, i$. *For any portfolio* $\{F_t\} \in \mathcal{S}^2$*, there exist portfolio weights* $w_{i,t}^k, k = 1, \ldots, m_i, i = 1, \ldots, n$*, such that for any* $0 \leq t \leq T$*, the replicating portfolio* $P_t(\boldsymbol{m})$ *satisfies*

$$P_t(\boldsymbol{m}) \xrightarrow{P} F_t,$$

as the number of assets $m_i, i = 1, \ldots, n$, *all go to infinity. As a result, there exist portfolio weights* $w_{i,t}^k, k = 1, \ldots, m_i, i = 1, \ldots, n$, *such that for any* $0 \le t \le T$,

$$P_t(\boldsymbol{m}) \xrightarrow{P} X_t^*,$$

as the number of assets $m_i, i = 1, \ldots, n$, *all go to infinity.*

Proposition 2 shows that the investor is indeed able to construct the optimal portfolio by investing in a large number of assets in each region. Appendix B gives one explicit example of how this can be done.

3 Analysis of the optimal portfolio exposure to jump risks and the effect of excitation asymmetry

In Section 2.2 we show that the optimal portfolio weights consist of a mean-variance demand, a myopic buy-and-hold demand and an intertemporal hedging demand. In Lemma 1, we have shown that the optimal Brownian risk exposure $\boldsymbol{\theta}^{W*}$ corresponds to the Merton mean-variance demand. In this section, we will be analyzing the properties of the optimal portfolio exposure to jump risks. In Section 3.1, we decompose the jump risk exposure $\boldsymbol{\theta}_t^{N*}$ into a Poisson jump risk exposure and a contagion risk exposure. In Section 3.2, we conduct comparative statics analysis of the contagion risk exposure with respect to jump risk parameters. In Section 3.3, we study the effect of excitation asymmetry on portfolio exposure to jump risks. We show that the optimal portfolio can be biased towards an equity market when the excitation structure is asymmetric.

3.1 Decomposition of exposure to jump risks

In this section we are going to show that the jump risk exposure $\boldsymbol{\theta}_t^{N*}$ can be decomposed into a Poisson jump risk exposure $\boldsymbol{\theta}^J$ which corresponds to the myopic buy-and-hold demand, and a contagion risk exposure $\boldsymbol{\theta}_t^C$ which corresponds to the intertemporal hedging demand.

Note that the exposure to a jump component $\theta_t^{N_i}$ is equal to $\log(1 + w_{i,t} z_i)$, where $w_{i,t}$ is the portfolio weight, $z_i < 0$ is the jump amplitude. If the investor longs the asset, i.e., $w_{i,t} > 0$, then it holds that $\theta_t^{N_i} = \log(1 + w_{i,t} z_i) < 0$. The more negative the exposure is, the more appealing the jump factor is to the investor. The portfolio exposure to the jump factor of region i can be written as

$$\begin{aligned}
\theta_t^{N_i*} &= -\frac{1}{\gamma}\log(1+\kappa_i) + \sum_{j=1}^{n} \beta_{ji} B_j(t) \\
&= \underbrace{-\frac{1}{\gamma}\log(1+\kappa_i)}_{\theta^{J_i}} + \underbrace{\overbrace{\beta_{ii} B_i(t)}^{\theta_t^{\mathrm{ts}_i}} + \sum_{j\neq i}^{n} \overbrace{\beta_{ji} B_j(t)}^{\theta_t^{\mathrm{cs}_{i,j}}}}_{\theta_t^{C_i}} \\
&=: \theta^{J_i} + \theta_t^{C_i}.
\end{aligned} \tag{23}$$

The static component θ^{J_i} is the portfolio exposure to Poisson jump risk. When $\boldsymbol{\beta} = 0$, the jumps are Poissonian with constant intensities, in which case the investor's optimal portfolio exposure to jump risks reduces to θ^{J_i}. The exposure to Poisson jump risks does not take into account the stochastic nature of the jump intensities and therefore plays the role of a myopic buy-and-hold demand.

An interesting comparison is to see what happens when the uncertainties brought by the jump risk are generated by Brownian motions that generate the same mean and variance. The jump factor $(dN_{i,t} - (1+\kappa_i)\lambda_{i,t}dt)$ has mean $-\kappa_i\lambda_{i,t}dt$ and variance $\lambda_{i,t}dt$. The instantaneous correlation between the jump components is 0. Consider instead a Brownian motion with the same mean and variance. Then, the investor will have an exposure of

$$\hat{\theta}^{J_i} = -\frac{1}{\gamma}\kappa_i. \tag{24}$$

One can show that

$$|\theta^{J_i}| < |\hat{\theta}^{J_i}|.$$

It implies that the exposure to a risk factor is smaller when it is recognized as a Poisson jump than a Brownian motion, given its mean (risk premium) and volatility (risk). In a situation where asset prices move continuously, the investor can rebalance the portfolio after any infinitesimal changes in value to avoid large losses. However, since the investor cannot anticipate jumps, his/her wealth can change substantially before the investor has an opportunity to perform any adjustment. For this reason, the investor is reluctant to take too much risk exposure for fear of disastrous events.

The horizon-dependent component $\theta_t^{C_i}$ is the portfolio exposure to contagion risks. Because jump risk factors drive the latent state variables as well as the asset prices, risky assets can be used to hedge the uncertainties in the state variables, which gives rise to the additional term in optimal portfolio jump exposure. $\theta_t^{C_i}$ therefore serves as an intertemporal hedging demand for jump risks.

The contagion risk exposure, $\theta_t^{C_i}$, can be decomposed further into exposure to time series contagion risk, $\theta_t^{\mathrm{ts}_i}$, as a result of self excitation of jump component i, and exposure to cross section contagion risk,

$\theta_t^{\text{cs}_{i,j}}$, as a result of cross section excitation from jump component i to jump component j, $j \neq i$.

3.2 Comparative statics analysis of the optimal portfolio exposure to contagion risks

In this section, we numerically show how the optimal portfolio exposure to contagion risks $\theta_t^{C_i}$ depends on the stochastic characteristics of the jump intensities. If risks are compensated properly so that investors long risky assets, then the exposure to Poisson jump risks θ^{J_i} is negative. The exposure to contagion risks increases the overall demand for jump risk in region i, if it has the same sign as θ^{J_i}. For better interpretation, we plot the negative of the exposure to contagion risks of the jump component N_i, $-\theta^{C_i}$, at the beginning of the investment horizon, understood as the hedging demand of the jump component. We suppress the time subscript to indicate that the exposure to contagion risks is evaluated at the beginning of the investment horizon. Larger hedging demand implies larger stake of the region in the portfolio.

We implement the model in a two-region market and calculate the optimal portfolio exposure to contagion risks given in Equation (23). We fix a set of base case parameter values and conduct comparative statics analysis of parameters of the intensity processes. Specifically, we set the mean reversion rate at $\alpha = 21$, the investment horizon at $T = 1$. We specify a symmetric excitation matrix with reasonable values, $\boldsymbol{\beta} = (15, 3; 3, 15)$, according to the parameter estimates in Aït-Sahalia et al. (2014). We also impose identical jump risk premia $\kappa_1 = \kappa_2 = 0.3$, so that the jump components of the two regions are not distinguishable. Then we only need to analyze the contagion exposure to one of the jump components. This allows us to see how parameters affect contagion exposure as straightforwardly as possible.

In addition, we fix the risk aversion parameter to be $\gamma = 3$. As noted by Liu (2007), when $\gamma > 1$, investors are more risk averse than those with a log utility function and choose to hedge changes in the state variables. When $\gamma < 1$, investors not only forgo the hedging potential, but seek the high risk premium by betting on the future outcome.[9] The comparative statics analysis in Figure 2 and 3 for investors with $\gamma < 1$ has opposite patterns to those with $\gamma > 1$, as is the case in Liu and Pan (2003) and Liu et al. (2003). Since it is unlikely that investors have such small risk aversion, we restrict our analysis to the case where $\gamma > 1$.

[9] Note that the investors with $0 < \gamma < 1$ are still risk averse.

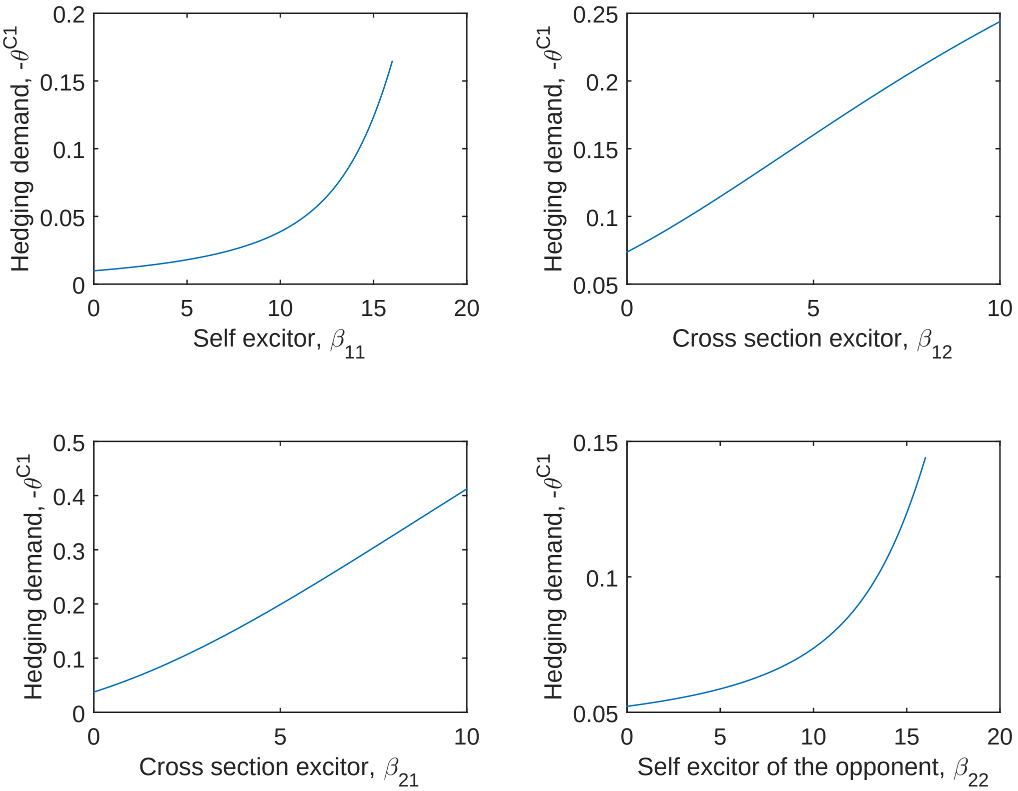

Figure 2: Comparative statics analysis of the hedging demand of jump component 1, $-\theta^{C_1}$. The market consists of two regions with identical jump risk factors. The hedging demand of jump component 1 (the other will be symmetric) is plotted as functions of elements in the excitation matrix $\boldsymbol{\beta}$. The base case parameters are $\alpha = 21,\ \boldsymbol{\beta} = (15, 3; 3, 15),\ T = 1,\ \kappa_1 = \kappa_2 = 0.3,\ \gamma = 3$.

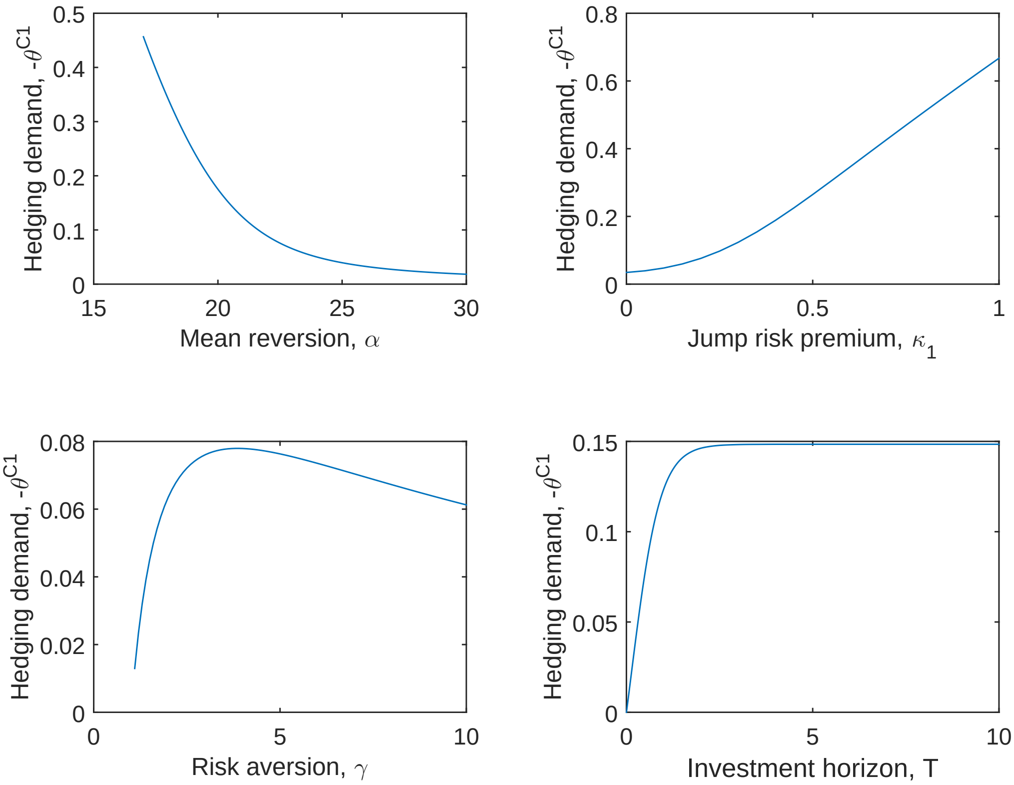

Figure 3: Comparative statics analysis of the hedging demand of jump component 1, $-\theta^{C_1}$. The market consists of two regions with identical jump risk factors. The hedging demand of jump component 1 (the other will be symmetric) is plotted as functions of the mean reversion rate α, jump risk premium κ_1, risk aversion γ and investment horizon T. The base case parameters are $\alpha = 21,\ \boldsymbol{\beta} = (15, 3; 3, 15),\ T = 1,\ \kappa_1 = \kappa_2 = 0.3,\ \gamma = 3$.

