International Correlation Asymmetries: Frequent-but-Small and Infrequent-but-Large Equity Returns

Bruno Solnik
Hong Kong University of Science and Technology

Thaisiri Watewai
Chulalongkorn University

We propose a novel regime-switching model to study correlation asymmetries in international equity markets. We decompose returns into frequent-but-small diffusion and infrequent-but-large jumps and derive an estimation method for many countries. We find that correlations due to jumps, not diffusion, markedly increase in bad markets, leading to correlation breaks during crises. Our model provides a better description of correlation asymmetries than do GARCH, copula, and stochastic volatility models. Good and bad regimes are persistent. Regime changes are detected rapidly, and risk diversification allocations are improved. Asset allocation results in- and out-of-sample are superior to other models, including the $1/N$ strategy. (*JEL* G01, G11, G15)

Received September 26, 2015; accepted May 25, 2016 by Editor Wayne Ferson.

Past crises have taught us that breaks in volatility and also in correlation could lead to huge losses in global portfolios, as illustrated by LTCM in the 1998 crisis or the 2008-2009 crisis. Stochastic volatility has been extensively studied and modeled, but correlation much less so. A simple example may help illustrate the importance of correlation. Let's consider a portfolio evenly spread between ten uncorrelated investments, each with a volatility of 10%. The portfolio volatility is 3.16%. If markets become more volatile with each volatility doubling from 10% to 20%, the portfolio volatility would increase to 6.32% as long as the correlation remains null. But if correlation breaks and

We would like to thank Wayne Ferson (editor) and an anonymous referee for their valuable comments that help improve this paper significantly. We thank Geert Bekaert, Ken Singleton, and seminar participants at Hong Kong University of Science and Technology, National University of Singapore, 2012 FMA Asian Conference, and 2011 Chulalongkorn Accounting and Finance Symposium for their useful comments and suggestions. Bruno Solnik gratefully acknowledges financial support from the HKUST Business School SBI; he is also a Distinguished Emeritus Professor at HEC-Paris. Thaisiri Watewai gratefully acknowledges financial support from Grants for Development of New Faculty Staff, Ratchadaphiseksomphot Endowment Fund, Chulalongkorn University, the Thailand Research Fund, and the Office of the Higher Education Commission (MRG-5380271). Send correspondence to Thaisiri Watewai, Department of Banking and Finance, Chulalongkorn Business School, Chulalongkorn University, Bangkok 10330, Thailand; telephone +66 22185691 ext. 13. E-mail: thaisiri@cbs.chula.ac.th.

doi:10.1093/rapstu/raw005 Advance Access publication July 5, 2016

shoots up to one, then the portfolio volatility would increase to 20%. The increase from 6.32% to 20% is solely explained by the correlation break. While simplistic, this example illustrates the need to model time-varying correlation in portfolio management. This paper focuses on the correlation of international equity markets.

Evidence of asymmetrical return behavior in international equity markets is well documented. One of the key findings in the literature is that returns tend to be more correlated when volatilities are high or when the markets go down.[1] Earlier studies focused on the link between higher volatility and higher international correlation. However, as pointed out by Stambaugh (1995), Boyer et al. (1999), and Forbes and Rigobon (2002), correlation is a complex nonlinear function of returns, and empirical estimates of correlations suffer from the relationship between correlation and volatility. This can introduce severe statistical biases in the estimation of correlation when conditioning on the level of returns or of volatility.[2] Hence, tests must directly model the return distributions. Longin and Solnik (2001) use extreme value theory to estimate correlations conditional on returns being extreme (*exceedance* correlation).[3] They find that extreme returns exhibit asymmetries. Correlations conditional on negative returns are higher than correlations conditional on positive returns. Furthermore, correlations conditional on negatively large returns tend to increase in the magnitude of returns, but correlations conditional on positively large returns do not. Their results suggest that direction (good or bad), not volatility per se, would induce international correlation asymmetry. This international "correlation break" for extreme negative returns (some authors call it "correlation breakdown") is a striking result that has been confirmed in subsequent studies. It implies that the benefits of international diversification are vastly reduced in bear markets, when it is most needed, and that international asset allocation should adapt to it. Many subsequent models have tried to better fit observed international data, but they usually fail to reproduce the observed break in correlation.

Early work develops international models with regime switching (Ang and Bekaert, 2002) or jumps (Das and Uppal, 2004). Following Longin and Solnik (2001), most recent works rely on estimated copula models with various heteroscedastic/asymmetric distributions or complex stochastic volatility models. We later estimate models from several important classes of these competitor models and compare their results with ours.

---

1 See, for example, Erb et al. (1994), Longin and Solnik (1995, 2001), Ramchand and Susmel (1998), Ball and Torous (2000), Ang and Bekaert (2002), Campbell et al. (2002), and Okimoto (2008), among others.

2 For example, assume that the return distribution of two markets is bivariate normal with a correlation of 0.5. If we condition on a market having a large absolute return (positive or negative), say the 10% largest returns in the distribution, then the computed correlation jumps to 0.77. But this is a spurious increase in correlation due to the choice of sample, as the true correlation is constant at 0.50. See Boyer et al. (1999).

3 Exceedance correlation is the correlation of returns that exceed a given threshold, for example, greater than 10% or lower than −10%.

In this paper, we aim to improve modeling return asymmetries, with a focus on international correlation. Besides standard econometric tests, our empirical analysis of competing models partly focuses on fitting the observed extreme correlation; breaks in correlation are important in managing global portfolios in changing markets. Taking into account those asymmetries should improve risk modeling to provide better asset allocation and risk management models. We derive optimal asset allocation weights for our model and conduct out-of-sample tests against other simpler models, including the famous "$1/N$" rule. We find significantly better results.

We propose a model that allows correlation asymmetries for both normal and extreme returns. Specifically, returns are decomposed into frequent-but-small and infrequent-but-large components and are referred to as diffusion and jumps, respectively.[4] We use systemic jumps to represent infrequent but strong comovements across markets. To capture asymmetries in diffusion and jump components, we use a regime-switching process that allows correlations and other parameters, including means and volatilities of diffusion, jump size distributions, and jump arrival rates to be stochastic. Hence, the whole distribution of diffusion and jumps can vary across regimes. Our relatively simple modeling choice allows for stylized facts, such as fat tails, skewness and time-varying volatility and correlations. It can also model crises with increased volatility and correlation breaks. After detailing our theoretical model, we proceed as follows.

We derive an efficient estimation method for our regime-switching model with jumps that allows us to estimate the model parameters for a large number of countries. While Das and Uppal (2004) impose a one-factor structure (perfect correlation) on the jump size distribution to reduce the number of parameters, our method is able to estimate the joint distribution of random jump sizes across multiple countries without such restrictions. This allows us to study the correlation asymmetries of jump components across regimes. As discussed below, the change in jump correlation across regimes is, in fact, a major cause of the correlation break, a drastic increase in the conditional correlation observed during bad markets. Our method also provides the estimates of the probabilities of the underlying market regime, which is assumed to be unobservable. These probabilities can be used to infer the likelihood of the current market conditions, which are essential in financial applications, such as portfolio choices and risk management. For example, our model enables us to detect rising crisis probabilities and an upward adjustment of the volatility and correlation after observing a short series of global returns with large negative values.

4 In this article we use the terminology "large/small" to refer to the absolute (unsigned) magnitude and "high/low" to refer to the signed magnitude. For example, a "large" return means a return of large size, whether positive or negative, and a "high" return means a high positive return.

We provide new empirical findings from the top ten major equity markets over the period 2001-2013, using weekly data. Our study finds strong evidence of the existence of two global market regimes (good and bad) and systemic jumps. Consistent with earlier works, returns are low, more correlated, and more volatile in one regime (the bad regime). The persistence of the regimes is an important feature. It shows that markets go through prolonged periods of good and bad regimes. The probability of remaining in the same regime is high and the expected regime duration is long (31 weeks for the bad regime and 67 weeks for the good regime). Estimating that markets enter a regime gives useful information for the future, providing support for a regime-switching model. Jump sizes are quite different in the two regimes. A large jump can lead to a rapid transition to the bad regime, and the jump regime is persistent. While our regimes are correlated with standard U.S. economic indicators, such as financial stress or VIX, the model gives earlier warnings of market crises. Our results suggest that the increased international correlation across markets can provide an early detection of a global crisis.

Interestingly, correlations due to diffusion are similar across the regimes, but correlations due to jumps increase markedly in bad markets and are much higher than those of diffusion. These highly correlated jump sizes can be viewed as the source of a correlation break during bad markets. These new findings help explain the observed asymmetries in correlations. We confirm this by running a horse race between our model and various models, including DCC GARCH, factor copula, and multivariate factor stochastic volatility. The results show that our model provides a much better fit to the actual data for asymmetry, the magnitude of exceedance correlations, and quantile dependence. We check whether the number of regimes should be increased or country-/region-specific jumps should be added. We fail to reject our simpler model.

We solve *dynamic* optimal portfolio choices in a multiple-country setting with *unobservable* market regimes and jumps. The regime unobservability is typically ignored in traditional dynamic portfolio choices with regime switching, such as Ang and Bekaert (2002) who assume that investors solve optimal portfolio choices based on the knowledge of the regime. In contrast, we assume that investors solve optimal portfolio choices based on their *beliefs* in the market regimes using a Bayesian framework. We are able to solve the problems with a relatively large set of countries with hidden regimes.[5] The analysis shows that asymmetries in correlations have significant effects on portfolio diversification. Modeling jumps in a regime-switching model allows

[5] Honda (2003) and Guidolin and Timmermann (2007) also follow this unobservable regime assumption. However, Honda (2003) only allows the mean return to be regime dependent in a geometric Brownian motion model and considers one risky asset, while we allow all model parameters to be regime dependent and consider ten countries. Guidolin and Timmermann (2007) assume a regime-switching VAR model and consider a problem with four assets (small-cap equity, large-cap equity, bond, and Treasury bill). Neither of these papers allows jumps in asset prices.

a better characterization of the good and bad regimes and markedly different optimal weights compared with a regime-switching model with only diffusion processes or a single-regime model with jumps. Our model also provides superior *out-of-sample* performance over various models, including the well-known robust $1/N$ portfolio. There are two reasons for this. First, bad regimes are detected early and with the increase in the bad-regime probability, investors reduce their exposure to risky assets or even move totally to the risk-free asset. Second, investors take advantage of the better estimation of time-varying correlation to select better-diversified asset allocations with higher Sharpe ratios. As regimes are persistent, these two effects produce strong benefits.

1. Literature Review

An early work in correlation-asymmetry modeling with asset allocation implications is Ang and Bekaert (2002) who estimate a Markov-switching multivariate-normal model. They have two regimes, and each regime has multivariate normal returns. They find some (weak) evidence of a bear regime characterized by low expected returns, high volatility, and high correlation and a normal regime associated with high expected returns, low volatility, and low correlation. They show that their model can replicate some of Longin and Solnik's (2001) results, although it underestimates the correlation for large negative exceedances (see also Okimoto, 2008). They also show that the asymmetric bivariate GARCH model, widely used in the past, cannot replicate these results. For computational reasons, they cannot estimate their regime-switching model for more than three markets, and the asset allocation results assume that investors know with certainty which regime they are in. Another early work is that of Das and Uppal (2004) who propose a model in which global shocks periodically affect all markets simultaneously. The return-generating process is Brownian with a jump component à la Merton (1976). Jumps arrive at the same time in all markets, and the jump sizes are *perfectly* correlated. Das and Uppal (2004) are able to derive explicit analytical solutions to the portfolio weights. However, returns in the single-regime model of Das and Uppal (2004) are basically i.i.d. with no persistence in jumps and have constant correlations conditional on the return history. The model cannot satisfactorily replicate the observed exceedance correlations (see Ang and Chen, 2002).

Following Longin and Solnik (2001), most recent works rely on some estimated copula models with various heteroscedasic/asymmetric distributions or complex stochastic volatility models. Okimoto (2008) uses a family of asymmetric copula models with lower-tail dependence (Longin and Solnik (2001) use one of these copula models). He further introduces two Markov-switching regimes to empirically fit the exceedance correlation. One regime has lower-tail dependence and the other does not. This rich empirical model allows for a good fit to actual returns and exceedance correlation in extreme events, but is limited to pair-wise analysis (two countries at a time) given the number of parameters to be estimated.

Estimating a multivariate model with asymmetries in correlation for many countries is challenging. To make estimation feasible, most multivariate models impose restrictive structures, leading to poor empirical fit, especially for the asymmetries of extreme correlations. For example, the popular dynamic conditional correlation or DCC GARCH of Engle (2002) usually relies on a small number of parameters (e.g., 3 for DCC(1,1)) to model the time-varying correlation matrix, imposing the same dynamic correlation structure between each pair of returns. Cappiello et al. (2006) propose the asymmetric generalized DCC or AG-DCC model that allows each correlation to have its own dynamic. However, they need to impose a diagonal covariance matrix when they estimate their model with many assets. Similarly, the multivariate copulas for three or more assets that allow asymmetries between the upper and lower tail dependences (such as multivariate Archimedean and skewed-t copulas) impose the same asymmetric dependence structure on all pairs of returns characterized by a single parameter. Christoffersen et al. (2012) model the correlation among a large set of countries with a *dynamic asymmetric copula* (or DAC) relying on the DCC structure and skewed-t copula. The interdependence among markets is modeled as a mean reversion, plus a constant time-trend adjustment in the conditional correlation. This approach detects a time trend in correlation among a large number of countries, but seems too restrictive for the study of asymmetries and changes in correlation during market crises.[6] Oh and Patton (2012) propose a *factor copula* for modeling dependence in a high-dimensional time series. They impose factor and block structures to limit the number of parameters. They fit the model for up to 100 time series with eight factors. Although the factor copula approach has been shown to be very powerful for handling a large dataset, we show in this paper that a factor copula model with multiple fat-tailed, asymmetric factors underestimates the extreme correlations in international equity markets.

Models such as DCC, DAC, and factor copula share the same drawback; that is, the (implied) covariances are assumed to be conditionally deterministic. *Stochastic volatility* models provide an alternative approach to modeling variances and covariances, by introducing latent parameters that drive the covariance matrix. This provides superior performance in both in-sample fitting and out-of-sample forecasting of financial time series.[7] A major disadvantage of multivariate stochastic volatility models is that they have no

6 For example, in their Figure 9, their exceedance correlation is too low for bad markets and it decreases as the threshold becomes increasingly negative, while the observed correlation increases with the (negative) threshold as found on actual data.

7 See, for example, Kim et al. (1998), and Yu (2002).

closed-form likelihood function. Estimation of such models is therefore computationally demanding, and derivation of explicit optimal portfolio choices is not possible. Omori and Ishihara (2012) propose a rich multivariate stochastic volatility model allowing for asymmetries, volatility clustering, and leverage effects. They impose a factor structure to make estimation feasible when the number of assets is large. Our empirical study shows that their model allows for asymmetries in exceedance correlations, but underestimates the magnitude of the correlations.

Ait-Sahalia, Cacho-Diaz, and Laeven (2015) propose a Hawkes jump-diffusion process for asset returns in which arrivals of jumps in one region increase the likelihood of future jumps in that and other regions. Their model is good for studying financial contagion. However, they assume constant correlation in diffusion and independent jump sizes, due to its complexity, and can estimate the model with only two assets. Ait-Sahalia and Xiu (2016) use high-frequency data (one minute or less) to study the impact of news on the pairwise correlation of future prices before and after the trading session. They find that correlations of diffusion and jumps increase in periods of crisis. They, however, do not study international equity correlation, correlation of extreme returns, or asset allocation.

Pukthuanthong and Roll (2015) present an empirical study of jump correlation across eighty-two countries using DataStream daily data. They detect daily jumps within each month using the BNS-G statistic developed by Barndorff-Nielsen and Shephard (2006). A significant negative BNS-G indicates a jump. For each country, they get a monthly time series of BNS-G and compute an average of pairwise country correlations of those BNS-G. They conclude that the international correlation of country jumps is low. They also compare the correlation of monthly returns of months with jumps ("jump months") and months with no jumps ("nonjump months") and find a much lower correlation for jump months. Their conclusion sharply contrasts with our result, and this is worth some discussion.[8] We provide three main reasons to explain the differences.

The first reason is the difference in the universe. The vast majority of their eighty-two markets are classified as emerging, frontier, or not classified by FTSE or MSCI. This could partly explain the differences as emerging and frontier markets can suffer from national socio-economic-political shocks that are country specific. On the other hand, their results on the BNS-G correlation between developed European countries suggest a conclusion closer to ours based on the top ten developed markets.

Second, the quality of DataStream daily data is poor for many markets, leading to the detection of spurious jumps caused by data errors. A major problem is that there are numerous months in which daily prices hardly

8 We thank Richard Roll for providing us with his data, computer codes, and output. Statistics reported below are based on these data.

move during the month, but then jump on the last day of the month. The last daily return captures the full monthly return, leading to extremely negative BNS-G, far larger than those during good-quality data periods. This is true for many emerging/frontier markets during many years, but also for some developed markets, such as Canada, Denmark, or Sweden.[9] These spurious and extreme BNS-G statistics induce biases in the BNS-G cross-country correlations, as well as in the return correlations conditional on jump or nonjump months. These errors lead to a downward bias in BNS-G correlations with other countries, except in a few cases in which the error periods overlap for the two countries.