Figure 2 plots the hedging demand of jump component 1, $-\theta^{C_1}$, as functions of excitation parameters. The figure shows that increasing any element of the excitation matrix $\boldsymbol{\beta}$ leads to increasing demand for jump component 1, whether it be the self excitor of jump component 1, β_{11} (top left), the cross section excitor from jump component 1 to component 2, β_{21} (bottom left), the cross section excitor from jump component 2 to component 1, β_{12} (top right), or the self excitor of the opponent jump component β_{22} (bottom right).

Figure 3 plots the hedging demand of jump component 1, $-\theta^{C_1}$, as functions of the mean reversion rate α (top left), jump risk premium κ_1 (top right), risk aversion γ (bottom left) and investment horizon T (bottom right). Larger risk premium and longer investment horizon result in increasing jump risk demand. On the contrary, faster mean reversion rate decreases the exposure to contagion risks and in turn decreases the demand for jump risk. Interestingly, increasing the risk aversion first increases then decreases the hedging demand.

Since a jump occurrence moves asset prices and state variables in opposite directions, jump excitation enables the risky assets to be used as a static hedge against the effects of jumps in the state variables, as pointed out by Liu, Longstaff, and Pan (2003). For instance, an occurrence in jump component i decreases the asset price (since $z_i^k < 0$) but increases the state variables $\boldsymbol{\lambda}_t$ by $\boldsymbol{\beta}_i \geq 0$.[10] When $\gamma > 1$, investors take extra exposure to jump risks to hedge changes in the state variables to reduce uncertainties of the indirect utility.

Larger excitation and slower mean reversion imply that the jump intensity process (11) is more volatile. As one may expect, the more uncertainty there is in the state variables, the larger hedging incentive investors have. Larger risk premium results in larger exposure to Poisson jump risks $\boldsymbol{\theta}^J$, which leads to larger jump risks in the portfolio to be hedged. Similarly, longer investment horizon leads to increased sensitivity of indirect utility to state variables. In short, hedging demand rises when there are increasing uncertainties in investor's indirect utility.

The effect of increasing the risk aversion, however, is not clear. On one hand, increasing the risk aversion decreases the exposure to Poisson jump risk $\boldsymbol{\theta}^J$, implying a smaller amount to be hedged thereby decreasing the hedging demand. On the other hand, a more risk averse investor is more inclined to hedge the changes in the state variables, and has a larger hedging demand. The final result depends on which effect is larger. Figure 3 shows that the effect of increasing risk aversion is not monotone: it first increases the jump risk demand and then reduces it.

3.3 The effect of asymmetric excitation

Interesting phenomena arise when the excitation structure becomes asymmetric. The excitor $\beta_{ji}, j \neq i$, measures the excitation capability of jump component N_i as the source jump component, i.e., how large an occurrence in N_i raises the intensities of another jump component j, whereas the excitor $\beta_{ij}, j \neq i$, measures the inclination to excitation of jump component N_i as the target jump component, i.e., how large an occurrence in some jump component j raises the intensity of N_i. As we see in reality, equity prices in other geographical markets usually crash in close succession with an equity plunge in the US, whereas the transmission in the reverse order is not as often observed. It implies that $\beta_{ji} > \beta_{ij}$ when N_i represents the jump component in the US equity.

Recall from Equation (23) that $\theta^{C_i} = \beta_{ii} B_i + \sum_{j \neq i}^n \beta_{ji} B_j$. The portfolio exposure to contagion risks of jump component i depends on the potential of jump component i to excite other jump components as well as itself. Observe that β_{ji} plays a different role from β_{ij} in determine θ^{C_i}: a larger cross section excitor

[10] The fact that asset prices and jump intensities jump in different directions is essential to generate hedging demand. If the equity jump $z_i^k > 0$, investors take smaller exposure to jump risks as a result of jump excitation. This result is consistent with Liu, Longstaff, and Pan (2003).

β_{ij} leads to a larger increase in the jump intensity of region i (as a result of a jump occurrence in region j), and a larger portfolio exposure to the jump risk factor of region j (instead of region i). Regions with comparable expected jump intensities may be weighted differently in the optimal investment portfolio due to asymmetric excitation.

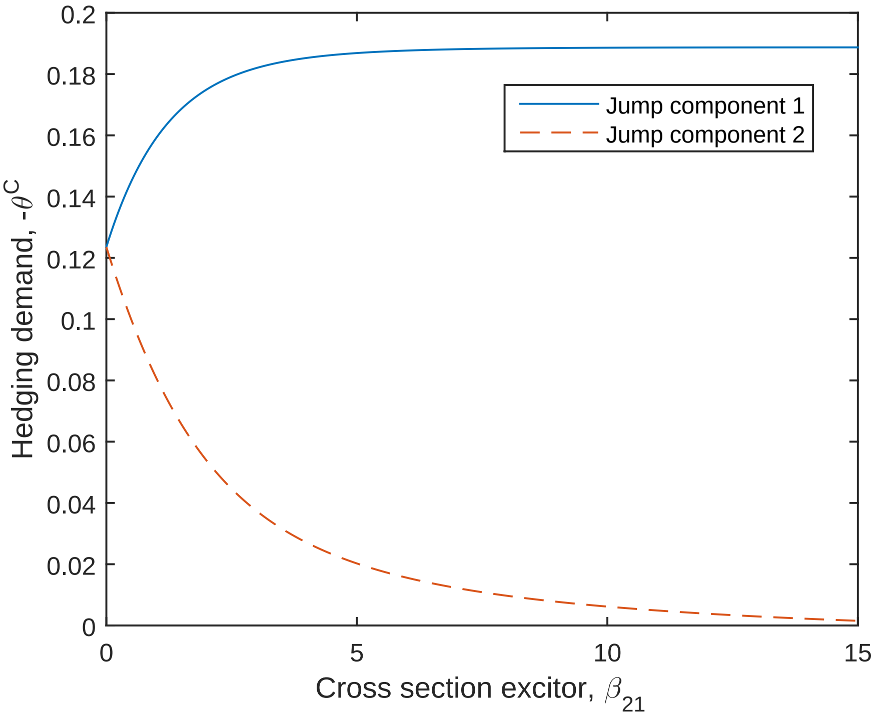

Figure 4: The hedging demand of the two jump components $-\theta^{C_1}$ (solid curve), $-\theta^{C_2}$ (dotted curve) as functions of the cross section excitor β_{21}. The excitation matrix is $\boldsymbol{\beta} = (18, 0; \beta_{21}, \beta_{22})$. We let β_{21} increase and find the corresponding β_{22} such that the expected jump intensities do not vary with the excitation matrix. The jump risk premium is set to be equal with $\kappa_1 = \kappa_2 = 0.3$. The other parameters are $\alpha = 21$, $\gamma = 3$, $T = 1$. Given the choice of parameter values, the expected jump intensity is $E[\lambda_{1,t}] = E[\lambda_{2,t}] = 2.1$.

Figure 4 gives a numerical illustration of the effect of excitation asymmetry in a two-region market. The figure plots the hedging demand $-\theta^{C_1}$ (solid curve), $-\theta^{C_2}$ (dotted curve) as functions of β_{21}. It shows how investors' demand for jump risks changes as the excitation structure becomes more asymmetric. We fix the first row of the excitation matrix to be $\beta_{11} = 18, \beta_{12} = 0$. We close the excitation channel from region 2 to region 1 by setting $\beta_{12} = 0$, so that the jump risk only propagates from region 1 to region 2 but not the other way round. We let β_{21} increase while finding the corresponding β_{22} that delivers the same expected jump intensity $E[\lambda_{2,t}]$. As β_{21} increases, the excitation matrix gets more asymmetric due to a larger difference between β_{21} and β_{12}, strengthening the bias towards region 1, with everything else (e.g., jump risk premia, expected jump intensities) unchanged. We observe from the figure that jump component 1 becomes more appealing to the investor as contagion becomes more asymmetric, even though the jump components have the same risk profile in all other aspects.

As mentioned in the previous section, jump excitation increases demand for jump risks because of their hedging potential. Consistent with this intuition, when jump components have heterogeneous capabilities in raising jump intensities, the jump component with a higher excitation capability is the more favorable risk factor due to its larger hedging potential against jump intensities. The end point on the x-axis in Figure 4 stands for $\boldsymbol{\beta} = (18, 0; 15, 3)$. An occurrence in jump component 1 raises $\lambda_{1,t}$ by 18, and $\lambda_{2,t}$ by 15, while an occurrence in jump component 2 only raises $\lambda_{2,t}$ by 3. Apparently, jump component 1 has a larger influence on the state variables (jump intensities) than jump component 2 and consequently larger hedging potential. Investors therefore tilt the portfolio towards region 1 for a more effective hedge of the state variables. One may expect that, everything else equal, the US equity will take a larger share in the investor's portfolio as compared to the classic model predictions due to excitation asymmetry.

4 Utility loss of suboptimal trading strategies

In the previous sections we have shown that jump propagation changes the risk profile of the optimal portfolio. In particular, we see that excitation asymmetry leads to larger investment towards the equity market which is capable of spreading jump risks. In this section, we examine the utility loss for an investor who fails to construct the equity portfolio optimally. We only focus on the case when the suboptimal portfolio is sufficiently diversified, i.e., $\boldsymbol{\theta}^Z \equiv 0$. A suboptimal portfolio is defined as the equity portfolio whose risk exposure to the systematic Brownian and jump risk factors are different from the optimal exposure. To quantify the utility loss of implementing suboptimal strategies, we adopt the measure introduced in Liu and Pan (2003). The utility loss L of a suboptimal portfolio $\hat{x}$ is defined as

$$L = \frac{\log x^* - \log \hat{x}}{T}, \tag{25}$$

where $x^*(\hat{x})$ is the certainty equivalent wealth of the optimal (suboptimal) portfolio strategy defined as

$$\begin{aligned}
\frac{x^{*1-\gamma}}{1-\gamma} &:= E_{0,x,\boldsymbol{\lambda}}\Big[\frac{X_T(\boldsymbol{\theta}^{W*}, \boldsymbol{\theta}^{N*})^{1-\gamma}}{1-\gamma}\Big] = J(0, x, \boldsymbol{\lambda}), \\
\frac{\hat{x}^{1-\gamma}}{1-\gamma} &:= E_{0,x,\boldsymbol{\lambda}}\Big[\frac{X_T(\hat{\boldsymbol{\theta}}^{W}, \hat{\boldsymbol{\theta}}^{N})^{1-\gamma}}{1-\gamma}\Big],
\end{aligned}$$

where $X_T(\cdot,\cdot)$ emphasizes the dependence of the terminal wealth on the portfolio risk exposure.

Effectively, L measures the loss in terms of the annualized, continuously compounded return in certainty equivalent wealth of the suboptimal portfolio strategy. The larger L is, the larger the utility loss of implementing the suboptimal portfolio strategy.

The next proposition computes the utility loss of implementing the portfolio strategy with a suboptimal risk exposure $(\hat{\boldsymbol{\theta}}_t^W, \hat{\boldsymbol{\theta}}_t^N)$.

Proposition 3 (Utility loss of suboptimal strategies)**.** *If the expected power utility investor implements a suboptimal portfolio strategy with risk exposure* $(\hat{\boldsymbol{\theta}}_t^W, \hat{\boldsymbol{\theta}}_t^N)$*, then the associated utility loss is given by*

$$L = \frac{1}{(1-\gamma)T}\Big(\gamma A(0) - \hat{A}(0) + (\gamma \boldsymbol{B}(0) - \hat{\boldsymbol{B}}(0))'\boldsymbol{\lambda}_0\Big), \tag{26}$$

where $\boldsymbol{B}(0), A(0)$ *are given by* (19)*, and* $\hat{\boldsymbol{B}}(0), \hat{A}(0)$ *can be solved from*

$$\dot{\hat{\boldsymbol{B}}}(t) =(1-\gamma)(1+\boldsymbol{\kappa}) \circ \hat{\boldsymbol{\theta}}_t^N + \alpha \hat{\boldsymbol{B}}(t) - (\hat{\boldsymbol{\theta}}_t^N + 1)^{1-\gamma} \circ e^{\boldsymbol{\beta}'\hat{\boldsymbol{B}}} + 1, \tag{27}$$

$$\dot{\hat{A}}(t) = -(1-\gamma)\Big(\hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{\eta} - \frac{1}{2}\hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{L}\boldsymbol{L}'\hat{\boldsymbol{\theta}}_t^W\Big) - \frac{1}{2}(1-\gamma)^2\hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{L}\boldsymbol{L}'\hat{\boldsymbol{\theta}}_t^W - \alpha\hat{\boldsymbol{B}}'\boldsymbol{\lambda}_\infty, \tag{28}$$

with $\hat{\boldsymbol{B}}(T) = \mathbf{0}, \hat{A}(T) = 0$*.*

Two relevant cases are when the investor fails to recognize the exciting nature of jump components and implements the portfolio strategy as if the equity price is generated by Poisson jump diffusion, and when the investor simply ignores the discontinuities in equity returns and implements the Merton mean-variance strategy. To calculate the utility loss associated with theses suboptimal strategies, according to proposition 3, we simply replace $\hat{\boldsymbol{\theta}}_t^W$ in Equation (28) by $\boldsymbol{\theta}^{W*}$, and $\hat{\boldsymbol{\theta}}_t^N$ in Equation (27) by $\boldsymbol{\theta}^J$, $\hat{\boldsymbol{\theta}}^J$, respectively.

Figure 5 plots the utility loss of the aforementioned two cases as functions of the currency jump intensity (top left), jump risk premium κ_1 (top right), risk aversion γ (bottom left) and investment horizon T (bottom right). The utility loss when the investor incorrectly recognizes the return generating model as Poisson jump diffusion is plotted in the solid curve, and the case when the investor incorrectly implements the Merton mean-variance strategy is plotted in the dotted curve. Notice that the two curves move together with the PJD curve above the Merton curve most of the time.

Although neither the optimal portfolio nor the suboptimal portfolio depends on the realization of the state variables, an investor's utility over terminal wealth is dependent on the current jump intensities. As a result, utility loss is sensitive to the current jump intensities, as shown in the upper left panel of Figure 5. During a financial crisis, jump intensities may build up to as large as 100 or higher as found by Aït-Sahalia et al. (2014). Ignoring jump excitation then costs a loss of over 4000 basis points in annualized portfolio returns.

The upper right panel plots the utility loss as a function of the jump risk premium, κ_1. The graph shows that the utility loss increases with the jump risk premium. For instance, when the jump risk premium $\kappa_1 = 0.5$ (which gives an equity jump premium of 5% as estimated by Bollerslev and Todorov (2011)), the expected annual return of a portfolio strategy that does not account for jump excitation is 1500 basis points lower than the true optimal portfolio. Recall from Figure 3 that the jump risk premium increases contagion exposure. A larger jump risk premium leads to larger deviations from the optimal jump exposure, which in turn leads to bigger utility loss. The same effect holds for the investment horizon shown in the lower left panel.

The lower right panel plots the utility loss as a function of the risk aversion γ. When $\gamma = 1$, the investor has log utility. Log utility investors are myopic, in the sense that they only care about the current realization of the state variables, and therefore do not have an incentive to hedge against future changes of the state variables. As a result, ignoring jump excitation generates no utility loss in PJD strategy for a log utility investor. The utility loss of the Merton strategy, however, is non-trivial even when the investor has log utility. As the investor becomes more risk averse, the utility loss of both suboptimal strategies start to increase. This is because the more risk averse the investor is, the more concerned he/she is about the changes in the state variables as a result of jump excitations.

A surprising fact is that the utility loss of ignoring jump excitation is even larger than that of ignoring jumps in total (except when γ is close to 1). In other words, if the true return generating model is mutually exciting jump diffusion, then it is better for the investor not to account for jumps at all than recognize jumps but mistake them for the wrong type. In Section 3.1, we have shown that the exposure to a risk factor is smaller when it is recognized as a Poisson jump than a Brownian motion. However, when jumps are mutually exciting, the investor increases the exposure to the jump risk factor in order to exploit the hedging potential inherited. The Poisson jump diffusion strategy turns out to be too conservative.

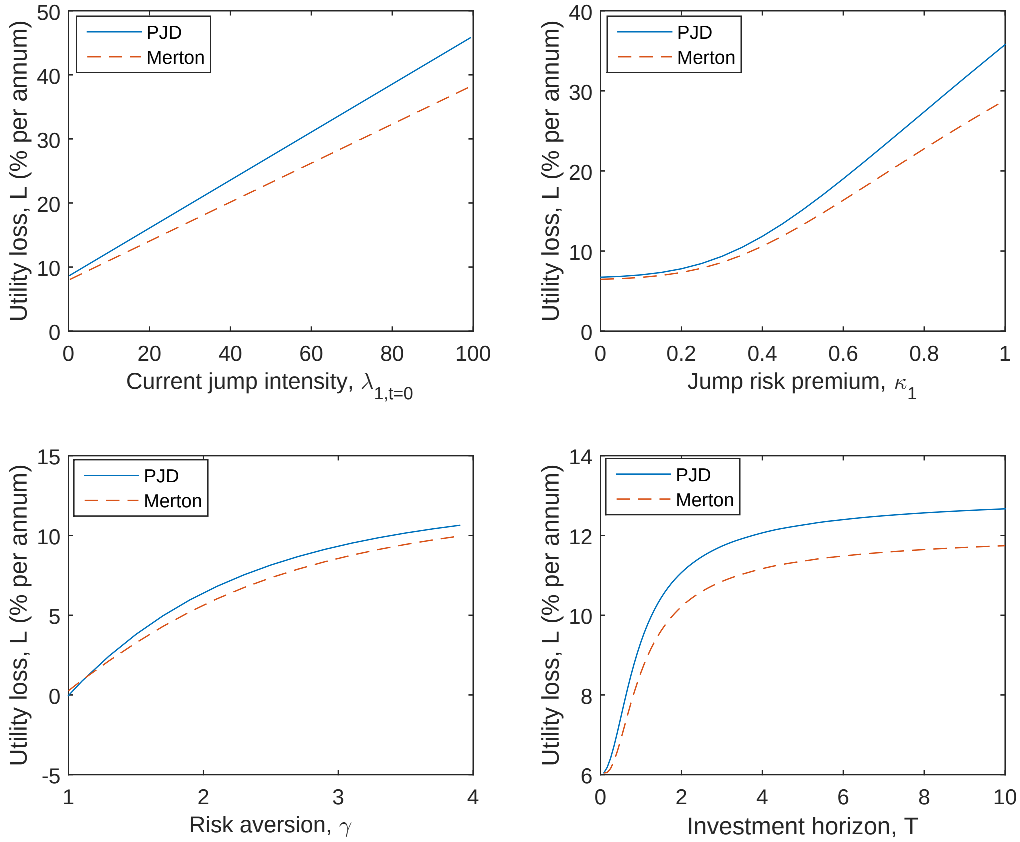

Figure 5: Utility loss in terms of the annualized continuously compounded return (% per annum) if the investor incorrectly assumes $\boldsymbol{\beta} = 0$ and implements the portfolio strategy as if the model is Poisson jump diffusion (solid curve), and if the investor implements the Merton mean-variance strategy (dotted curve). The "true" base case parameters are $\boldsymbol{\eta} = (0.3; 0.3)$, $\alpha = 21$, $T = 1$, $\boldsymbol{\lambda}_{\infty} = (0.3; 0.3; 0.3)$, $\boldsymbol{\kappa} = (0.3; 0.3)$, $\boldsymbol{\lambda}_{t=0} = (2; 2)$, $\boldsymbol{\beta} = (15, 3; 3, 15)$, $\gamma = 3$. The top left panel plots the utility loss as a function of the current value of the intensity of jump component 1, $\lambda_{1,t=0}$, while all other parameters remain the same. The top right graph plots the utility loss as a function of the risk premium of jump component 1, κ_1. The bottom left panel plots the utility loss as a function of the risk aversion γ. The bottom right graph plots the utility loss as a function of the investment horizon T.

5 Application to international equity returns

In this section, we estimate the mutually exciting jump diffusion model to index returns of US, Japan and Europe. We show that the underdiversification of the market portfolio, especially the over-weighting of the US equity of the market portfolio compared to classic asset allocation models, can be explained by an asymmetric excitation structure. In Appendix C, we also show how the portfolio tilts towards the home market for country representative investors who are ambiguity averse towards foreign equity markets. In Section 5.1, we describe the equity index data we used in the empirical analysis. We disentangle jumps from continuous returns and estimate the diffusion and jump parameters in Section 5.2. Section 5.3 compares the empirical market portfolio exposure to risk factors to model predictions and shows that excitation asymmetry is able to generate the observed US bias in the market portfolio.

5.1 Data

We consider three well-developed stock markets: the United States, Japan, and Europe, and make the simplifying assumption that these regions represent the global financial market. We choose these three equity markets because: (1) these equity markets have little barriers to international investment; (2) these three equity markets already represented around 63% of the world equity market capitalization at the end of 2012. More markets could be included theoretically, but parameter estimation is likely to become cumbersome. Therefore we do not go beyond three equity markets in the empirical analysis.[11]

We examine excess returns on Morgan Stanley Capital International (MSCI) indices of US, Japan and Europe, over *local* risk free rates in *local* currencies. The MSCI indices are value weighted stock indices. There are 609 individual stocks included in MSCI US index, 320 stocks in MSCI Japan index, and 434 stocks in MSCI Europe index by the end of 2012. Expected returns are estimated as the sample mean of the log returns on the MSCI *total return* index from Jan 1970 to Dec 2012. The total return index has been adjusted for dividends and other noncash payments to shareholders. We estimate the covariance and jump parameters using the *daily price* index from Jan 3$^{\text{rd}}$, 1972 to Dec 31$^{\text{th}}$, 2012, a total of 10696 observations.[12] Excitation parameters would not be estimated accurately on less frequent data such as weekly or monthly, since multiple jumps could happen within adjacent observations. We use US 3 month Treasury bill rates, Japan base discount rates, UK 3 month Libor rates as proxies for the local risk free rates of the three regions.

Table 2 contains the descriptive statistics of the annualized excess returns on MSCI indices. Japan has the lowest mean return of 2.45% and Europe the highest with 4.16% on an annual basis. Return volatilities vary from around 15% to 19%. Comparing the risk return tradeoffs of these three major equity markets, the European market generates a fairly high mean with the lowest volatility. By contrast, the Japanese equity is characterized by the lowest expected return and the highest volatility. In spite of

[11] Similar assumptions that a few representative markets make up the world economy are also made in, for example, Ang and Bekaert (2002), Uppal and Wang (2003), Das and Uppal (2004). A more extensive empirical analysis is beyond our scope and left to future research.

[12] We do not model dividends separately in our model. Preferably we would like to use daily returns on the *total return* index in all estimations. Unfortunately, the *daily* data of total returns of MSCI indices are only available since 2001. We compare the sample covariance of the daily returns of price index and total return index in the overlapping period. They turn out to be almost identical. Same holds for the jump intensities. We conclude that on an index level, most actions in daily total returns come from price moves, not dividends.

the unfavorable risk-return tradeoff of Japanese equity, the correlation between returns on the Japanese market and those on the US market is as small as 0.11. The correlation between returns of Japan and Europe is also lower than that between US and Europe. The relatively small dependence between the Japanese market and the other two equity markets makes the Japanese equity a better candidate for international diversification. All returns are left-skewed, implying larger extreme losses than gains. For all regions, the excess kurtosis is substantially larger than 0, as would be caused by jumps.

Descriptive statistics of annualized excess returns

	US	JA	EU
Number of constituents	609	320	434
Mean	4.12%	2.45%	4.16%
Standard deviation	17.50%	18.81%	15.26%
Correlation	1	0.11	0.46
		1	0.31
			1
Skewness	-1.04	-0.31	-0.41
Excess kurtosis	25.96	12.05	9.86

Table 2: Descriptive statistics of annualized excess log returns on MSCI indices over local risk free rates. The sample mean is computed using the total returns data from Jan 1970 to Dec 2012. Higher moments are computed using daily price index data from Jan 3$^{\text{rd}}$ 1972 to Dec 31$^{\text{st}}$ 2012.

5.2 Parameter estimation of nested models

The equity indices, by construction, are well diversified local portfolios of a large number of individual assets. We henceforth assume that the equity indices are representative assets free of idiosyncratic risks and follow the dynamics

$$\frac{dS_{i,t}}{S_{i,t^-}} = \sigma_i(dW_{i,t} + \eta_i dt) + z_i(dN_{i,t} - (1+\kappa_i)\lambda_{i,t}dt), \quad i = 1,2,3, \tag{29}$$

with instantaneous 3 by 3 covariance matrix of local portfolios $\mathbf{\Sigma} = \boldsymbol{\sigma}\boldsymbol{L}\boldsymbol{L}'\boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ is a diagonal matrix with σ_i on the diagonals.

We consider the following nested models: the diffusion only model, Poisson jump diffusion model (PJD), self exciting jump diffusion model (SEJD) and finally the full-fledged mutually exciting jump diffusion model (MEJD). Table 3 lists nested models as special cases of the mutually exciting jump diffusion model with proper parameter restrictions. If we restrict jump amplitudes $\boldsymbol{z} = (z_1; z_2; z_3) = \mathbf{0}$, i.e., asset prices do not jump at all, we have the classic diffusion only model in which asset prices follow geometric Brownian motions. The model becomes a Poisson jump diffusion model if $\boldsymbol{\beta} = \mathbf{0}$, implying that jumps do not excite and intensities are kept constant at the level $\boldsymbol{\lambda}_\infty$. More generally, if we only close the cross section excitation channels and restrict $\boldsymbol{\beta}$ to be a diagonal matrix, asset prices will follow self exciting jump diffusion processes. In the most general case where no restrictions are imposed, we have a full-fledged mutually exciting jump diffusion model.

Nested models as special cases

Nested models	Parameter restrictions
Diffusion only	$\boldsymbol{z} = \mathbf{0}$
PJD	$\alpha = 0, \boldsymbol{\beta} = 0$
SEJD	$\boldsymbol{\beta}$ diagonal
MEJD	none

Table 3: Nested models as special cases of the mutually exciting jump diffusion model with proper parameter restrictions. "PJD" stands for the Poisson jump diffusion model; "SEJD" represents the self exciting jump diffusion model; "MEJD" represents the mutually exciting jump diffusion model.

We are going to estimate each model listed in Table 3 using the historical returns on MSCI indices. Parameter estimates of the diffusion only model can be easily obtained through the first and second moments of the log returns reported as the summary statistics in Table 2. For jump diffusion models, similar to Liu, Longstaff, and Pan (2003), we disentangle jumps from the continuous part of log returns in order to estimate diffusion and jump parameters separately.[13] We define truncation thresholds as negative three times the sample standard deviation. Daily log returns that fall below the thresholds are regarded as jump returns. Figure 6 plots the filtered jump occurrences in US, Japan and EU. We observe jump clustering during the Asian crisis (1997-1999), the stock market downturn (2002) and the Subprime mortgage crisis (2007-2012) .

[13]Jumps are infrequent by nature. The parametrization of the mutually exciting jump diffusion model is rich and econometrically challenging. Ideally, we would like to apply the Generalized Method of Moments used by Aït-Sahalia et al. (2014) which minimizes the effects of the "Peso problem" inherited. However, giving the 3-dimensional nature of our problem, we use this informal two-step maximum likelihood approach to get some idea of the excitation structure of equity markets. A formal econometric treatment is beyond the scope of this paper.

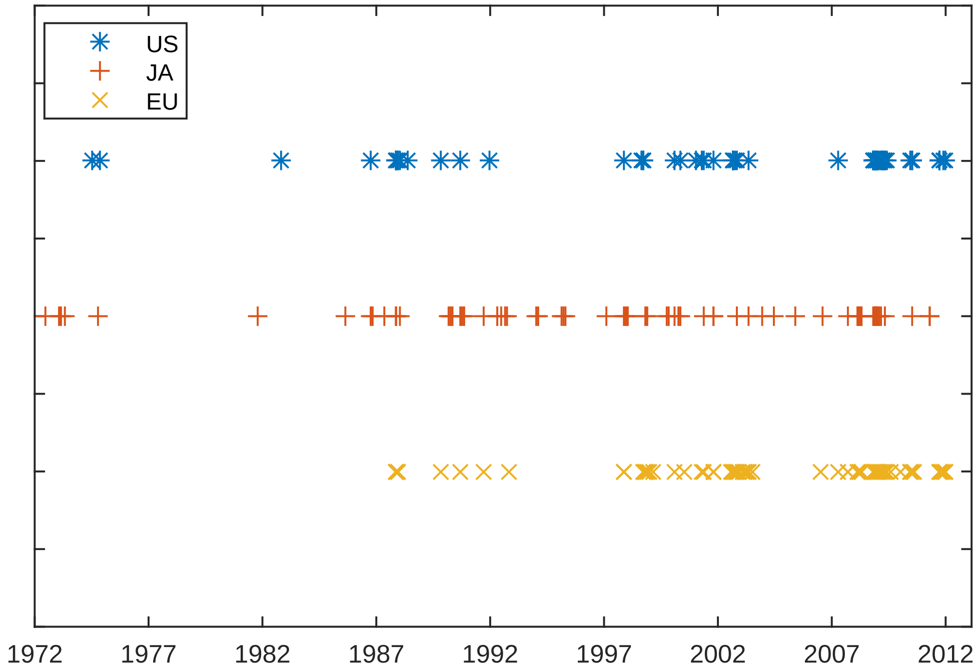

Figure 6: Filtered jump occurrences in US, Japan and EU. Each mark indicates a jump occurrence. The top row stands for the jump occurrences in the US equity market, the middle row Japan and the bottom row Europe. Daily log returns that fall below negative three times the sample standard deviation thresholds are recorded as jump occurrences.