Third, differences exist between the two methodologies.[10] The BNS methodology is designed to detect discontinuity in the continuous-time price path. The monthly estimated BNS-G is a test statistic of jump detection within a given month, an indicator based on a mixture of second and fourth moments of returns. The precise relationship between the BNS-G correlation and the existence of common jumps with time-varying means, volatility ratio, and correlation is not clear.[11] We model systemic jumps and directly focus on the correlation of returns and jump sizes. In our regime-switching model, we find that the ratio of jump volatility to diffusion volatility is much larger in the bad regime and so are the mean jump size and the correlation of jump sizes. But smaller jumps frequently occur in the good regime.

2. Modeling Return Asymmetries with Jumps

Returns are decomposed into two components. The first component is to capture the relatively small normal movements. We refer to this component as frequent but small ("diffusion" in continuous time). The second is to capture movements that rarely occur, but once they occur, they have a

[9] Let's take the example of Canada from Sep 1971 to Dec 1985 (172 months); the daily price variation is minuscule, but not zero, except for a big jump on the last day of the month. Any jump during that period always occurs on the last day of the month. It leads to hugely negative BNS-G for most months. The BNS-G is asymptotically unit normal if no jumps occur. But the mean and median of BNS-G for Canada over that period are –235 and –53. So roughly 50% of months (86 months) have a BNS-G lower than –53. Over the remaining period of Jan 1986 to Oct 2009 (286 months), the median BNS-G is –0.37, an order of magnitude found for other countries without data problems, and the largest negative BNS-G over 1986-2009 is only –8.8. Hence, all large jumps detected for Canada are caused by data errors.

[10] An apparent error in the BNS-G formula leads to a severe under-detection of jumps. While the pairwise BNS-G correlation is mostly invariant to the scale factor caused by the error, the error seriously affects the correlation of raw returns conditional on months with or without jumps (Table 7, panel B). For example, the mean correlation of jump months increases from 0.083 to 0.259, when the correct formula is applied to these data. That is because the wrong formula fails to detect many jumps during the periods with good-quality daily data, while the correct BNS-G formula finds roughly four times as many jumps. More specific details are available from the authors.

[11] Simply comparing the magnitude of the cross-country correlation of a jump-detection indicator and of returns is not meaningful as these are very different concepts. Simple simulations show that correlation of returns can be much larger than correlation of BNS-G even though all jumps are common jumps (i.e., systemic), and all parameters are time invariant. For example, we find the monthly return correlation is five times larger than the BNS-G correlation (0.415 vs. 0.082) for a daily-return model with two stocks in which the volatility and correlation of the diffusion components are 0.02 and 0.4, while those for jump sizes are doubled (0.04 and 0.8), and jumps are systemic with jump arrival rate of 0.01. Both diffusion and jumps have zero means.

relatively large magnitude. We refer to this component as infrequent but large ("jumps" in continuous time). Correlation across markets comes from correlations between diffusion components (frequent-but-small) and correlations between jumps (infrequent-but-large). Although our model is presented as a discrete-time model, the modeling concept can be easily extended to a continuous-time model.[12]

To model asymmetries across the market regimes, the parameters of both return components are allowed to be regime dependent and hence stochastic. Specifically, let $\{Y_t, t = 1, 2, \ldots\}$ denote a time-homogeneous K–state discrete-time Markov chain representing the market regime, and $p_{y,z}$ denotes the associated transition probability from regime y to z. Also, let $\varrho_y = P(Y_1 = y)$ denote the probability of the initial state. The log-return process of asset i is denoted $\{R_{i,t}, t = 1, 2, \ldots\}$, where

$$R_{i,t} = Z_{i,t} + \sum_{m=1}^{\Delta N_t} \delta_{i,t}^m \quad i = 1, \ldots, n, \quad t = 1, 2, \ldots. \tag{1}$$

Conditional on the current regime $Y_t = y$, $Z_t = [Z_{1,t}, \ldots, Z_{n,t}]'$ is an n–dimensional Gaussian random vector representing the diffusion component with mean vector $\mu(y) = [\mu_1(y), \ldots, \mu_n(y)]'$, and variance-covariance matrix $\Sigma(y) = [\sigma_{ij}^{\Sigma}(y)]$, ΔN_t is the random number of jumps that occur during time period t for which the distribution follows a Poisson distribution with a scalar parameter $\lambda(y)$ representing the expected number of jumps, and $\delta_t^m = [\delta_{1,t}^m, \ldots, \delta_{n,t}^m]'$ is the vector of *random* jump sizes from the m^{th} jump for which the distribution is Gaussian with mean vector $\eta(y) = [\eta_1(y), \ldots, \eta_n(y)]'$ and variance-covariance matrix $\Omega(y) = [\sigma_{ij}^{\Omega}(y)]$. The jump sizes δ_t^m are assumed to be conditionally independent across m given Y_t. That is, the distributions of the diffusion, the number of jumps, and the jump sizes depend on the market regime. We further assume that the random vector Z_t, the number of jumps ΔN_t, and the jump sizes δ_t are conditionally independent given Y_t.

Jumps are assumed to be *systemic* or *global-wide*. That is, an arrival of a jump affects returns of *all* assets at the instant of the arrival, but the jump size of each asset can be different.[13] It is technically easy to augment the model to include additional types of jumps whose arrivals affect the returns of a *subset* of assets. This can be used, for example, to model *country-specific* or *region-specific* jumps. However, we find that including country-specific or region-specific jumps, *in addition to the global-wide jumps*, leads to overfitting (Section 4.4).

[12] The detail of a continuous-time counterpart is available on request.

[13] If jumps were primarily country-specific jumps rather than systemic (global) jumps, with only one country $\delta_{i,t}$ being nonzero at any arrival time, our model with a single systemic jump would be poorly specified, and we would find that the correlation of jump sizes is very small. However, we do find high correlation, suggesting systemic jumps. Furthermore, we reject the need to add country-specific jumps to significantly improve the fit of the model.

Our model can be viewed as a generalized discrete-time version of Das and Uppal (2004) who consider these types of Gaussian systemic jumps. They, however, assume that the jump sizes δ_i are perfectly correlated, while we estimate the correlations of jump sizes from the data. They also assume a single regime, but allow the jump intensity to be regime dependent in their earlier working paper. In contrast, all model parameters are regime dependent in this paper.

Note that correlations between asset returns come from two components: the vector of correlated Gaussian random variables Z and the systemic jump with correlated jump sizes δ. The correlations due to Z can be viewed as a measure of the degree of small comovements between asset returns (the frequent-but-small or diffusion component).[14] On the other hand, the correlations due to δ can be viewed as a measure of the degree of large comovements between asset returns (the infrequent-but-large component or jumps). This enables us to explicitly investigate the regime shifts of different types of correlations. As a result, the correlation asymmetries of small and large comovements can be examined separately. To summarize, our model is fully characterized by the following set of parameters $\Theta = \{\varrho, p, \mu, \Sigma, \lambda, \eta, \Omega\}$.

Finally, the mean, standard deviation, and correlation of log-returns can be derived from Θ as follows. The conditional mean return of asset i is

$$\overline{\mu}_i(y) = \mu_i(y) + \lambda(y)\eta_i(y), \tag{2}$$

the conditional standard deviation of asset i is

$$\overline{\sigma}_i(y) = \sqrt{\sigma_{ii}^{\Sigma}(y) + \lambda(y)[\sigma_{ii}^{\Omega}(y) + \eta_i(y)^2]}, \tag{3}$$

and the conditional correlation between assets i and j is

$$\overline{\rho}_{ij}(y) = \frac{\rho_{ij}^{\Sigma}(y)\sigma_i^{\Sigma}(y)\sigma_j^{\Sigma}(y) + \lambda(y)[\rho_{ij}^{\Omega}(y)\sigma_i^{\Omega}(y)\sigma_j^{\Omega}(y) + \eta_i(y)\eta_j(y)]}{\overline{\sigma}_i(y)\overline{\sigma}_j(y)}, \tag{4}$$

where $\sigma_i^{\Sigma} = \sqrt{\sigma_{ii}^{\Sigma}}$, $\sigma_i^{\Omega} = \sqrt{\sigma_{ii}^{\Omega}}$, and ρ_{ij}^{Σ} and ρ_{ij}^{Ω} are the correlations implied from the covariance matrices Σ and Ω, respectively.[15]

The model defined by (1) is a hidden Markov model (HMM) with unobservable states Y_t and observable variables $R_{i,t}$, which are log-returns of asset i, $i = 1, \ldots, n$. Note that given the current state Y_t, the distribution of $R_{i,t}$ is an infinite mixture of normal distributions due to the randomness of the number of jumps with Gaussian jump sizes. Maximizing the unconditional log-likelihood function of these non-i.i.d. infinite-mixture returns using a search algorithm for a large number of assets is computationally infeasible.

[14] In our approach we model the full covariance structure of returns and do not impose any factor structure, such as global or regional factors. Empirical comparisons with factor models provided in Section 5 show that our model captures observed extreme correlation asymmetries much better than the factor models.