Parameter estimates of the nested jump diffusion models are reported in Table 4. Jump amplitudes of log returns, $\log(1+\boldsymbol{z})$, are estimated as the differences between average jump returns and non-jump returns. In a Poisson jump diffusion model, the constant Poisson jump intensities are obtained by dividing the sum of total jump occurrences detected in each MSCI return by the number of years. In case of self excitation and mutually excitation, jump excitation parameters $\alpha, \boldsymbol{\beta}$ are estimated using the maximum likelihood, while $\boldsymbol{\lambda}_\infty$ is estimated such that the unconditional expected jump intensity $E[\boldsymbol{\lambda}]$ is equal to the average jump occurrences per year. The algorithm of computing the likelihood functions is discussed in Appendix D. Having identified the jumps, we construct the truncated returns by removing detected jumps from returns, so that the truncated data can be regarded as being generated by the continuous part of the model. We estimate the instantaneous volatility $\boldsymbol{\sigma}$ and the correlation matrix $\boldsymbol{LL}'$ from the truncated data for models with jumps.

We see from Table 4 that equities with higher volatility have larger jump amplitudes but not necessarily more frequent jumps. This is not unreasonable since volatility measures the variation in the bulk of the data, while jumps contribute to the very left tails of the return distribution. Among three regions, the

Japanese market has the largest jump amplitude, -4.97% with an average of 2.02 jumps per year. Europe has the smallest jump amplitude -4.07%, but has the most frequent jumps – average 2.63 jumps occur per year. The US equity market has moderate jump amplitude and jump frequency. Jumps display statistically significant self excitation as well as cross section excitation. The jump propagation from US to the other two regions are significant and large in magnitude. US, on the other hand, can only be excited by itself or the European market. The spillover effect from Japan to the US is almost zero. No sound evidence of the cross section excitation between Japan and Europe is found. The estimated excitation structure is in line with the pairwise estimation results in Aït-Sahalia et al. (2014), who show that US always has a larger cross section excitor as the source jump component than as the target jump component when paired with other economies.

5.3 Empirical vs implied portfolio exposure to risk factors

In this section, we calculate the exposure to risk factors of the market portfolio and compare it to the model predictions.

First, we infer the market exposure to systematic risk factors using market portfolio weights. In Section 2.4, we show that the optimal wealth for an expected power utility investor is given by

$$\frac{dX_t^*}{X_{t^-}^*} = \sum_{i=1}^{n} \Big\{ \theta^{W_i*}(dW_{i,t} + \eta_i dt) + \big(\exp(\theta_t^{N_i*}) - 1\big)\big(dN_{i,t} - (1+\kappa_i)\lambda_{i,t}dt\big) \Big\}.$$

Denote the weight on the local equity index $S_{i,t}$ in the market equity portfolio M_t by $h_{i,t}$, with $\sum_i h_{i,t} = 1$. The dynamics of $S_{i,t}$ are given by Equation (29). It holds that

$$\frac{dM_t}{M_{t^-}} = \sum_{i=1}^{n} h_{i,t}\Big(\frac{dS_{i,t}}{S_{i,t^-}}\Big) \tag{30}$$

$$= \sum_{i=1}^{n} h_{i,t}\Big(\sigma_i(dW_{i,t} + \eta_i dt) + z_i(dN_{i,t} - (1+\kappa_i)\lambda_{i,t}dt)\Big). \tag{31}$$

In equilibrium, the market portfolio M_t should reflect the optimal wealth X_t^* of the representative investor:

$$M_t \approx X_t^*. \tag{32}$$

Notice that we have an approximation instead of an equality here. As shown in Section 2.5, the replicated portfolio converges to the optimal wealth process as the number of assets goes to infinity. In reality, unfortunately, the equity indices are made up of finite (although many) individual assets. The market portfolio M_t, therefore, are not completely free of idiosyncratic risks.

Parameter estimates of nested jump diffusion models

Parameters	PJD	SEJD	MEJD
$\boldsymbol{\sigma}$		$\begin{pmatrix} 16.17\% & 0 & 0 \\ 0 & 17.38\% & 0 \\ 0 & 0 & 13.71\% \end{pmatrix}$	
$\boldsymbol{LL'}$		$\begin{pmatrix} 1 & 0.10 & 0.40 \\ 0.10 & 1 & 0.25 \\ 0.40 & 0.25 & 1 \end{pmatrix}$	
$\boldsymbol{z}$		$\begin{pmatrix} -4.93\% \\ -4.97\% \\ -4.07\% \end{pmatrix}$	
$\boldsymbol{\lambda}_\infty$	$\begin{pmatrix} 1.75 \\ 2.02 \\ 2.63 \end{pmatrix}$	$\begin{pmatrix} 0.57 \\ 1.03 \\ 0.58 \end{pmatrix}$	$\begin{pmatrix} 0.47 \\ 0.77 \\ 0.31 \end{pmatrix}$
α	0	17.52^{***}	29.35^{***}
$\boldsymbol{\beta}$	$\mathbf{0}$	$\begin{pmatrix} 11.80^{***} & 0 & 0 \\ 0 & 8.55^{***} & 0 \\ 0 & 0 & 13.66^{***} \end{pmatrix}$	$\begin{pmatrix} 13.07^{***} & 0.00 & 5.60^{**} \\ 9.04^{***} & 6.96^{***} & 2.52 \\ 24.87^{***} & 2.60 & 7.35^{***} \end{pmatrix}$

Table 4: Parameter estimates of nested jump diffusion models. “PJD” stands for the Poisson jump diffusion model; “SEJD” represents the self exciting jump diffusion model; “MEJD” represents the mutually exciting jump diffusion model. *, **, *** indicate significance at 95%, 97.5%, and 99.5% confidence levels, respectively.

Equation (32) implies that

$$\begin{cases} \theta^{W_i*} \approx \sigma_i h_{i,t}, \\ \theta_t^{N_i*} \approx z_i h_{i,t}. \end{cases} \tag{33}$$

Note that Equation (33) holds as long as assets follow jump diffusion processes, regardless of whether the jumps are mutually exciting, self exciting or Poissonian.

Table 5 reports the normalized portfolio weights and (approximated) exposure to risk factors, $\frac{\boldsymbol{\theta}^{W*}}{\boldsymbol{\iota}'\boldsymbol{\theta}^{W*}}$ and $\frac{\boldsymbol{\theta}^{N*}}{\boldsymbol{\iota}'\boldsymbol{\theta}^{N*}}$, of the market portfolio. The market values (US dollar denominated) of MSCI US, Japan and Europe at the end of 2012 serve as proxy of the market portfolio. The portfolio weights are calculated by dividing the market values of each region by the sum of the market values, so that the weights on US, Japan and Europe add up to 1. Risk exposure is calculated using the approximation (33). The reason that the portfolio exposure to Brownian risks and that to jump risks are similar is that a region with higher volatility has on average larger jump amplitude (see Table 4). Normalization further narrows the differences. In fact, the over-exposure to the US Brownian risks and to the US jump risks stem from different reasons, as we will see shortly.

Empirical market portfolio weights and exposure to risk factors

	US	JA	EU
Portfolio weights	0.58	0.10	0.32
Exposure to systematic Brownian risks	0.60	0.11	0.29
Exposure to jump risks	0.61	0.11	0.28

Table 5: Empirical market portfolio weights and exposure to risk factors as of the end of 2012. The market weights are computed from the MSCI market values. Risk exposures are calculated according to Equation (33). Statistics are normalized to add up to 1 on the rows.

Next, we derive the model predicted risk exposure. As discussed in Section 3.1, optimal exposure to mutually exciting jumps can be decomposed into three components (see Equation (23)): exposure to Poisson jump risks, exposure to time series contagion risks, and exposure to cross section contagion risks. The optimal jump exposure predicted by any nested model, as a result, is a combination of these components, depending on the risk features of the jump factors. Table 6 lists the portfolio exposure to risk factors predicted by nested models. Notice that although the term $\boldsymbol{\theta}^{ts}$ (exposure to time series contagion risks) appears in the self exciting jump model as well as the mutually exciting jump model, they have different values due to the fact the excitation matrices of SEJD and MEJD are different not only in off-diagonal elements but also in diagonal elements (c.f. Table 4). For comparability, we also include the risk exposure of a pure diffusion model. In a pure diffusion model, the Brownian risk factors remain unchanged. Investors regard the jump factors as if they were generated by Brownian motions. $\hat{\boldsymbol{\theta}}^J$, the exposure to jump risk factors when jump components are recognized as Brownian motions, is given by (24).

Optimal risk exposure of nested models

Nested models	Brownian exposure	Jump exposure
Diffusion only	$\boldsymbol{\theta}^{W*}$	$\hat{\boldsymbol{\theta}}^J$
PJD	$\boldsymbol{\theta}^{W*}$	$\boldsymbol{\theta}^J$
SEJD	$\boldsymbol{\theta}^{W*}$	$\boldsymbol{\theta}^J + \boldsymbol{\theta}^{ts}(\boldsymbol{\beta}^D)$
MEJD	$\boldsymbol{\theta}^{W*}$	$\boldsymbol{\theta}^J + \boldsymbol{\theta}^{ts}(\boldsymbol{\beta}) + \boldsymbol{\theta}^{cs}(\boldsymbol{\beta})$

Table 6: Model predicted exposure to risk factors. "PJD" stands for the Poisson jump diffusion model; "SEJD" represents the self exciting jump diffusion model; "MEJD" represents the mutually exciting jump diffusion model.

Now that all models predict the same exposure to Brownian risks, we focus on the comparison of the model prediction on the exposure to jump risks. We calibrate the Brownian risk premium $\boldsymbol{\eta}$ such that $\boldsymbol{\theta}^{W*}$ coincide with the Brownian risk exposure of the market portfolio. To do that, we first decompose the equity premium into variance premium and jump premium. Let $\bar{\boldsymbol{\mu}}$ be the variance-corrected expected excess log return: $\bar{\boldsymbol{\mu}} = E[d\boldsymbol{S}_t./\boldsymbol{S}_{t^-}] = E[\log(\boldsymbol{S}_t)] + \frac{1}{2}E[(\log(\boldsymbol{S}_t) - E[\log(\boldsymbol{S}_t)])^2]$.[14] The total equity premium $\bar{\boldsymbol{\mu}}$ can therefore be divided into the variance premium and the jump premium:

$$
\begin{aligned}
\bar{\boldsymbol{\mu}} &= \bar{\boldsymbol{\mu}}_{\text{variance}} + \bar{\boldsymbol{\mu}}_{\text{jump}}, \\
&:= \boldsymbol{\sigma\eta} + (-\boldsymbol{\kappa} \circ \boldsymbol{z} \circ E[\boldsymbol{\lambda}_t]). \qquad (34)
\end{aligned}
$$

Since we consider the *relative* allocations to US, Japan and Europe, $\boldsymbol{\eta}$ is only identified up to a positive constant: $\boldsymbol{\eta} \propto \boldsymbol{LL}' \frac{\boldsymbol{\theta}^{W*}}{\boldsymbol{\iota}'\boldsymbol{\theta}^{W*}}$. The risk premium parameters reported in Table 7 are calibrated using

$$
\boldsymbol{\eta} = \boldsymbol{LL}' \frac{\boldsymbol{\theta}^{W*}}{\boldsymbol{\iota}'\boldsymbol{\theta}^{W*}}, \qquad (35)
$$

$$
\boldsymbol{\kappa} = (\boldsymbol{\sigma\eta} - \bar{\boldsymbol{\mu}})./\boldsymbol{z}./E[\boldsymbol{\lambda}_t]. \qquad (36)
$$

As shown in Table 7, the market portfolio exposure to Brownian risk implies a large variance premium in the US and a small variance premium in Japan. The jump premium, calculated as the equity premium less the variance premium, is comparable across regions, with Europe having the largest jump premium

[14]We are aware of the fact that the first moment of equity returns cannot be consistently estimated using a sample of 41 years alone. The main purpose here is to show that for given risk premium estimates, while acknowledging this fact, the estimated excitation structure gives rise to the US bias observed in the market portfolio. Other papers also use the empirical first moment for asset allocation purposes. See, among others, French and Poterba (1991), Liu, Longstaff, and Pan (2003), Liu and Pan (2003), Das and Uppal (2004), Jin and Zhang (2012).

and Japan the lowest.

Calibrated risk premium parameters

	η	κ	Variance premium	Jump premium	Total equity premium
US	0.11	0.42	1.82	3.60	5.43
JA	0.04	0.33	0.65	3.31	3.96
EU	0.09	0.37	1.17	3.93	5.10

Table 7: Calibrated risk premium parameters and the corresponding variance and jump risk premiums. The parameters $\boldsymbol{\eta}, \boldsymbol{\kappa}$ are calibrated using Equations (35), (36). The variance and the jump premium are calculated with (34) and are denoted in percentage per annum.

The jump risk premium in US reported in Table 7 is consistent with Pan (2002), who estimates the S&P 500 average mean excess rate of return demanded for the jump risk to be 3.5% per year. The estimates of the jump risk premium from other papers can vary. For instance, Bollerslev and Todorov (2011) non-parametrically estimate the average US jump risk premium to be approximately 5%. Santa-Clara and Yan (2010) find that the jump risk premium is on average more than half of the total equity premium. In a self-exciting jump diffusion model using the US equity and option data, Boswijk et al. (2015) estimate the jump risk premium to be around 8.82 times the spot jump intensity, implying a κ of the US market of 0.79. The actual choice of risk premiums does not have a qualitative impact on the presence of the US bias.[15]

The property of no market timing of the optimal portfolio discussed in section 2.4 allows us to construct unconditional optimal portfolios without having to estimate the latent state variable process – jump intensities in our case. All portfolios are constructed using static parameter estimates in Table 4.