[15] The terms $\eta_i(y)^2$ in (3) and $\eta_i(y)\eta_j(y)$ in (4) come from the randomness of the number of jumps ΔN_t.

In the next section, we derive a powerful estimation method that allows us to obtain the maximum likelihood estimators of this model.

3. Estimation Method

We derive an efficient estimation method based on the framework of the expectation maximization (EM) algorithm for our model. Although tractable EM algorithms for certain regime switching models have been proposed in the literature, deriving a tractable EM algorithm for *multivariate* models with regime-switching *and* jumps is nontrivial. Nevertheless, we are able to obtain a tractable algorithm for a large number of assets as described below.

The general framework of the EM algorithm was first proposed by Dempster, Laird, and Rubin (1977). The EM algorithm is an iterative method for computing maximum likelihood estimators of model parameters when some variables are missing or unobservable. More specifically, let $\mathbb{X}_t$ denote the *observed* or *incomplete* data up to time t, $\mathbb{Y}_t$ the *unobserved* data up to time t, and $\mathbb{C}_t = \mathbb{X}_t \cup \mathbb{Y}_t$ the *complete* data up to time t. Given a set of parameters $\Theta^{(p)}$ at iteration p, the algorithm finds the expected value of the complete-data log-likelihood given $\mathbb{X}_T$ (E-step):

$$Q(\Theta, \Theta^{(p)}) = \mathbb{E}[log\, L(\mathbb{C}_T \mid \Theta) \mid \mathbb{X}_T, \Theta^{(p)}], \tag{5}$$

where $L(\mathbb{C}_T \mid \Theta)$ is the likelihood of the complete data with parameter Θ. The conditional expectation is taken over the random unobserved data $\mathbb{Y}_T$ for which the distribution is specified by parameter $\Theta^{(p)}$. Then the expectation is maximized to obtain a new set of parameter estimates (M-step):

$$\Theta^{(p+1)} = arg\, \max_{\Theta}\, Q(\Theta, \Theta^{(p)}). \tag{6}$$

These two steps are performed alternately until convergence. It can be shown that the log-likelihood of the observed data $L(\mathbb{X}_T \mid \Theta^{(p)})$ is nondecreasing in each iteration p (see Dempster et al., 1977) and that under regularity conditions the algorithm converges to a local maximum solution (see Wu, 1983). Initializing the parameters at various points throughout the parameter space may increase the chance of getting to a global maximum solution.

In this paper, the set of the parameters of our model is $\Theta = \{\varrho, p, b, \Sigma, \lambda, \eta, \Omega\}$, and we set $\mathbb{X}_t = \{r_1, \ldots, r_t\}$ and $\mathbb{Y}_t = \{Y_1, \ldots, Y_t, Z_1, \ldots, Z_t, \Delta N_1, \ldots, \Delta N_t, \delta_1, \ldots, \delta_t\}$, where r_t denotes the observed vector of asset returns at time t. That is, we treat returns as observable variables, but treat the regime, diffusion components, number of jumps, and jump sizes as unobserved variables.[16] In general, one has to compute conditional

[16] Pickard, Kempthorne, and Zakaria (1986) propose to include only the number of jumps (ΔN) in the unobserved data, and Duncan, Randal, and Thomson (2009) propose to include the diffusion components (Z), number of jumps (ΔN), and jump sizes (δ) in the unobserved data, leading to simpler implementation. However, both consider univariate jump diffusion models without regime switching.

expectations in the E-step (5) and solve maximization problems in the M-step (6), both of which are model specific and often require simulations and numerical optimization. This is especially the case when observations are non-i.i.d. and infinite-mixture, and this mix leads to computationally demanding algorithms for multivariate models. With our choice of unobserved variables, we are able to analytically derive the explicit equations for computing related conditional expectations in the E-step and obtain closed-form solutions for maximization problems in the M-step. In addition, we obtain the filtered and smoothed probabilities of the market regimes, as well as those of the diffusion and jump components as by-products of our method. The details of the estimation method are given in Internet Appendix A.

Our model have a large number of parameters, especially when the number of assets is large. So the identification of our model should be discussed. In Internet Appendix B we provide an outline of a proof that our model is indeed identified. The identification of our model lies in the fact that any finite mixture of the return distributions within each regime is identified and that regime switching provides structure to the model. To illustrate the identification of our model, first consider a single-regime model with ten assets. The diffusion component captures 10 means and 55 variances/covariances with 65 parameters. Introducing jumps adds 66 more parameters (10 means, 55 variances/covarainces, and 1 jump arrival rate), but allows a better match to additional 220 skewness/coskewness.[17] With two regimes, the number of parameters from the two regimes doubles to 262, and the initial and transition probabilities add three more, making it equal to 265 in total. The number of the first three unconditional moments remains at 285 (65 + 220) for ten assets, and this number is still higher than the number of parameters. When the number of regimes increases, our model provides a better match to higher unconditional moments, such as kurtosis/cokurtosis. Having more moments than parameters is merely a necessary condition for identifiability, but our proof confirms it.

4. Empirical Results

4.1 Data

We use weekly log-return data from ten Morgan Stanley Capital International (MSCI) country indices: Australia (AU), Canada (CA), France (FR), Germany (GE), Hong Kong (HK), Japan (JP), Spain (SP), Switzerland (SW), the United Kingdom (UK), and the United States (US). The data are obtained from the Thomson Reuters Datastream database for January 2001 to March 2013 (639 observations). In terms of market

[17] By skewness and coskewness, we mean the cross-moments of the form $\mathbb{E}[R_i^3]$, $\mathbb{E}[R_i^2 R_j]$ and $\mathbb{E}[R_i R_j R_k]$, where R_i, R_j and R_k are returns of three different assets i, j and k. For ten assets, we have 10 terms for $\mathbb{E}[R_i^3]$, 90 terms for $\mathbb{E}[R_i^2 R_j]$, and 120 terms for $\mathbb{E}[R_i R_j R_k]$.

Table 1
Summary statistics of weekly log-returns of country equity indices

	Average (%)	SD (%)	Skewness	Excess kurtosis	Median (%)	Min (%)	Max (%)
AU	0.252	3.682	–1.787	13.856	0.621	–34.30	14.94
CA	0.155	3.416	–1.181	8.915	0.444	–26.05	17.76
FR	0.048	3.610	–1.005	5.923	0.420	–26.69	13.88
GE	0.083	3.873	–0.868	4.905	0.493	–26.06	15.20
HK	0.142	3.037	–0.300	2.293	0.290	–17.11	10.32
JP	0.024	2.771	–0.399	2.207	0.091	–16.40	11.02
SP	0.097	3.992	–0.970	5.016	0.329	–26.07	13.43
SW	0.127	2.920	–1.118	9.612	0.333	–23.91	13.12
UK	0.076	3.125	–1.295	12.281	0.389	–27.57	16.28
US	0.065	2.649	–0.846	7.119	0.195	–20.05	11.58

This table provides summary statistics of weekly log-returns of ten MSCI country equity indices, including Australia (AU), Canada (CA), France (FR), Germany (GE), Hong Kong (HK), Japan (JP), Spain (SP), Switzerland (SW), United Kingdom (UK), and United States (US). The log-returns are computed from the total return indices in U.S. dollars. The data cover the period from January 2001 to March 2013 (639 observations).

capitalization, these ten markets are among the largest investable markets in the world as described in the MSCI Global Investable Market Index methodology. Some countries, such as India and China, are not included because their domestic markets are not open to foreign investors. However, Hong Kong companies are strongly linked to the Chinese economy, and many mainland Chinese companies are listed in Hong Kong. Weekly returns are used instead of daily returns to reduce the effect from nonsynchronous data.[18] Past research has often used monthly data over a long period. For example, Ang and Bekaert (2002) have 335 monthly return observations from 1970 to 1997. Although the number of data points is smaller because of the monthly frequency, the likelihood of extreme events and different regimes is higher over a longer period. However, the stability of these regimes over several decades is more questionable. We use total returns in U.S. dollars to take the viewpoint of a U.S. investor.

Table 1 shows the summary statistics of weekly returns of each index. All index returns have negative skewness, confirming the asymmetric distributions of log-returns. This negativity is well known in equity index returns. The minimum and maximum values show that negative shocks could be much larger than positive shocks. In particular, during the study period of January 2001 to March 2013, the most severe minimum weekly return is –34.30%, while the largest positive weekly return is + 17.76%. Such extreme returns, say beyond three standard deviations, have an extremely small probability of occurring under the normal distribution (0.27%), but we do observe them repeatedly during crises. Jumps can help explain those extreme returns. The returns also exhibit large positive excess kurtosis (fat tails). The value of excess kurtosis ranges from 2.21 for Japan to 13.86 for Australia. These very large values of excess kurtosis could be due to large movements in the equity indices

[18] Christoffersen et al. (2012) also used weekly returns.

and the nonstationarity of the return time series. These two possibilities are modeled by jumps and regime switching in our model. Regime switching addresses nonstationarity by including a bad/crisis regime, and it is well known that a jump component can be used to model fat-tail distribution (see, e.g., Liu et al., 2003).

4.2 Model selection: Do we need global regimes and systemic jumps?

Before presenting the detailed results of our model, we test whether a regime-switching model with jumps is a significant improvement over the models used in the past literature. We fit the model using the EM-based algorithm developed in Section 3 for the ten equity index returns. We run the model with four specifications: (1) one regime without jumps, (2) one regime with jumps, (3) two regimes without jumps, and (4) two regimes with jumps.

In choosing the best model among those with a different number of regimes and return stochastic processes, the usual likelihood ratio test is not applicable because the parameters associated with additional regimes of the model with a higher number of regimes are unidentified under the null hypothesis of the model with a lower number of regimes (see, e.g., Hansen, 1992). Similarly, parameters associated with jumps are unidentified under the null hypothesis of the model without jumps. To identify the number of regimes and the existence of jumps, we use the upper bound of the p–value derived by Davies (1987) for the hypothesis testing of the nested case, and we apply the test of Rivers and Vuong (2002) for the nonnested case. The details of the tests are given in Internet Appendix C.

We also compute the Akaike information criterion (AIC) given by $AIC =$ – 2(log-likelihood – number of parameters), which is a model selection criterion that relies on the trade-off between the goodness of fit (log-likelihood) and the complexity of the models (number of parameters). AIC values and the results of the p-values of hypothesis testing are provided in Table 2. According to the values of the AIC reported in panel A, our two-regime model with jumps is the most preferred (lowest AIC), and the two-regime model without jumps is preferred to the one-regime model with jumps. The one-regime without jumps (i.e., multivariate-normal model) is, as expected, the least preferred.

Panel B reports the upper bounds of the p-value of various Davies' hypothesis tests for the nested case (top row and rightmost column), and that of Rivers and Vuong's test for the nonnested case. The results strongly reject the null hypothesis of a single regime, as well as the null hypothesis of the models without jumps against our model with two regimes and jumps. A single-regime model with jumps, as suggested by Das and Uppal (2004), or a two-regime model without jumps, as suggested by Ang and Bekaert (2002), are both rejected at 0.01% significance level. The test of Rivers and Vuong strongly rejects the single-regime model with jumps in favor of the two-regime model without jumps, although both are dominated by our two-regime model

Table 2
Model selection

A. Model selection criteria

Criteria	Model			
	One regime without jumps	One regime with jumps	Two regimes without jumps	Two regimes with jumps
No. of parameters	65	131	133	265
Log-likelihood	16,632	17,082	17,188	17,396
AIC	−33,135	−33,903	−34,111	−34,261

B. Upper bound of the p-*value*

Null hypothesis	Alternative hypothesis		
	One regime with jumps	Two regimes without jumps	Two regimes with jumps
One regime without jumps	0.0000	0.0000	0.0000
One regime with jumps		0.0001	0.0000
Two regimes without jumps			0.0000

Panel A of this table provides the number of parameters, log-likelihood, and Akaike information criterion (AIC) for four different models: one regime without jumps, one regime with jumps, two regimes without jumps, and two regimes with jumps. The upper bounds of the *p*-value of the null hypothesis against the alternative hypothesis are provided in panel B. When the model under the null hypothesis is nested in the model under the alternative hypothesis (top row and rightmost column), the upper bounds are based on Davies (1987). For the nonnested case (one regime with jumps vs. two regimes without jumps), the upper bound is based on the test statistic of Rivers and Vuong (2002).

with jumps. Hence, our results imply the existence of global regimes and of systemic jumps among these ten markets.

Previous tests rely on statistical significance. But whether the model improves our understanding of the observed data and is economically meaningful is also important. A first step below is to show that our model yields a better fit to the moments of the unconditional distribution of observed returns. In Section 5 we will also show that our model provides a much better description of the observed international correlation structure, especially for extreme returns. Finally, Section 6 will discuss the economic importance in terms of asset allocation implications.

We examine the moments implied from the estimates of each model specification by computing the averages of unconditional values of mean, standard deviation, correlation, skewness, and excess kurtosis of the data, and those implied from the fitted models. The averages are computed across all ten countries, or forty-five country pairs for the correlation, and are reported in Table 3.[19] The maximum likelihood estimates from all four model specifications obtained from the EM-based algorithm can very well match the first and second moments (mean, standard deviation, and correlation) of the data.

[19] The formula of unconditional covariances, skewness, and kurtosis for the models with two regimes is available from the authors.

Table 3
Averages of unconditional moments

	Data	One regime without jumps	One regime with jumps	Two regimes without jumps	Two regimes with jumps
Mean (%)	0.107	0.107	0.107	0.109	0.107
SD (%)	3.307	3.305	3.318	3.298	3.327
Correlation	0.709	0.709	0.711	0.709	0.712
Skewness	−0.977	0.000	−0.586	−0.321	−0.931
Excess kurt	7.213	0.000	5.928	2.023	8.289

This table provides the unconditional moments of returns computed from the data and those implied from the estimates of four different models: one regime without jumps, one regime with jumps, two regimes without jumps, and two regimes with jumps. The values are the averages across all ten countries for the mean, standard deviation, skewness and excess kurtosis, and across all forty-five country pairs for the correlation.

The one-regime model without jumps is simply a multivariate normal model and naturally fails to produce the observed negative skewness and positive excess kurtosis. The one-regime model with jumps does detect return asymmetries and fat tails, but it strongly underestimates the negative skewness and kurtosis. The two-regime model without jumps fares even more poorly (despite a larger number of estimated parameters). Our two-regime model with jumps provides the best match for the negative skewness and kurtosis. These results highlight the importance of including jumps to properly model return moments. But the complexity of the international return distribution does not stop at the first few unconditional moments. All models seem to fit the observed unconditional correlation, but international correlation exhibits breaks during crises, an observation of great importance in asset allocation.

4.3 Number of regimes

In our model we assume that all of the ten equity markets share the same *global* regime. This assumption provides parsimony of our return-generating process and hence allows us to deal with a large number of markets.[20] This assumption is consistent with the two-country analysis of Ang and Bekaert (2002). They fail to reject the common regime model against the country-specific regime model. Similarly, Okimoto (2008) finds that the U.S. regime and the U.K. regime mostly coincide in his two-country model, and assumes the common regime classification in the estimation of his joint models.

To check whether our model needs more than two regimes, we fit a three-regime model without jumps. Using the test of Rivers and Vuong (2002), we can reject the three-regime model in favor of our model at 0.35% significance level. Based on the estimated model, one of the three regimes is not persistent with expected duration of only less than three weeks, and the markets are in this regime less than 2% of the time. We also find that this regime tries to capture a few outliers, and some of its associated parameters have large

20 If we allow each country to have two regimes of its own, the multivariate model would contain $2^{10} = 1,024$ distinct regimes. This extremely large number of regimes is not practical, and its marginal contribution would never outweight the computational burden and estimation error.

standard errors. Similar problems are found for a three-regime model with jumps. See also Ryden, Terasvirta, and Asbrink (1998), and Alexander and Lazar (2006) for similar findings. Therefore, we keep our model parsimonious with two regimes.

4.4 Country-specific and region-specific jumps

Our model allows systemic jumps of different sizes for different markets. This model can be generalized to include country-specific or region-specific jumps in a subset of countries. To check whether our two-regime model with *systemic* jumps further needs country-specific or region-specific jumps, we fit the two-regime model with (1) both systemic *and* ten country-specific jumps and (2) both systemic *and* three region-specific (Asia Pacific, Europe, North America) jumps. The tests fail to reject our model (with *only* systemic jumps) against each of the other two models at 10% significance level. In other words, the systemic jump component in our two-regime model is statistically sufficient to capture jumps among these ten countries.

4.5 Characteristics of global regimes and systemic jumps

Recall that returns as given by model (1) are the combinations of diffusion and jump components. These two components are estimated by the EM-based algorithm, and then the implied means, standard deviations, and correlations of returns, defined by (2) - (4), are computed from these quantities.[21] To characterize the regimes, we first look at return statistics as provided in Table 4. We find that returns in regime 1 have lower means, have higher volatilities, and are more correlated than those in regime 2. We choose to call regime 1 the “bad” or “bear” regime and regime 2 the “good” or “bull” regime. For all countries, mean returns are negative in the bad regime and positive in the good regime. The standard deviations of returns (volatility) are roughly twice as high in the bad regime as in the good regime. All cross-country correlations are higher in the bad regime. These findings tend to be consistent with the empirical literature in asymmetries of returns in equity markets. But two-regime models of return without jumps have a difficult time distinguishing between good and bad markets. For example, Ang and Bekaert (2002) fail to reject the equality of correlation across regimes at a 20% confidence level for their joint test. With two regimes and jumps, we find a better characterization of the regimes. All correlations are higher in the bad regime than in the good regime. A joint test over all correlation pairs rejects the null hypothesis of equal correlation in both regimes at a 0.02% confidence level. Similarly, we reject the null hypothesis of equal volatilities in both regimes at a 0.00% confidence level. However, any model faces the difficulty of estimating conditional mean returns as their standard errors are large.