Table 8 reports the optimal jump risk exposure under the four nested models for various coefficients of relative risk aversion and investment horizons. Observe that when jumps do not excite, the normalized exposure to jump risks does not change with investors' risk aversions or investment horizons. Compared to the prediction generated by the classic diffusion only model, neither the jump itself nor jump self excitation has a noticeable impact on the relative jump risk allocation of the optimal portfolio. Although Poisson jumps and self exciting jumps reduce or increase the *total* investment in risky assets, they barely affect the composition *within* the equity portfolio. The diffusion only, no excitation as well as the self excitation model predict comparable exposure to jump risks of the three regions.

When jumps are mutually exciting, investors demand more US jump exposure than the cases of self excitation or no excitation. As either risk aversion or investment horizon increases, the bias towards US

[15] See Appendix F1 for robustness checks on this matter.

jump factors becomes more prominent. Self exciting jumps, which in fact have a symmetric excitation structure, are not able to generate the US bias. It is only when an asymmetric excitation matrix $\boldsymbol{\beta}$ comes into play do we observe a shift from Japanese jump factors to US jump factors. For example, when $\gamma = 5, T = 10$ (bold cells), of the total jump exposure of the portfolio, around 58% comes from US, and only 14% from Japan. Therefore we conclude that the serial dependence of jumps alone does not lead to the US bias, but rather the excitation asymmetry contributes to the over-weighting of the US equity market in the optimal portfolio.

It becomes clear that the over-exposure to the US Brownian risk is due to the large variance premium in the US (as shown in Table 7), while the over-exposure to the US jump risk is due to the asymmetric excitation structure. If the over-exposure to the US jump factor were, too, caused by a higher jump premium in the US, then we would observe the US bias in the portfolio predictions of the other models as well.

Model implied exposure to jump risks

Risk aversion	Horizon	Diffusion only	PJD	SEJD	MEJD
1.5	month	0.37	0.37	0.37	0.40
		0.30	0.30	0.29	0.28
		0.33	0.33	0.33	0.32
	quarter	0.37	0.37	0.37	0.42
		0.30	0.30	0.28	0.26
		0.33	0.33	0.34	0.32
	10-year	0.37	0.37	0.37	0.44
		0.30	0.30	0.27	0.25
		0.33	0.33	0.36	0.31
3	month	0.37	0.37	0.37	0.43
		0.30	0.30	0.29	0.26
		0.33	0.33	0.34	0.31
	quarter	0.37	0.37	0.38	0.49
		0.30	0.30	0.26	0.21
		0.33	0.33	0.36	0.30
	10-year	0.37	0.37	0.36	0.53
		0.30	0.30	0.21	0.18
		0.33	0.33	0.43	0.29
5	month	0.37	0.37	0.37	0.44
		0.30	0.30	0.28	0.25
		0.33	0.33	0.34	0.31
	quarter	0.37	0.37	0.38	0.52
		0.30	0.30	0.25	0.19
		0.33	0.33	0.37	0.29
	10-year	0.37	0.37	0.34	**0.58**
		0.30	0.30	0.17	**0.14**
		0.33	0.33	0.49	**0.28**

Table 8: Optimal exposure to jump risks in a jump diffusion market where investors specify the return generating process to be pure diffusion ("Diffusion only"), Poisson jump diffusion ("PJD"), self exciting jump diffusion ("SEJD") and mutually exciting jump diffusion ("MEJD"). Each column is in the order of "US, Japan, Europe". Parameter values are contained in Table 4 and 7. The figures are normalized so that the exposure to the three regions adds up to 1.

6 Conclusion

Inspired by the empirical findings that investors tend to over-invest in US equities compared to classic model implications and the finding that the US equity plays a leading role in international stock returns, we postulate a mutually exciting jump diffusion model for equity prices that (1) accounts for the lead-lag relationships of international returns; (2) theoretically generates the US bias in a representative investor's equity portfolio.

In particular, we allow for asymmetric jump excitation in international equity prices. We show that the leading role of the US equity can be generated by having larger cross section excitor(s) as the source jump component than the cross section excitor(s) as the target jump component. We solve the asset allocation problem in closed form in this market using the martingale approach and apply the theoretical work to historical returns on MSCI indices. We show that the optimal portfolio exhibits the US bias, i.e., the US equity is over-weighted in the market portfolio compared to the classic portfolio predictions.

The analytical nature of the solution helps establish the economic intuitions of the optimal portfolio risk exposure, which can be summarized into the following properties: (1) The optimal portfolio is sufficiently diversified in the sense that it includes a large number of individual stocks to diversify away idiosyncratic risks; (2) Similar to the Merton mean-variance portfolio, it exploits the covariance structure of the Brownian risks; (3) The exposure to jump risks includes exposure to Poisson jump risk, which is a risk-return tradeoff term, and exposure to contagion risks, which increases the total jump risk exposure since jump factors can be used to hedge against changes in the state variables; (4) The portfolio exposure to jump risks is biased towards the US equity which is capable of spreading jump risks worldwide and is not prone to foreign equity markets turmoil. Using parameter estimates on MSCI indices, we show that the US bias observed in the market portfolio can be explained by excitation asymmetry.

Appendix

A Proofs

Proof of Proposition 1. Since the market is free of arbitrage opportunities and admits a unique martingale measure Q given by Equation (13), the portfolio optimization problem defined in (1) is equivalent to

$$\sup_{X_T \in \mathcal{X}} E_0\Big[\frac{X_T^{1-\gamma}}{1-\gamma}\Big],$$

where $\mathcal{X}$ is the set of admissible square integrable terminal wealth:

$$\mathcal{X} = \{X_T \text{ is } \mathcal{F}_T - \text{measurable} : E_0[\pi_T X_T] \leq e^{rT} x_0\}.$$

The corresponding Lagrange function is constructed as

$$\mathcal{L} = \frac{X_T^{1-\gamma}}{1-\gamma} + y(e^{rT} x_0 - \pi_T X_T).$$

Take the first order condition with respect to X_T. The optimal terminal wealth is given by

$$X_T^* = (y\pi_T)^{-\frac{1}{\gamma}},$$

where y is the Lagrange multiplier that satisfies

$$E_0[\pi_T X_T^*] = e^{rT} x_0,$$

which implies

$$y^{-\frac{1}{\gamma}} = \frac{e^{rT} x_0}{E_0[\pi_T^{1-1/\gamma}]}.$$

Since the wealth process satisfies the no arbitrage assumption

$$X_t \pi_t = E_t[X_T \pi_T],$$

we have

$$X_t^* = \frac{E_t[X_T^* \pi_T]}{\pi_t} = \frac{E_t[y^{-\frac{1}{\gamma}} \pi_T^{1-\frac{1}{\gamma}}]}{\pi_t} = e^{rt} \frac{x_0}{\pi_t} \frac{E_t\Big[\pi_T^{\frac{\gamma-1}{\gamma}}\Big]}{E_0\Big[\pi_T^{\frac{\gamma-1}{\gamma}}\Big]}.$$

The solution requires the computation of the expected value of the exponential of a stochastic integral. Let $M_t := \log(\pi_t)$. It holds that

$$
\begin{aligned}
dM_t =& \frac{d\pi_t}{\pi_t} - \frac{1}{2}\frac{d\langle\pi_t\rangle}{\pi_t^2} \\
=& \Big(-\frac{1}{2}\boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta} - \sum_{i=1}^{n}\kappa_i\lambda_{i,t}\Big)dt - \boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}d\boldsymbol{W}_t + \sum_{i=1}^{n}\log(1+\kappa_i)dN_{i,t}.
\end{aligned}
$$

Write $\boldsymbol{Y}_t = (M_t; \boldsymbol{\lambda}_t)$. We have

$$
d\boldsymbol{Y}_t = \boldsymbol{\mu}(\boldsymbol{Y}_t)dt + \boldsymbol{\sigma}(\boldsymbol{Y}_t)d\boldsymbol{W}_t + d\boldsymbol{N}_t.
$$

One can easily check that $\boldsymbol{Y}_t$ is well-behaved in the sense of Duffie et al. (2000), for finite T. In addition, $\boldsymbol{Y}_t$ admits the affine structure:

$$
\begin{cases}
\boldsymbol{\mu}(\boldsymbol{Y}_t) = \boldsymbol{K}_0 + \boldsymbol{K}_1\boldsymbol{Y}_t, \\
[\boldsymbol{\sigma}(\boldsymbol{Y}_t)\boldsymbol{\sigma}(\boldsymbol{Y}_t)']_{ij} = [\boldsymbol{H}_0]_{ij} + [\boldsymbol{H}_1]_{ij}\cdot\boldsymbol{Y}_t, \\
\lambda^i(\boldsymbol{Y}_t) = l_0^i + \boldsymbol{l}_1^i\cdot\boldsymbol{Y}_t,
\end{cases}
$$

where

$$
\boldsymbol{K}_0 = \begin{pmatrix} -\frac{1}{2}\boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta} \\ \alpha\lambda_{1,\infty} \\ \vdots \\ \alpha\lambda_{n,\infty} \end{pmatrix}, \ \boldsymbol{K}_1 = \begin{pmatrix} 0 & -\kappa_1 & \dots & -\kappa_n \\ 0 & -\alpha & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -\alpha \end{pmatrix},
$$

$$
\boldsymbol{H}_0 = \begin{pmatrix} \boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \ \boldsymbol{H}_1 = 0,
$$

$$
l_0^i = 0, \ \boldsymbol{l}_1^i = (0; \boldsymbol{e}_i), i = 1, \dots, n.
$$

Here, $\boldsymbol{e}_i$ denotes a vector of zeros with 1 at the i^{th} entry. The jump transform $\zeta^i(c)$ that determines the jump size distribution of jump component i is $\zeta^i(c) = (1+\kappa_i)^{c_1}\exp(\sum_{j=1}^{n} c_{j+1}\beta_{ji}), i = 1, \dots, n$. According to Duffie et al. (2000), the conditional expectation takes the form

$$
E_t\Big[\pi_T^{\frac{\gamma-1}{\gamma}}\Big] = \pi_t^{\frac{\gamma-1}{\gamma}}\exp(A(t) + \boldsymbol{B}'(t)\boldsymbol{\lambda}_t),
$$

where $A(t)$ and $\boldsymbol{B}(t)$ satisfy

$$\begin{cases} \dot{\boldsymbol{B}}(t) = \frac{\gamma-1}{\gamma}\boldsymbol{\kappa} + \alpha \boldsymbol{B}(t) - (\boldsymbol{\kappa}+1)^{\frac{\gamma-1}{\gamma}} \circ e^{\boldsymbol{\beta}'\boldsymbol{B}(t)} + 1, \\ \dot{A}(t) = \frac{\gamma-1}{2\gamma}\boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta} - \alpha \boldsymbol{B}'(t)\boldsymbol{\lambda}_\infty, \end{cases}$$

with $A(T) = 0, \boldsymbol{B}(T) = \boldsymbol{0}$. Therefore it holds that

$$X_t^* = e^{rt}x_0\pi_t^{-1/\gamma}\exp(A(t) + \boldsymbol{B}'(t)\boldsymbol{\lambda}_t - A(0) - \boldsymbol{B}'(0)\boldsymbol{\lambda}_0),$$

from which we derive the SDE of the optimal wealth path

$$\begin{aligned} \frac{dX_t^*}{X_{t^-}^*} =& \Big(r + \frac{1}{\gamma}\boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta}\Big)dt + \frac{1}{\gamma}\boldsymbol{\eta}'(\boldsymbol{L}\boldsymbol{L}')^{-1}d\boldsymbol{W}_t \\ &+ \sum_{l=1}^{n}\Big((1+\kappa_l)^{-\frac{1}{\gamma}}\exp(\boldsymbol{\beta}_l'\boldsymbol{B}(t)) - 1\Big)(dN_{l,t} - (1+\kappa_l)\lambda_{l,t}dt) \\ =& \boldsymbol{\theta}^{W*\prime}(d\boldsymbol{W}_t + \boldsymbol{\eta}dt) + \sum_{l=1}^{n}\Big(\exp\big(\theta_t^{N_l*}\big) - 1\Big)\big(dN_{l,t} - (1+\kappa_l)\lambda_{l,t}dt\big). \end{aligned}$$

□

Proof of Lemma 1. The economy has constant investment opportunity with asset prices generated by geometric Brownian motions. Let $\boldsymbol{\sigma}$ be a diagonal matrix with σ_i on the diagonals. We can write the asset price dynamics in matrix form as

$$\operatorname{diag}(\boldsymbol{S}_t)^{-1}d\boldsymbol{S}_t = (r + \boldsymbol{\sigma}\boldsymbol{\eta})dt + \boldsymbol{\sigma}d\boldsymbol{W}_t.$$

Then according to Merton (1969), the portfolio weights are given by

$$\boldsymbol{w}^* = \frac{1}{\gamma}(\boldsymbol{\sigma}\boldsymbol{L}\boldsymbol{L}'\boldsymbol{\sigma}')^{-1}\boldsymbol{\sigma}\boldsymbol{\eta}.$$

The portfolio wealth process follows

$$\frac{dX_t^*}{X_t^*} = (r + \boldsymbol{w}^{*\prime}\boldsymbol{\sigma}\boldsymbol{\eta})dt + \boldsymbol{w}^{*\prime}\boldsymbol{\sigma}d\boldsymbol{W}_t.$$

It holds that

$$\boldsymbol{\theta}_{\text{Merton}}^{W*} = \boldsymbol{\sigma}'\boldsymbol{w}^* = \frac{1}{\gamma}(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta} = \boldsymbol{\theta}^{W*}.$$

□

Proof of Lemma 2. To follow the price dynamics of $F_t \in \mathcal{S}^2$ given by Equation (22), it requires that for every $i = 1, \dots, n$,

$$\begin{cases} w_{i,t}^{k=1}\sigma_i^{k=1} + w_{i,t}^{k=2}\sigma_i^{k=2} = g_i, \\ w_{i,t}^{k=1}z_i^{k=1} + w_{i,t}^{k=2}z_i^{k=2} = h_i. \end{cases} \tag{37}$$