[21] Detailed parameter estimates are available from the authors.

Table 4
Estimates of means, standard deviations, and correlations of weekly log-returns

A. Mean and SD (%)

	Mean ($\overline{\mu}$)			SD ($\overline{\sigma}$)		
	Regime 1 (bad regime)	Regime 2 (good regime)	*p*-value	Regime 1 (bad regime)	Regime 2 (good regime)	*p*-value
AU	−0.137 (0.399)	0.431 (0.135)	0.1843	5.251 (0.515)	2.721 (0.149)	0.0000
CA	−0.314 (0.380)	0.371 (0.116)	0.0883	5.025 (0.471)	2.358 (0.103)	0.0000
FR	−0.412 (0.380)	0.260 (0.133)	0.0996	5.049 (0.423)	2.712 (0.129)	0.0000
GE	−0.557 (0.426)	0.378 (0.131)	0.0382	5.650 (0.478)	2.678 (0.111)	0.0000
HK	−0.388 (0.310)	0.386 (0.113)	0.0198	4.173 (0.294)	2.313 (0.091)	0.0000
JP	−0.419 (0.244)	0.228 (0.121)	0.0191	3.262 (0.212)	2.505 (0.090)	0.0012
SP	−0.178 (0.389)	0.223 (0.167)	0.3510	5.143 (0.452)	3.369 (0.187)	0.0004
SW	−0.291 (0.323)	0.320 (0.098)	0.0731	4.271 (0.386)	2.008 (0.081)	0.0000
UK	−0.317 (0.351)	0.257 (0.104)	0.1197	4.610 (0.459)	2.110 (0.092)	0.0000
US	−0.295 (0.294)	0.232 (0.087)	0.0878	3.936 (0.304)	1.744 (0.082)	0.0000

B. Correlation ($\overline{\rho}$)

Regime 1 (bad regime)

	AU	CA	FR	GE	HK	JP	SP	SW	UK
CA	0.8555 (0.0290)								
FR	0.8313 (0.0320)	0.8584 (0.0273)							
GE	0.7945 (0.0385)	0.8211 (0.0339)	0.9475 (0.0103)						
HK	0.7559 (0.0410)	0.7378 (0.0432)	0.6990 (0.0486)	0.6859 (0.0500)					
JP	0.6704 (0.0509)	0.5974 (0.0603)	0.6023 (0.0589)	0.5612 (0.0637)	0.5942 (0.0569)				
SP	0.8196 (0.0346)	0.8009 (0.0380)	0.9083 (0.0175)	0.8991 (0.0191)	0.6728 (0.0522)	0.5693 (0.0631)			
SW	0.7834 (0.0418)	0.7767 (0.0420)	0.8856 (0.0219)	0.8528 (0.0287)	0.6733 (0.0518)	0.5604 (0.0636)	0.8722 (0.0248)		
UK	0.8479 (0.0316)	0.8663 (0.0277)	0.9240 (0.0149)	0.8860 (0.0224)	0.7089 (0.0484)	0.5641 (0.0647)	0.8576 (0.0281)	0.8614 (0.0284)	
US	0.7477 (0.0446)	0.8030 (0.0378)	0.8215 (0.0325)	0.8326 (0.0306)	0.6534 (0.0526)	0.5180 (0.0664)	0.7495 (0.0434)	0.7742 (0.0394)	0.8240 (0.0326)

Regime 2 (good regime)

	AU	CA	FR	GE	HK	JP	SP	SW	UK
CA	0.7271 (0.0284)								
FR	0.6735 (0.0350)	0.7212 (0.0289)							

(continued)

Table 4
Continued

	AU	CA	FR	GE	HK	JP	SP	SW	UK
	Regime 2 (good regime)								
GE	0.6473	0.7134	0.9383						
	(0.0345)	(0.0275)	(0.0066)						
HK	0.6019	0.4546	0.4909	0.4951					
	(0.0368)	(0.0439)	(0.0447)	(0.0431)					
JP	0.5424	0.4663	0.4427	0.4843	0.5139				
	(0.0373)	(0.0407)	(0.0426)	(0.0396)	(0.0377)				
SP	0.5485	0.5758	0.8596	0.8076	0.3481	0.3263			
	(0.0507)	(0.0447)	(0.0179)	(0.0220)	(0.0555)	(0.0491)			
SW	0.6634	0.6513	0.8420	0.8262	0.4885	0.4631	0.6939		
	(0.0320)	(0.0321)	(0.0165)	(0.0172)	(0.0417)	(0.0405)	(0.0325)		
UK	0.7246	0.7535	0.8737	0.8455	0.5338	0.4613	0.7229	0.8238	
	(0.0283)	(0.0245)	(0.0144)	(0.0161)	(0.0395)	(0.0408)	(0.0326)	(0.0175)	
US	0.5982	0.7515	0.7947	0.7914	0.4772	0.4413	0.6427	0.6692	0.7758
	(0.0407)	(0.0252)	(0.0236)	(0.0218)	(0.0449)	(0.0426)	(0.0408)	(0.0315)	(0.0232)

This table provides the estimated means, standard deviations (panel A), and correlations (panel B) given by Equations (2) - (4). All are estimated from the weekly log-returns using the developed EM-based algorithm. Standard errors are given in parentheses.

We find that all countries exhibit mean returns of opposite signs in the two regimes with large standard errors. We only reject the hypothesis of equal means against that of mean returns being lower in bad regimes at the 11% confidence level in a joint test. The difference is more significant for jump means. The mean jump size is lesser in the bad regime (average –2.8%) than in the good regime (average –1.0%), with a *p*–value of 4.71%. To alleviate this common problem of large standard errors in the mean, we will use Bayesian mean shrinkage estimators for the conditional means in our asset allocation study. Nevertheless, the introduction of jumps allows a better characterization of the risk parameters, correlation and volatility asymmetries. We will refine the discussion when we look at correlation breaks in Section 5.

Figure 1 shows the filtered (dashed line) and smoothed (solid line) probabilities of being in the bad regime. Table 5 shows that both regimes are persistent with a high probability of remaining in the same regime. The expected duration is 31 weeks for the bad regime (low return, high volatility, and high correlation) and 67 weeks for the good regime (high return, low volatility, and low correlation). The results also show that the arrival rate of jumps is significantly different from zero, thereby confirming the existence of systemic jumps in both regimes. Jump sizes are quite different in the bad and good regimes, with strong negative mean, high standard deviation, and high correlation in the bad regime.

The persistence of the regimes is an important feature. It shows that markets go through prolonged periods of good and bad regimes. The probability to stay in the same regime is high, and the expected regime duration is long. Our estimate that markets enter a regime gives useful information for the future and provides support for a regime-switching model. Jump sizes are

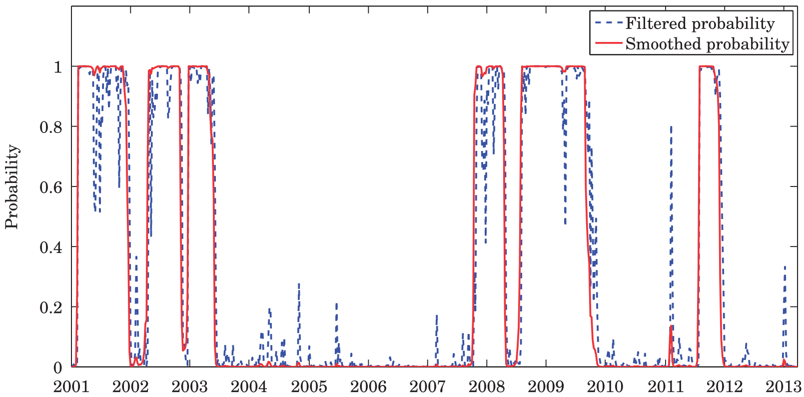

Figure 1
Probabilities of bad regime
This figure shows the filtered (dashed line) and smoothed (solid line) probabilities of being in the bad regime.

Table 5
Regime and jump statistics

	Prob of staying in the same regime (per week)	Expected duration (weeks)	Jump Arrival rate (per week)	Averages of jump size		
				Mean (%)	SD (%)	Corr
Bad regime	0.967	30.71	0.203	−2.829	6.446	0.903
	(0.015)	(13.98)	(0.040)	(1.270)	(0.925)	(0.037)
Good regime	0.985	66.47	0.187	−1.006	2.802	0.669
	(0.007)	(30.00)	(0.020)	(0.401)	(0.296)	(0.078)

This table provides the estimates of weekly transition probabilities, expected durations, jump arrival rates, and jump size statistics of both regimes obtained from the EM-based algorithm. Standard errors are given in the parentheses. The expected duration for a given regime is computed from the inverse of the rate at which the regime will shift to another regime.

quite different in the two regimes. A large negative jump can lead to a rapid transition to the bad regime, and the jump regime is persistent. In other words, a bad regime can be quickly detected, and this is useful information for asset allocation as the regime is persistent and all parameters, including international correlation, change drastically. This is confirmed by the out-of-sample results in Section 6.

4.6 Economic insights on regimes

Some indications of how well our model detects bear and bull markets can be gleaned in Figure 2A, which charts the periods of the bad regime and the returns on an equal-weight global portfolio. As stated above, mean returns are of opposite signs in bad and good regimes, but with rather low statistical power to differentiate. Nevertheless, we can see from the figure that the model quickly detects the start of a bad regime. This is true in all six bad periods detected in the model.

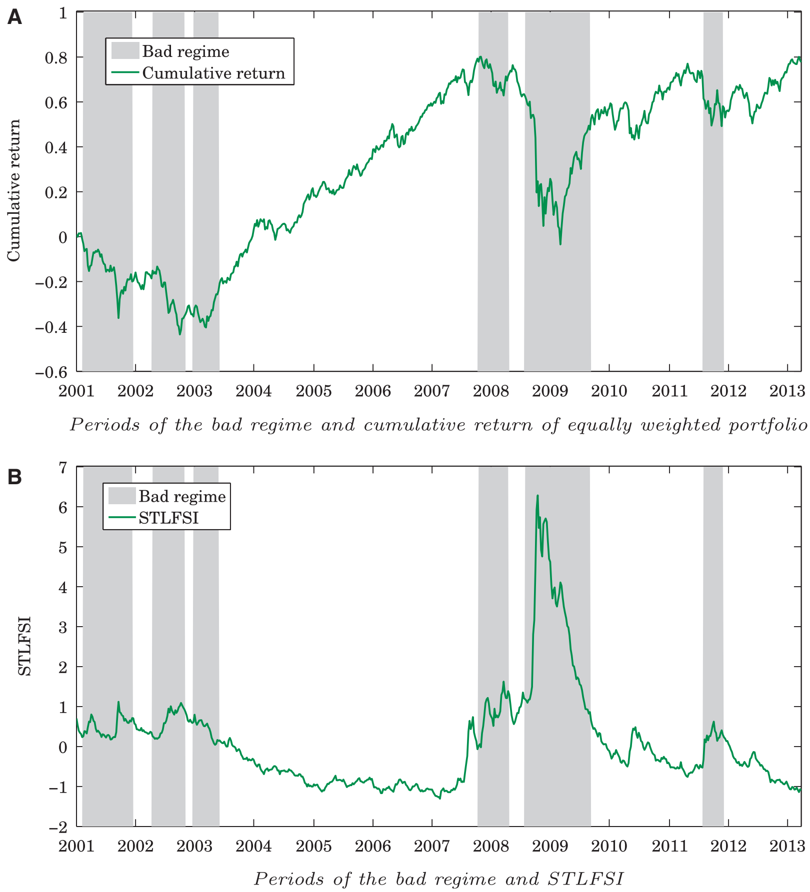

Figure 2
Global portfolio, STLFSI, and the bad regime periods
This figure shows the cumulative return of an equally weighted portfolio of the ten countries (*A*) and STLFSI (*B*) in solid lines. The bad regime periods are represented by the shaded areas.

Systemic jumps are a major determinant in our model. Large systemic shocks lead to high international market correlation and volatility. The increased correlation and volatility due to persistent big jumps allow for an early detection of bad regimes. One weakness of our model lays in its diminished ability to quickly detect the return to a bull market, such as after the 2003 and 2009 crises. At the start of a market recovery, observed returns, although generally positive, still exhibit a fairly large volatility and correlation for a while. It takes time for the model to confirm that observed returns are smoother and less correlated. On the other hand, large negative shocks easily signal a bad regime.

We now study whether the bad regimes are detected by some other early financial indicators. There are few high-frequency (weekly) indicators of financial crisis, and they are domestic ones, primarily from the United States. We focus on the STLFSI (Saint Louis Fed's Financial Stress Index), which is a broad weekly index of financial stress derived from a principal components analysis of eighteen financial indicators based on market prices. The VIX (Chicago Board Options Exchange Market Volatility Index) is a major component of STLFSI with a correlation of 0.9. Figure 2B charts the periods of the bad regime and STLFSI. Unsurprisingly, we can see that STLFSI is correlated with the regime (the regression R^2 is 0.45 and the slope coefficient is highly significant). But its indications tend to lag at the start of the bad regime. Our regime probability quickly shoots up at the start of crises, while STLFSI takes a longer time to reach a high value. To confirm, we run Granger causality tests between STLFSI and the filtered regime probability. We find that the regime probability explains the future values of STLFSI (*p*–value of 0.10%), but not the other way around. Thus, simply looking at implied volatility from option prices or from more comprehensive market indicators, such as STLFSI, provides less information than does our model with systemic jumps and international correlation breaks.

5. Asymmetries in Correlations

Conditional correlation received a lot of attention with the recent financial crises. A stylized fact about equity returns is that correlation is higher in bad markets than in good markets. In this section we explain how our model naturally captures the asymmetries in correlations and empirically show that our model indeed provides a much better fit than the models with a single regime, the models without jumps, or many other classes of standard multivariate models, including CCC and DCC GARCH models, asymmetric factor copula models, and multivariate factor stochastic volatility models.

5.1 Correlation breaks in bad regimes

In this section we investigate the source of correlation asymmetries; doing so, will be helpful in explaining why our model captures the asymmetries much better than the other models in the next section. Panel A of Table 6 provides the averages of correlations of returns ($\overline{\rho}$), diffusion (ρ^{Σ}), and jump sizes (ρ^{Ω}) over the forty-five country pairs for each regime. For the two-regime model with jumps, we find that the average return correlation increases from 0.64 in good markets to 0.76 in bad markets. A joint test indicates that we can reject the null hypothesis of no increase in correlation at the 0.01% significance level (one-sided test).

A major contribution of our paper is to introduce both diffusion and jump processes in the two regimes. As we estimate the parameters of the diffusion and jump processes in the two regimes, we can look at the relative contribution of each process to the increase in correlation in bad markets. As seen in

Table 6
Average correlations, means, and standard deviations of diffusion, jump sizes, and returns

	One regime without jumps	One Regime with jumps	Two regimes without jumps		Two regimes with jumps	
			Bad regime	Good regime	Bad regime	Good regime
A. Correlation						
Diffusion (ρ^{Σ})	0.709	0.629	0.745	0.631	0.668	0.636
Jump size (ρ^{Ω})		0.813			0.903	0.669
Return ($\overline{\rho}$)	0.709	0.711	0.745	0.631	0.763	0.638
B. Mean (%)						
Diffusion (μ)	0.107	0.394	−0.561	0.414	0.243	0.497
Jump size (η)		−1.369			−2.829	−1.006
Return ($\overline{\mu}$)	0.107	0.107	−0.561	0.414	−0.331	0.309
C. Standard deviation (%)						
Diffusion (σ^{Σ})	3.305	2.325	4.832	2.187	3.339	2.055
Jump size (σ^{Ω})		4.911			6.446	2.802
Return ($\overline{\sigma}$)	3.305	3.318	4.832	2.187	4.637	2.452

This table provides the averages of the estimated correlations (panel A), means (panel B), and standard deviations (panel C) of diffusion components, jump sizes, and log-returns for each of the four models: one regime without jumps, one regime with jumps, two regimes without jumps, and two regimes with jumps. The averages are provided for each regime for the two-regime models.

panel A of Table 6, the correlation between diffusion does not increase much *between the good and bad regimes*; the average correlations of diffusion are 0.64 in good markets and 0.67 in bad markets. On the other hand, the average correlation of jump sizes increases from 0.67 in the good regime to 0.90 in the bad regime. The jump sizes are highly correlated in the bad regime, with the correlations above 0.80 for each of the forty-five country pairs. Thus, the correlation asymmetry of returns between the two regimes is caused by the jump component, and not by the diffusion component. We can also see from panels B and C of Table 6 that diffusion provides higher means and lower volatility in the good regime. Both regimes are subject to jump risk, but jumps have a much larger negative expected size and higher correlation and volatility in the bad regime. So jumps in bad regimes bring large negative returns to all markets with high correlation.

Now we focus on the difference of average correlations between the diffusion component and jump component *within a given regime*. In the good regime, the average correlations of diffusion and jump components are similar and equal to 0.64 and 0.67, respectively. On the contrary, we find that the average correlation of diffusion (0.67) is much lower than that of jumps (0.90) in the bad regime. We refer to this drastic increase in correlation during the bad regime conditional on a jump arrival as *correlation breaks*. To further illustrate, let's take the case of the Hong Kong market. Systemic jumps are common to all countries but infrequent and have a random jump size for each country. In the good regime, the correlation of jump sizes between Hong Kong and the other countries is rather small (on average 0.47), providing good risk diversification benefits and illustrating the Chinese-diversification property of Hong Kong. But in the bad regime, systemic jumps affect all countries in a similar fashion and jump sizes are highly correlated (on average 0.92). On the other hand the correlations of the diffusion components are around 0.5 in both regimes. In bad markets, systemic jumps lead to a correlation break for Hong Kong, making it less attractive as a diversification investment. These results also intuitively explain why introducing systemic jumps allows for a better characterization of the return process compared with a two-regime model without jumps.