Hence at time t, for every region i, we have two unknowns and two equations that are not linearly dependent. We can solve for the weighting vector at any time t and replicate the price dynamic of F_t in continuous time. Since $\{P_t\}$ and $\{F_t\}$ are solutions of the same stochastic differential equations, they are indistinguishable processes. □

Proof of Proposition 2. Inspired by Merton (1980), the portfolio weights $w_{i,t}^k$ can be restricted such that they satisfy

$$\begin{cases} \sum_{k=1}^{m_i} w_{i,t}^k\sigma_i^k = g_i, \\ \sum_{k=1}^{m_i} w_{i,t}^k z_i^k = h_i, \end{cases} \tag{38}$$

and can be represented as

$$w_{i,t}^k =: \frac{\omega_{i,t}^k}{m_i}, \tag{39}$$

where $\omega_{i,t}^k$ is a finite constant, independent of the total number of assets.[16] The replicating portfolio P_t thus follows

$$\begin{aligned} \frac{dP_t}{P_{t^-}} =& \sum_{i=1}^n g_i(dW_{i,t} + \eta_i dt) + \sum_{i=1}^n h_i(dN_{i,t} - (1+\kappa_i)\lambda_{i,t}dt) \\ &+ \sum_{i=1}^n \sum_{k=1}^{m_i} w_{i,t}^k \nu_i^k dZ_{i,t}^k \\ =& \frac{dF_t}{F_{t^-}} + \sum_{i=1}^n d\zeta_{i,t}(m_i), \end{aligned}$$

where

$$d\zeta_{i,t}(m_i) := \sum_k^{m_i} w_{i,t}^k \nu_i^k dZ_{i,t}^k.$$

Let $d\varepsilon_{i,t}^k = \omega_{i,t}^k \nu_i^k dZ_{i,t}^k$. Then $d\varepsilon_{i,t}^k$ is a random variable following normal distribution with zero mean and variance $(\omega_{i,t}^k \nu_i^k)^2 dt$. The variance of $d\varepsilon_{i,t}^k$ is bounded and independent of the total number of assets m_i by the assumption on $\omega_{i,t}^k$. We have that

$$d\zeta_{i,t}(m_i) = \sum_k^{m_i} w_{i,t}^k \nu_i^k dZ_{i,t}^k = \frac{\sum_{k=1}^{m_i} d\varepsilon_{i,t}^k}{m_i}.$$

[16]Such representation is indeed possible due to Lemma 2. Appendix B gives an explicit example of such representation in case of the optimal wealth path.

Note that $d\varepsilon_{i,t}^k$ are independent, since each $dZ_{i,t}^k$ represents the idiosyncratic risk of asset k in region i. By the Law of Large Numbers, it holds that, for all i,

$$d\zeta_{i,t}(m_i) \xrightarrow{P} 0,$$

which implies that

$$\frac{dP_t}{P_{t-}} \xrightarrow{P} \frac{dF_t}{F_{t-}}.$$

Therefore

$$P_t \xrightarrow{P} F_t, \forall t.$$

□

Proof of Proposition 3. For a given well-diversified portfolio (i.e., free of idiosyncratic risks), suppose its exposure to the systematic Brownian risks is $\hat{\boldsymbol{\theta}}_t^W$, and its exposure to the jump risks is $\hat{\boldsymbol{\theta}}_t^N$. The wealth process is given by

$$d\log \hat{X}_t = \Big(-\frac{1}{2}\hat{\boldsymbol{\theta}}_t^{W\prime} LL' \hat{\boldsymbol{\theta}}_t^W + \hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{\eta} - \sum_{l=1}^{n} \big(e^{\hat{\boldsymbol{\theta}}_t^{N_l}} - 1\big)(1+\kappa_i)\lambda_{l,t}\Big) dt + \hat{\boldsymbol{\theta}}_t^{W\prime} d\boldsymbol{W}_t + \sum_{l=1}^{n} \hat{\boldsymbol{\theta}}_t^{N_l} dN_{i,t},$$

with $X_0 = x_0$. Similar to the proof of Proposition 1, we employ the formula in Duffie, Pan, and Singleton (2000) to evaluate the expected utility of terminal wealth and write $\boldsymbol{Y}_t = (\log \hat{X}_t, \boldsymbol{\lambda}_t)$. We have

$$d\boldsymbol{Y}_t = \boldsymbol{\mu}(\boldsymbol{Y}_t)dt + \boldsymbol{\sigma}(\boldsymbol{Y}_t)d\boldsymbol{W}_t + d\boldsymbol{N}_t.$$

$\boldsymbol{Y}_t$ admits the affine structure:

$$\begin{cases} \boldsymbol{\mu}(\boldsymbol{Y}_t) = \boldsymbol{K}_0 + \boldsymbol{K}_1 \boldsymbol{Y}_t, \\ [\boldsymbol{\sigma}(\boldsymbol{Y}_t)\boldsymbol{\sigma}(\boldsymbol{Y}_t)']_{ij} = [\boldsymbol{H}_0]_{ij} + [\boldsymbol{H}_1]_{ij} \cdot \boldsymbol{Y}_t, \\ \lambda^i(\boldsymbol{Y}_t) = l_0^i + \boldsymbol{l}_1^i \cdot \boldsymbol{Y}_t, \end{cases}$$

where

$$
\boldsymbol{K}_0 = \begin{pmatrix} -\frac{1}{2}\hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{L}\boldsymbol{L}'\hat{\boldsymbol{\theta}}_t^{W} + \hat{\boldsymbol{\theta}}^{W\prime}\boldsymbol{\eta} \\ \alpha\lambda_{1,\infty} \\ \vdots \\ \alpha\lambda_{n,\infty} \end{pmatrix},
$$

$$
\boldsymbol{K}_1 = \begin{pmatrix} 0 & -(1+\kappa_1)\big(e^{\hat{\theta}_t^{N_1}}-1\big) & \dots & -(1+\kappa_n)\big(e^{\hat{\theta}_t^{N_n}}-1\big) \\ 0 & -\alpha & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -\alpha \end{pmatrix},
$$

$$
\boldsymbol{H}_0 = \begin{pmatrix} \hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{L}\boldsymbol{L}'\hat{\boldsymbol{\theta}}_t^{W} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \ \boldsymbol{H}_1 = 0,
$$

$$
l_0^i = 0, \ \boldsymbol{l}_1^i = (0;\boldsymbol{e}_i), \ i = 1\dots,n.
$$

The jump transform $\zeta^i(c)$ is given by $\zeta^i(c) = \exp\big(c_1\hat{\theta}_t^{N_i} + \sum_{j=1}^n c_{j+1}\beta_{ji}\big), i = 1,\dots,n$. The conditional expectation takes the form

$$
\frac{1}{1-\gamma}E[\hat{X}_T^{1-\gamma}] = \frac{(e^{rT}x_0)^{1-\gamma}}{1-\gamma}\exp(\hat{A}(0) + \hat{\boldsymbol{B}}'(0)\boldsymbol{\lambda}_0), \tag{40}
$$

where

$$
\begin{aligned}
\dot{\hat{\boldsymbol{B}}}(t) =& (1-\gamma)(\boldsymbol{\kappa}+1)\circ\big(e^{\hat{\boldsymbol{\theta}}_t^N}-1\big) + \alpha\hat{\boldsymbol{B}}(t) - \exp\big((1-\gamma)\hat{\boldsymbol{\theta}}_t^N + \boldsymbol{\beta}'\hat{\boldsymbol{B}}\big) + 1, \\
\dot{\hat{A}}(t) =& -(1-\gamma)\Big(\hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{\eta} - \frac{1}{2}\hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{L}\boldsymbol{L}'\hat{\boldsymbol{\theta}}_t^{W}\Big) - \frac{1}{2}(1-\gamma)^2\hat{\boldsymbol{\theta}}_t^{W\prime}\boldsymbol{L}\boldsymbol{L}'\hat{\boldsymbol{\theta}}_t^{W} - \alpha\hat{\boldsymbol{B}}'\boldsymbol{\lambda}_\infty,
\end{aligned}
$$

with $\hat{\boldsymbol{B}}(T) = \boldsymbol{0}, \hat{A}(T) = 0$. The certainty-equivalent wealth is

$$
\hat{x}_0 = x_0 e^{rT}\exp\Big(\frac{1}{1-\gamma}(\hat{A}(0) + \hat{\boldsymbol{B}}'(0)\boldsymbol{\lambda}_0)\Big).
$$

□

Proof of Corollary **??**. If we restrict $\boldsymbol{\beta} = 0$, to match the patterns of the mutually exciting jumps, we need to impose that

$$
\boldsymbol{\lambda} = E[\boldsymbol{\lambda}_t],
$$

$$\alpha = 0,\ \boldsymbol{\lambda}_\infty = \boldsymbol{\lambda}.$$

Since the optimal jump exposure $\boldsymbol{\theta}_t^{N*}$ is only a function of $\boldsymbol{\beta}$, independent of $\alpha, \boldsymbol{\lambda}_\infty$, the jump exposure restricting $\boldsymbol{\beta} = 0$ is simply the Poisson exposure:

$$\hat{\boldsymbol{\theta}}_t^N = -\frac{1}{\gamma}\log(1+\boldsymbol{\kappa}).$$

The Brownian exposure remains unchanged:

$$\hat{\boldsymbol{\theta}}_t^W = \boldsymbol{\theta}^{W*}.$$

Applying Proposition 3 using $\hat{\boldsymbol{\theta}}_t^W = \boldsymbol{\theta}^{W*}, \hat{\boldsymbol{\theta}}_t^N = -\frac{1}{\gamma}\log(1+\boldsymbol{\kappa})$, one gets the desired results. □

B Portfolio construction with a large basket of assets

In this section we demonstrate how to represent the portfolio weights by Equation (39) in a given region, thereby completing the proof of Proposition 2.

Let $(\sigma^k, \nu^k, z^k), k = 1, \ldots, m$, characterize the price dynamics of individual asset k in a given region (for which we omit the subscript denoting the region identify). First, we randomly pair the m assets into $m/2$ pairs, denoted by $p, p = 1, \ldots, m/2$.[17] For any pair p with asset S^k, S^l, let ω_t^k, ω_t^l be the weights on assets S^k, S^l as if S^k, S^l make up the entire portfolio in that region. Since the optimal regional portfolio produces the risk exposure f, g, it holds that

$$\begin{cases} \omega_t^k \sigma^k + \omega_t^l \sigma^l = f, \\ \omega_t^k z^k + \omega_t^l z^l = g. \end{cases}$$

Since assets are not linearly dependent, the linear equation system has a unique solution for ω_t^k, ω_t^l. Define the paired portfolio P_t^p as

$$dP_t^p := \omega_t^k dS_t^k + \omega_t^l dS_t^l.$$

It holds that the price of any paired portfolio X_t^p follows the dynamic of the local optimal portfolio plus some tracking errors:

$$\frac{dP_t^p}{P_{t^-}^p} = f(dW_t + \eta dt) + g(dN_t - (1+\kappa)\lambda_t dt) + d\zeta_t^p,$$

[17]In case m is odd, we simply create a $m+1$ asset by, for example, $S^{m+1} = \frac{1}{2}S^m + \frac{1}{2}S^{m-1}$. Therefore we assume m is even without loss of generality.

where

$$d\zeta_t^p := \omega_t^k \nu^k dZ_t^k + \omega_t^k \nu^k dZ_t^l =: \nu_t^p dZ_t^p.$$

Here, dZ_t^p is the idiosyncratic Brownian motion of the paired portfolio, independent of all other paired portfolios. The last equality of the previous equation holds in distribution.

Now that we have $m/2$ paired portfolios. Any weighted average of these paired portfolios produces the optimal exposure to systematic risk factors. Denote the weights on the paired portfolios by $\boldsymbol{a}(m) = (a^1, \ldots, a^{m/2})'$. The replicating portfolio $P_t(m)$is given by

$$dP_t(m) = \boldsymbol{a}(m)' dP_t^p.$$

For example, in an equal weighted scheme, each paired portfolio is assigned the same weight $a^p = 2/m, p = 1, \ldots, m/2$. Then asset k in the m-asset portfolio gets weight

$$w_t^k = \frac{2\omega_t^k}{m},$$

in which $2\omega_t^k$ is independent of the number of asset m. We have therefore showed that the representation proposed in (39) is indeed feasible.

C Portfolios that exhibit home bias

The mutually exciting jump diffusion model proposed in Section 2.3 allows for ambiguity averse preferences, modeled by, for example, the multiple prior preferences proposed by Gilboa and Schmeidler (1989). An investor from region i does not have the full knowledge of the true probability law of asset prices. Instead of optimizing the expected utility under the physical measure, the investor specifies plausible candidate ambiguity measures $G_i \in \mathcal{G}$ and optimizes the expected utility under the worst case scenario:

$$\sup_{\boldsymbol{\theta}^W, \boldsymbol{\theta}_t^N} \inf_{G_i \in \mathcal{G}} E_0^G[u(X_T)]. \tag{41}$$

In general, Investors from different regions have different ambiguity levels towards other regions. Equation (41) describes the optimization problem for an investor from region i with ambiguity measure $\mathcal{G}$. We omit the region identity i for notation simplicity.

In order to have a tractable solution, we further assume that ambiguity comes from parameter uncertainty. Instead of relying entirely on the point estimates of risk premiums, investors construct confidence intervals of parameter estimates, and optimize under the worst case parameter values. Based on the pure diffusion model in Garlappi, Uppal, and Wang (2007), we restrict the Brownian risk premium $\boldsymbol{\eta}$ to lie within $[\underline{\boldsymbol{\eta}}, \overline{\boldsymbol{\eta}}]$, and similarly, the jump risk premium parameter $\boldsymbol{\kappa}$ to lie within $[\underline{\boldsymbol{\kappa}}, \overline{\boldsymbol{\kappa}}]$. Then we may characterize the ambiguity measure G by $\boldsymbol{\eta}^G, \boldsymbol{\kappa}^G$ and write $G(\boldsymbol{\eta}^G, \boldsymbol{\kappa}^G)$, where

$$\boldsymbol{\eta}^G \in [\underline{\boldsymbol{\eta}}, \overline{\boldsymbol{\eta}}], \quad 0 \leq \underline{\boldsymbol{\eta}} \leq \boldsymbol{\eta} \leq \overline{\boldsymbol{\eta}}, \quad [\boldsymbol{\eta}^G]_i = \eta_i, \tag{42}$$

$$\boldsymbol{\kappa}^G \in [\underline{\boldsymbol{\kappa}}, \overline{\boldsymbol{\kappa}}], \quad 0 \leq \underline{\boldsymbol{\kappa}} \leq \boldsymbol{\kappa} \leq \overline{\boldsymbol{\kappa}}, \quad [\boldsymbol{\kappa}^G]_i = \kappa_i. \tag{43}$$

All operators are element-wise comparisons. The constraints (42) and (43) specify the ambiguity level – the larger the sets $[\underline{\boldsymbol{\eta}}, \overline{\boldsymbol{\eta}}], [\underline{\boldsymbol{\kappa}}, \overline{\boldsymbol{\kappa}}]$, the more ambiguity averse the investor is. Nevertheless, however large the investor's ambiguity aversion is, he/she has no ambiguity towards the home equity, which is reflected through $[\boldsymbol{\eta}^G]_i = \eta_i, [\boldsymbol{\kappa}^G]_i = \kappa_i$. Since $\boldsymbol{\eta} \in [\underline{\boldsymbol{\eta}}, \overline{\boldsymbol{\eta}}], \boldsymbol{\kappa} \in [\underline{\boldsymbol{\kappa}}, \overline{\boldsymbol{\kappa}}]$, the ambiguity sets contain the true measure P. Notice that the ambiguity set $\mathcal{G}$ is convex and compact.

Under the ambiguity measure $G(\boldsymbol{\eta}^G, \boldsymbol{\kappa}^G)$, asset k from region i follows the dynamics

$$\frac{dS_{i,t}^k}{S_{i,t^-}^k} = \sigma_i^k (dW_{i,t}^G + \eta_i^G dt) + \nu_i^k dZ_{i,t}^{k,G} + z_i (dN_{i,t}^G - (1 + \kappa_i^G)\lambda_{i,t} dt),$$

where $\boldsymbol{W}_t^G, \boldsymbol{Z}_t^G$ are martingales and $\boldsymbol{N}_t^G$ are mutually exciting jumps with intensities $\big(1 + (\boldsymbol{\kappa} - \boldsymbol{\kappa}^G)\big) \circ \boldsymbol{\lambda}_t$ under the ambiguity measure G. Consistent with the asset dynamics under G is the measure change process $\frac{dG}{dP} \mid_{\mathcal{F}_t} = \xi_t$ that follows

$$\frac{d\xi_t}{\xi_{t^-}} = -(\boldsymbol{\eta} - \boldsymbol{\eta}^G)'(\boldsymbol{L}\boldsymbol{L}')^{-1} d\boldsymbol{W}_t + \sum_{i=1}^n \Big(\kappa_i - \kappa_i^G\Big)\Big(dN_{i,t} - \lambda_{i,t} dt\Big), \quad \xi_0 = 1. \tag{44}$$

Expected utility is nested when the ambiguity set $\mathcal{G}$ is a singleton in which the physical measure P is the only element. The optimal risk exposure $\boldsymbol{\theta}^{W*}, \boldsymbol{\theta}_t^{N*}$ derived in Proposition 1 can therefore be regarded as a special case of the more general function $\boldsymbol{\theta}^{W*}(\boldsymbol{\eta}^G), \boldsymbol{\theta}_t^{N*}(\boldsymbol{\kappa}^G)$, when $\boldsymbol{\eta}^G = \boldsymbol{\eta}$, $\boldsymbol{\kappa}^G = \boldsymbol{\kappa}$. The following proposition is a generalization of Proposition 1 to incorporate ambiguity averse preferences.

Proposition 4 (Optimal portfolio choice with ambiguity aversion)**.** *In a contagious economy with asset prices given by* (2.3)*, suppose that a representative investor from a certain region is ambiguity averse and aims to solve*

$$\sup_{\boldsymbol{\theta}^W, \boldsymbol{\theta}_t^N} \inf_{G \in \mathcal{G}} E^G \Big[\frac{X_T^{1-\gamma}}{1-\gamma} | X_0 = x_0\Big], \gamma > 1, \tag{45}$$

with

$$\mathcal{G} = \Big\{ G(\boldsymbol{\eta}^G, \boldsymbol{\kappa}^G) : \frac{dG}{dP}\Big|_{\mathcal{F}_t} = \xi_t \Big\},$$

where ξ_t is given by (44). Then the optimal portfolio X^* follows the dynamics

$$\frac{dX_t^*}{X_{t-}^*} = \boldsymbol{\theta}^{W*}(\underline{\boldsymbol{\eta}})'(d\boldsymbol{W}_t + \boldsymbol{\eta}dt) + \sum_{i=1}^{n}\left(\exp(\theta_t^{N_i*}(\underline{\boldsymbol{\kappa}})) - 1\right)\Big(dN_{i,t} - (1+\kappa_i)\lambda_{i,t}dt\Big),$$

where the risk exposure is given by

$$\begin{cases} \boldsymbol{\theta}^{W*}(\underline{\boldsymbol{\eta}}) = \frac{1}{\gamma}(\boldsymbol{L}\boldsymbol{L}')^{-1}\underline{\boldsymbol{\eta}}, \\ \boldsymbol{\theta}_t^{N*}(\underline{\boldsymbol{\kappa}}) = -\frac{1}{\gamma}\log(1+\underline{\boldsymbol{\kappa}}) + \boldsymbol{\beta}'\boldsymbol{B}(\underline{\boldsymbol{\kappa}};t). \end{cases} \tag{46}$$

Here, $\boldsymbol{B}(\underline{\boldsymbol{\kappa}};t)$ is given by

$$\dot{\boldsymbol{B}}(\underline{\boldsymbol{\kappa}};t) = \frac{\gamma-1}{\gamma}\underline{\boldsymbol{\kappa}} + \alpha\boldsymbol{B}(\underline{\boldsymbol{\kappa}};t) - (\underline{\boldsymbol{\kappa}}+1)^{\frac{\gamma-1}{\gamma}} \circ e^{\boldsymbol{\beta}'\boldsymbol{B}(\underline{\boldsymbol{\kappa}};t)} + 1, \tag{47}$$

with $\boldsymbol{B}(\underline{\boldsymbol{\kappa}};T) = 0$.

Proof of Proposition 4. Since any prior $G \in \mathcal{G}$ is equivalent to P, and $\mathcal{G}$ is by construction convex and compact, it holds that[18]

$$\sup_{\boldsymbol{\theta}^W,\boldsymbol{\theta}_t^N}\inf_{G\in\mathcal{G}} E^G\Big[\frac{X_T^{1-\gamma}}{1-\gamma}|X_0 = x_0\Big] = \inf_{G\in\mathcal{G}}\sup_{\boldsymbol{\theta}^W,\boldsymbol{\theta}_t^N} E^G\Big[\frac{X_T^{1-\gamma}}{1-\gamma}|X_0 = x_0\Big].$$

For a given measure, we first solve the inner supremum problem. The result in Proposition 1 can be directly applied. For any $G(\boldsymbol{\eta}^G,\boldsymbol{\kappa}^G) \in \mathcal{G}$, the optimal portfolio exposure to risk factors is given by

$$\begin{cases} \boldsymbol{\theta}^{W*}(\boldsymbol{\eta}^G) = \frac{1}{\gamma}(\boldsymbol{L}\boldsymbol{L}')^{-1}\boldsymbol{\eta}^G, \\ \boldsymbol{\theta}_t^{N*}(\boldsymbol{\kappa}^G) = -\frac{1}{\gamma}\log(1+\boldsymbol{\kappa}^G) + \boldsymbol{\beta}'\boldsymbol{B}(\boldsymbol{\kappa}^G;t). \end{cases}$$

where

$$\dot{\boldsymbol{B}}(\boldsymbol{\kappa}^G;t) = \frac{\gamma-1}{\gamma}\boldsymbol{\kappa}^G + \alpha\boldsymbol{B}(\boldsymbol{\kappa}^G;t) - (\boldsymbol{\kappa}^G+1)^{\frac{\gamma-1}{\gamma}} \circ e^{\boldsymbol{\beta}'\boldsymbol{B}(\boldsymbol{\kappa}^G;t)} + 1,$$

with $\boldsymbol{B}(\boldsymbol{\kappa}^G;T) = 0$. Having solved the inner supremum problem, one can easily show that the indirect utility function given in (18) is strictly decreasing in $\boldsymbol{\eta},\boldsymbol{\kappa}$ for $\boldsymbol{\eta} \geq 0, \boldsymbol{\kappa} \geq 0$. Therefore it suffices to replace $\boldsymbol{\eta}^G$ by $\underline{\boldsymbol{\eta}}$, and $\boldsymbol{\kappa}^G$ by $\underline{\boldsymbol{\kappa}}$ in $\boldsymbol{\theta}^{W*}(\boldsymbol{\eta}^G), \boldsymbol{\theta}_t^{N*}(\boldsymbol{\kappa}^G)$. □

The proposition confirms that the results of Garlappi et al. (2007), who find that ambiguity aversion towards expected return in a pure diffusion market is equivalent to a lower risk premium, can be readily

[18] The proof of this equality can be found in, for example, Theorem 2 of Schied and Wu (2005).

extended to our jump diffusion model. The coexistence of home bias and foreign bias found in investors' equity portfolios can therefore be generated by taking realistic values of the ambiguity parameters.

D Transition density of jump arrivals

We first give the algorithm of computing the log transition density function of jump arrivals for the univariate case.

Algorithm 1 (Univariate)**.** *The algorithm of computing the transition densities for a univariate self exciting jump process, given the K jump arrival times $\{u_1, \ldots, u_K\}$ within a time span $[0, T]$, conditional on $\lambda_0 = \lambda_\infty, N_0 = 0$:*

1. *Set the initial conditions: $u_0 = 0$, $\lambda_t = \lambda_\infty$, $t \in [0, u_1^-]$ and $k \in \{1, 2, \ldots, K\}$. Define $u_{K+1} := T$.*

2. *Denote the log likelihood of observing a jump occurrence at time u_k conditional on the information available by time u_{k-1} by $f(u_k|\mathcal{F}_{k-1})$. It holds that*

$$f(u_k|\mathcal{F}_{k-1}) = \log \lambda_{u_k^-} - \Lambda(k),$$

where

$$\Lambda(k) := -\frac{1}{\alpha}(\lambda_{u_{k-1}} - \lambda_\infty)(e^{-\alpha(u_k - u_{k-1})} - 1) + \lambda_\infty(u_k - u_{k-1}).$$

3. *Record the jump intensity at u_k to be*

$$\lambda_{u_k} = \lambda_{u_k^-} + \beta.$$

4. *Compute the intensity just before the next jump arrival u_{k+1}:*

$$\lambda_{u_{k+1}^-} = (\lambda_{u_k} - \lambda_\infty)e^{-\alpha(u_{k+1} - u_k)} + \lambda_\infty.$$

5. *Repeat step 2-4 until $k = K$.*

6. *The total log likelihood L is*

$$L = \sum_{1}^{K} f(u_k|\mathcal{F}_{k-1}) - \Lambda(K + 1).$$

Proof. Given the k^{th} jump arrival $u_k, k = 1, \ldots, K$, the intensity of the point process at $t \in [u_{k-1}, u_k]$ follows

$$d\lambda_t = \alpha(\lambda_\infty - \lambda_t)dt, \ u_{k-1} \leq t < u_k,$$

with

$$\lambda_{u_k} = \lambda_{u_k^-} + \beta.$$

Apparently, the differential equation admits the solution

$$\lambda_t = \begin{cases} (\lambda_{u_{k-1}} - \lambda_\infty) e^{-\alpha(t-u_{k-1})} + \lambda_\infty, & \text{if } u_{k-1} \le t < u_k, \\ (\lambda_{u_{k-1}} - \lambda_\infty) e^{-\alpha(t-u_{k-1})} + \lambda_\infty + \beta, & \text{if } t = u_k. \end{cases}$$

Conditional on the jump arrival u_{k-1}, until the next jump arrival, the point process is a time-inhomogeneous Poisson jump process with exponentially decaying intensities. Denote the probability of a jump occurrence at time u_k and no jump occurrences between time u_{k-1} and u_k^- by $P(u_k|\mathcal{F}_{u_{k-1}})$. It holds that

$$\begin{aligned} P(u_k|\mathcal{F}_{u_{k-1}}) &= P(\text{waiting time} = (u_k - u_{k-1})|\mathcal{F}_{u_{k-1}}) \\ &= \lambda_{u_k^-} \exp(-\int_{u_{k-1}}^{u_k^-} \lambda_s ds). \end{aligned}$$

Define

$$\Lambda(k) := \int_{u_{k-1}}^{u_k^-} \lambda_s ds.$$

It holds that

$$\begin{aligned} \Lambda(k) &= \int_{u_{k-1}}^{u_k} \Big((\lambda_{u_{k-1}} - \lambda_\infty) e^{-\alpha(t-u_{k-1})} + \lambda_\infty \Big) dt \\ &= -\frac{1}{\alpha} (\lambda_{u_{k-1}} - \lambda_\infty)(e^{-\alpha(u_k - u_{k-1})} - 1) + \lambda_\infty (u_k - u_{k-1}). \end{aligned}$$

When $k = K + 1$, the probability of no jump occurrences until T can be computed as

$$\begin{aligned} P(u_{K+1}|\mathcal{F}_{u_K}) &:= P(N_{u_{k+1}} - N_{u_k} = 0|\mathcal{F}_{u_k}) \\ &= P(\text{waiting time} > (T - u_K)|\mathcal{F}_{u_K}) \\ &= \exp(-\int_{u_K}^{T} \lambda_s ds) \\ &= \exp(-\Lambda(K+1)). \end{aligned}$$

□

The algorithm can be easily generalized to a multivariate setting.

Algorithm 2 (Multivariate)**.** *The algorithm of computing the transition densities for a D-dimensional multivariate mutually exciting jump process, given the K joint jump times $\{u_1, \ldots, u_K\}$ within a time* *span* $[0,T]$, *conditional on* $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}_\infty$, $\boldsymbol{N}_0 = \boldsymbol{0}$:

1. *Set the initial conditions:* $u_0 = 0$, $\boldsymbol{\lambda}_t = \boldsymbol{\lambda}_\infty$, $t \in [0, u_1^-]$, *and* $k \in \{1, \ldots, K\}$. *Define* $u_{K+1} := T$.

2. *Decide* u_k *belongs to which jump component. Denote the jump component by* d. *The log transition probability of* u_k *is*

$$f(u_k|\mathcal{F}_{k-1}) = \log \lambda_{d,u_k^-} - \sum_{j=1}^{D} \Lambda(k,j), \tag{48}$$

where

$$\Lambda(k,j) := -\frac{1}{\alpha}(\lambda_{j,u_{k-1}} - \lambda_{j,\infty})(e^{-\alpha(u_k - u_{k-1})} - 1) + \lambda_{j,\infty}(u_k - u_{k-1}). \tag{49}$$

3. *Record the jump intensity at* u_k *to be*

$$\boldsymbol{\lambda}_{u_k} = \boldsymbol{\lambda}_{u_k^-} + \boldsymbol{\beta}_d,$$

where $\boldsymbol{\beta}_d$ *is the* d^{th} *column of the excitation matrix* $\boldsymbol{\beta}$.

4. *Compute the intensities just before the next jump arrival* u_{k+1}: *for* $j = 1, \ldots, D$,

$$\lambda_{j,u_{k+1}^-} = (\lambda_{j,u_k} - \lambda_{j,\infty})e^{-\alpha(u_{k+1} - u_k)} + \lambda_{j,\infty}.$$

5. *Repeat step 2-4 until* $k = K$.

6. *The total log likelihood* L *is*

$$L = \sum_{1}^{K} f(u_k|\mathcal{F}_{k-1}) - \sum_{j=1}^{D} \Lambda(K+1, j). \tag{50}$$

E Small sample behavior

To examine the small sample behavior of the maximum likelihood estimators, we run 5,000 Monte Carlo simulation experiments. We set the data generating parameters (DGP) to be the MEJD parameter estimates given in Table 4. We generate 41 years of jump arrivals using the exact simulation algorithm proposed by Dassios and Zhao (2013). For each simulated sample, we estimate $\alpha, \boldsymbol{\beta}$ using the maximum likelihood, using the same starting values as in the empirical estimation. Table 9 reports the mean and the quartiles of estimates.

MEJD estimation

	α	β_{11}	β_{21}	β_{31}	β_{12}	β_{22}	β_{32}	β_{13}	β_{23}	β_{33}
DGP	29.35	13.07	9.04	24.87	0.00	6.96	2.60	5.60	2.52	7.35
Mean	29.68	12.29	9.15	25.37	0.34	6.59	2.72	5.49	2.70	6.83
25% quantile	27.05	9.95	6.97	21.79	0.00	5.02	1.45	3.92	1.32	4.95
Median	29.46	12.39	9.03	25.05	0.00	6.56	2.45	5.38	2.51	6.79
75% quantile	32.05	14.63	11.18	28.53	0.45	8.09	3.66	6.98	3.83	8.58

Table 9: Mean and quartiles of parameter estimates from 5,000 Monte Carlo experiments.

F Robustness checks

F.1 Risk premium calibration

All parameters which are used to produce Table 8 are estimated from the historical data except for the risk premium parameters $\boldsymbol{\eta}, \boldsymbol{\kappa}$. In this section, we vary the risk premium parameters, while keeping the sum of the variance and jump premium equal to the historical equity premium. We will see in Table 10 and 11 that varying the risk premiums does not have a qualitative impact on the US bias.

Table 10 reports the model predicted jump exposure for different combinations of the variance premium and the jump premium. The variance premium is restricted such that the predicted Brownian exposure $\boldsymbol{\theta}^{W*}$ coincide with the Brownian risk exposure of the market portfolio.

Optimal portfolio exposure to jump risks

Variance premium	Jump premium	Diffusion only	PJD	SEJD	MEJD
0.73	4.70	0.40	0.39	0.39	0.61
0.26	3.70	0.27	0.28	0.14	0.12
0.47	4.63	0.32	0.32	0.48	0.28
1.27	4.15	0.39	0.38	0.36	0.60
0.45	3.50	0.28	0.29	0.15	0.13
0.82	4.28	0.33	0.33	0.48	0.28
1.82	3.60	0.37	0.37	0.34	0.58
0.65	3.31	0.30	0.30	0.17	0.14
1.17	3.93	0.33	0.33	0.49	0.28
2.73	2.69	0.34	0.34	0.30	0.54
0.97	2.98	0.32	0.32	0.22	0.17
1.76	3.34	0.34	0.34	0.49	0.29
3.64	1.78	0.28	0.29	0.25	0.46
1.30	2.66	0.36	0.36	0.28	0.23
2.35	2.76	0.35	0.35	0.48	0.31

Table 10: Model predicted jump exposure for different combinations of the variance premium and the jump premium. The models under consideration are pure diffusion ("Diffusion only"), Poisson jump diffusion ("PJD"), self exciting jump diffusion ("SEJD") and mutually exciting jump diffusion ("MEJD"). Within each scenario, every column is in the order of "US, Japan, Europe". The "Variance premium" and "Jump premium" are reported in percentage per annum. The sum of the variance premium and the jump premium is equal to the equity premium, and is held fixed at historical levels. Variance premium is calibrated such that the model predicted Brownian exposure coincides with that of the market portfolio. Parameter values are contained in Table 4 with $\gamma = 5,\ T = 10$. The figures are normalized so that the exposure to the three regions adds up to 1.

Table 11 reports both the model predicted Brownian and jump exposure without imposing the restriction that the Brownian exposure $\boldsymbol{\theta}^{W*}$ is equal to the Brownian risk exposure of the market portfolio. Observe that given the variance premium, all four models predict the same optimal Brownian risk exposure $\boldsymbol{\theta}^{W*}$. We report the optimal risk exposure under equal jump risk premiums across all regions (the first scenario), under equal variance premium across all regions (the second scenario), and when the total equity premium in each region is divided equally into the variance premium and jump premium (the third scenario).

Optimal portfolio exposure to risk factors

Variance premium	Jump premium	$\boldsymbol{\theta}^{W*}$	Diffusion only	PJD	SEJD	MEJD
2.43	3.00	0.46	0.37	0.37	0.36	0.55
0.96	3.00	0.08	0.32	0.32	0.23	0.17
2.10	3.00	0.46	0.30	0.31	0.41	0.28
3.00	2.43	0.31	0.49	0.48	0.48	0.57
3.00	0.96	0.33	0.17	0.18	0.13	0.12
3.00	2.10	0.37	0.34	0.35	0.39	0.31
2.71	2.71	0.36	0.42	0.41	0.41	0.55
1.98	1.98	0.23	0.26	0.27	0.20	0.16
2.55	2.55	0.40	0.32	0.32	0.39	0.29

Table 11: Model predicted portfolio exposure to Brownian and jump risk factors for different combinations of the variance premium and the jump premium. The models under consideration are pure diffusion ("Diffusion only"), Poisson jump diffusion ("PJD"), self exciting jump diffusion ("SEJD") and mutually exciting jump diffusion ("MEJD"). Within each scenario, every column is in the order of "US, Japan, Europe". The "Variance premium" and "Jump premium" are reported in percentage per annum. The sum of the variance premium and the jump premium is equal to the equity premium, and is held fixed at historical levels. The first block reports the optimal risk exposure under equal jump risk premiums across all regions; the second block reports the optimal risk exposure under equal variance premium across all regions; the last block reports the optimal risk exposure when the total equity premium in each region is divided equally into the variance premium and jump premium. Parameter values are contained in Table 4 with $\gamma = 5,\ T = 10$. The figures are normalized so that the exposure to the three regions adds up to 1.

F.2 Time zone differences

The econometric estimation is conducted using daily data on international equity returns. To account for differences in market opening times, we re-estimate the excitation parameter estimates by lagging the US returns by one day. Table 12 reports the resulting excitation parameters of the SEJD and MEJD models. Observe that the SEJD parameters are not affected, because the SEJD model does not take into account the interdependence structure among jumps. With respect to the parameter estimates of the MEJD model, compared to Table 4, the self excitor of the US is smaller and not significant at the 95% level. Nevertheless, the cross section excitors from the US to Japan and EU are statistically significant and large in magnitude compared to the reverse directions.

Excitation parameter estimates: Lagging US returns by one day

	SEJD			MEJD		
α	14.39***			31.17***		
$\boldsymbol{\beta}$	11.80***	0	0	4.31	0.00	13.80***
	0	8.55***	0	9.14***	7.35***	2.90
	0	0	13.66***	15.35***	3.18	13.30***

Table 12: Excitation parameter estimates when the US returns are lagged by one day. "SEJD" represents the self exciting jump diffusion model and "MEJD" represents the mutually exciting jump diffusion model. *, **, *** indicate significance at 95%, 97.5%, and 99.5% confidence levels, respectively.

F.3 Sub-sample estimation

As a third robustness check, we estimate the excitation parameters over a subsample of the full sample, excluding data from the first one-third of the sample. Table 13 reports the subsample parameter estimates. The excitation structure is not qualitatively different from the full-sample estimation results.

Excitation parameter estimates: Excluding the starting one-third of the sample

	SEJD			MEJD		
α	18.01***			36.75***		
$\boldsymbol{\beta}$	11.86***	0	0	14.65***	0.00	8.88**
	0	8.96***	0	9.92**	6.71**	4.34
	0	0	13.03***	28.04***	5.36	5.27*

Table 13: Excitation parameter estimates over a subsample of the full sample, excluding the starting one third of the sample. "SEJD" represents the self exciting jump diffusion model and "MEJD" represents the mutually exciting jump diffusion model. *, **, *** indicate significance at 95%, 97.5%, and 99.5% confidence levels, respectively.

References

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*(2), 556–584.

Aït-Sahalia, Y., J. Cacho-Diaz, and R. J. Laeven (2014). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics, forthcoming*.

Aït-Sahalia, Y. and T. Hurd (2012). Portfolio choice in markets with contagion. *working paper*.

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

Ang, A. (2014). *Asset Management: A Systematic Approach to Factor Investing*. Oxford University Press.

Ang, A. and G. Bekaert (2002). International asset allocation with regime shifts. *Review of Financial Studies 15*(4), 1137–1187.

Ang, A. and J. Chen (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics 63*(3), 443–494.

Bekaert, G. and X. Wang (2009). Home bias revisited. *Columbia Business School, working paper*.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies 22*(11), 4463–4492.

Bollerslev, T. and V. Todorov (2011). Tails, fears, and risk premia. *Journal of Finance 66*(6), 2165–2211.

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

Boyle, P., L. Garlappi, R. Uppal, and T. Wang (2012). Keynes meets Markowitz: The trade-off between familiarity and diversification. *Management Science 58*(2), 253–272.

Branger, N., C. Schlag, and E. Schneider (2008). Optimal portfolios when volatility can jump. *Journal of Banking & Finance 32*(6), 1087–1097.

Buraschi, A., P. Porchia, and F. Trojani (2010). Correlation risk and optimal portfolio choice. *The Journal of Finance 65*(1), 393–420.

Chacko, G. and L. M. Viceira (2005). Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *Review of Financial Studies 18*(4), 1369–1402.

Chan, K., V. Covrig, and L. Ng (2005). What determines the domestic bias and foreign bias? Evidence from mutual fund equity allocations worldwide. *The Journal of Finance 60*(3), 1495–1534.

Christoffersen, P., K. Jacobs, and C. Ornthanalai (2012). Dynamic jump intensities and risk premiums: Evidence from S&P 500 returns and options. *Journal of Financial Economics 106*(3), 447–472.

Coval, J. D. and T. J. Moskowitz (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance 54*(6), 2045–2073.

Cox, J. and C. Huang (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory 49*, 33–83.

Cox, J. C. and S. A. Ross (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics 3*(1), 145–166.

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

Dassios, A. and H. Zhao (2013). Exact simulation of hawkes process with exponentially decaying intensity. *Electronic Communications in Probability 18*(62).

Diyarbakirlioglu, E. (2011). Domestic and foreign country bias in international equity portfolios. *Journal of Multinational Financial Management 21*(5), 301–329.

Duffie, D. (2010). *Dynamic asset pricing theory.* Princeton University Press.

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

Epstein, L. and J. Miao (2003). A two-person dynamic equilibrium under ambiguity. *Journal of Economic Dynamics and Control 27*(7), 1253–1288.

Eun, C. S. and S. Shim (1989). International transmission of stock market movements. *Journal of Financial and Quantitative Analysis 24*(02), 241–256.

Ferreira, M. A. and P. Matos (2008). The colors of investors money: The role of institutional investors around the world. *Journal of Financial Economics 88*(3), 499–533.

Forbes, K. J. (2010). Why do foreigners invest in the United States? *Journal of International Economics 80*(1), 3–21.

French, K. and J. Poterba (1991). Investor diversification and international equity markets. *American Economic Review 81*, 222–226.

Garlappi, L., R. Uppal, and T. Wang (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies 20*(1), 41–81.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

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

Hawkes, A. (1971a). Point spectra of some self-exciting and mutually exciting point processes. *Journal of the Royal Statistics Society, Series B 33*, 438–443.

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

Jin, X. and A. Zhang (2012). Decomposition of optimal portfolio weight in a jump-diffusion model and its applications. *Review of Financial Studies 25*(9), 2877–2919.

Kang, J.-K., R. Stulz, et al. (1997). Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan. *Journal of Financial Economics 46*(1), 3–28.

King, M. A. and S. Wadhwani (1990). Transmission of volatility between stock markets. *Review of Financial studies 3*(1), 5–33.

Liu, J. (2007). Portfolio selection in stochastic environments. *Review of Financial Studies 20*(1), 1–39.

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

Liu, J. and J. Pan (2003). Dynamic derivative strategies. *Journal of Financial Economics 69*(3), 401–430.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics 51*(3), 247–57.

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

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics 8*(4), 323–361.

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

Rapach, D. E., J. K. Strauss, and G. Zhou (2013). International stock return predictability: What is the role of the United States? *The Journal of Finance 68*(4), 1633–1662.

Santa-Clara, P. and S. Yan (2010). Crashes, volatility, and the equity premium: Lessons from s&p 500 options. *The Review of Economics and Statistics 92*(2), 435–451.

Schied, A. and C.-T. Wu (2005). Duality theory for optimal investments under model uncertainty. *Statistics & Decisions 23*(3), 199–217.

Sialm, C., Z. Sun, and L. Zheng (2013). Home bias and local contagion: Evidence from funds of hedge funds. *working paper*.

Solnik, B. (1974). Why not diversify internationally rather than domestically? *Financial Analysts Journal 30*(4), 48–52.

Uppal, R. and T. Wang (2003). Model misspecification and underdiversification. *Journal of Finance 58*(6), 2465–2486.

Wachter, J. A. (2002). Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal of Financial and Quantitative Analysis 37*(01), 63–91.