5.2 Correlation of extreme returns

Using extreme-value modeling, Longin and Solnik (2001) find that correlations conditional on returns that are lower than a negative threshold increase as the threshold becomes more negative, and those conditional on returns that are higher than a positive threshold decrease somewhat as the threshold becomes more positive. This asymmetric exceedance correlation is also documented by Ang and Bekaert (2002), Ang and Chen (2002), and Okimoto (2008). The conclusion that correlation is higher in bad markets and increases for extreme negative returns is very important in risk management and asset allocation. With nonnormal distributions, portfolio risk management has to focus on what happens in periods of extreme negative returns. As an anecdote, let's consider LTCM, which was badly hit in the 1998 crisis because its extensive diversification across countries and markets failed to provide benefits due to the drastic rise in correlation. A similar scenario occurred in the 2008-2009 crisis.

Models for risk management and asset allocation optimization should be able to deal with extreme returns and should lead to useful asset allocation implications for periods of crisis. We now focus our discussion on correlations during the extreme bad markets and will turn to asset allocation implications in Section 6.

We first investigate how important regime switching and jumps are for reproducing the exceedance correlation structure observed in the data, especially for extreme negative returns. We estimate four different models: one-regime model without jumps, one-regime model with jumps, two-regime model without jumps, and two-regime model with jumps. We compute the exceedance correlations implied by the models. If jumps alone can capture the increase in the correlation during bad markets, the one-regime model with jumps should provide a good fit. Likewise, if the increase in the correlation of the diffusion in bad regimes is the only source of the correlation asymmetries, the two-regime model without jumps should provide a good fit. As we will see

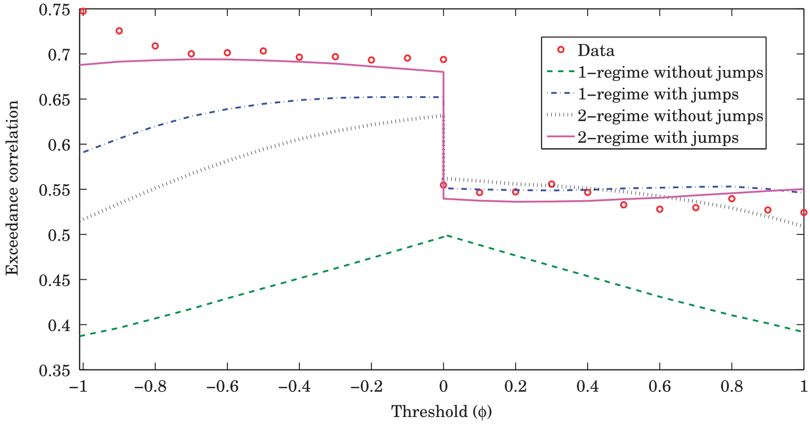

Figure 3
Exceedance correlations
This figure shows the average exceedance correlations from forty-five country pairs. It contains the average exceedance correlations computed from observed data (circles), one-regime model without jumps (dashed line), one-regime model with jumps (dot-dashed line), two-regime model without jumps (dotted line), and two-regime model with jumps (solid line).

below, neither is true. Every model is estimated based on the EM-based algorithm, which maximizes the likelihood function, so there is no guarantee that models with a larger number of parameters will better reproduce the observed exceedance correlations.

Let's denote returns r_1 on market of country 1 and r_2 on market of country 2 with unconditional means $\overline{r}_1$ and $\overline{r}_2$ and unconditional standard deviations s_1 and s_2. The normalized return of country i is defined by $\tilde{r}_i = (r_i - \overline{r}_i)/s_i$. We compute the exceedance correlation for a positive threshold ϕ as $Corr(\tilde{r}_1, \tilde{r}_2 \,|\, \tilde{r}_1 > \phi, \tilde{r}_2 > \phi)$ and for a negative threshold ϕ as $Corr(\tilde{r}_1, \tilde{r}_2 \,|\, \tilde{r}_1 < \phi, \tilde{r}_2 < \phi)$. For each model, we simulate 500,000 returns and compute the exceedance correlations. The results for the average of all country pairs are presented in Figure 3, which plots the average exceedance correlation computed from the actual data, as well as those estimated from the four models. Note that the actual exceedance correlation data is not smooth when we increase the magnitude of the threshold because we have few data points once we consider extreme returns. We limit the calculation to have at least 20 data points for every country pair.

To understand this figure, let's consider the case of a multivariate normal distribution (dashed line). As stressed in Galambos (1978) and Longin and Solnik (2001), the conditional correlation of a multivariate normal distribution decreases with the threshold and reaches zero for extreme returns. In Figure 3, the correlation of the multivariate normal process conditional on both returns being below (or above) their means is equal to the exceedance correlation measured at a zero threshold (e.g., 0.50 for the average exceedance correlation) and the conditional exceedance correlation decreases as the absolute size of the threshold increases, hence the symmetric inverted V shape. The conditional correlation goes asymptotically to zero for extreme returns.

We focus the discussion on negative exceedances as risk lies in extreme negative returns, not in positive returns. On average, the two-regime model with jumps does a good job of capturing the exceedance correlation for negative thresholds observed in the data in both direction and magnitude. On the other hand, the other three models do a poor job of capturing the observed exceedance correlations for negative thresholds; this failure increases as the thresholds become more negative (extreme negative returns). All models (except the one-regime model without jumps) yield fairly similar results for positive exceedances.

Now we provide an intuitive explanation for the success of our two-regime model with jumps in capturing asymmetries in extreme correlations. As mentioned above, correlation breaks occur during bad markets and are associated with jumps. Observing large negative returns, investors would form a strong belief that markets are in the bad regime and those returns come from jumps. Because jumps in the bad regime have the highest average correlation of 0.90 (panel A of Table 6), when returns becomes more negative, the conditional correlation increases, generating increasing negative exceedance correlation as observed in the data. However, when ignoring jumps in the two-regime model, one has to rely on the less-correlated diffusion in the bad regime, for which the average correlation is only 0.75. Similarly, if only jumps, but not regime-switching, are accounted for, less-correlated good-regime jumps and more-correlated bad-regime jumps are mixed into one jump component, reducing the average jump correlation to only 0.71. The values for the latter two cases are too low to match the observed correlation breaks during bad markets.

Overall, jumps or regime switching alone is not enough to model asymmetric exceedance correlations observed among these ten countries during the past decade, while the model with both jumps and regime switching significantly improves the fit of the observed extreme correlation, particularly for the negative thresholds. As stressed before, risk is really about negative returns, and our model seems to do better than the others for negative returns. Indeed, we show in Section 6 that our richer modeling leads to superior *out-of-sample* portfolio performance.

5.3 Comparison with other benchmarks

In this section we conduct an extensive empirical study to investigate how well other standard volatility models in the finance literature can capture the observed exceedance correlations of extreme returns compared with our model. We consider three main classes of multivariate models: multivariate GARCH, factor copula, and multivariate factor stochastic volatility. As in the previous section, each model is fitted to the data by its standard estimation method (e.g., maximum likelihood estimation for GARCH models), rather than by trying to minimize the exceedance correlation fitting errors. This allows us to understand how much the asymmetries of extreme correlations implied by the return data each of these benchmark models can capture. In Figure 4, we provide the plots of exceedance correlations implied from data (circles) and from our two-regime model with jumps (solid line) compared with those implied from the benchmark models. Internet Appendix D provides the details of each benchmark model.

We first consider the standard multivariate GARCH models with constant conditional correlation (CCC) of Bollerslev (1990) and dynamic conditional correlation (DCC) of Engle (2002). The returns of each country are assumed to follow a univariate GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993) with normal residuals. Figure 4A shows the exceedance correlation plots of the CCC and DCC models. Both CCC and DCC models generate symmetric inverted V-shaped plots similar to the multivariate normal model and significantly underestimate the exccedance correlations for both positive and negative returns. This emphasizes the undesirable conditionally normal joint distribution of these GARCH models.

Next, we consider the factor copula models of Oh and Patton (2012) in which the joint distribution is constructed using the copula implied from a factor model. Following Oh and Patton (2012), we consider three variants of factor models: one-factor model with the same factor loading, one-factor model with different factor loadings, and four-factor model with one global factor, and three regional factors (Asia-Pacific, Europe, and North America) and assume that factors have skewed-t distribution, which allows asymmetries between upturn and downturn movements, and correlated crashes. The implied exeedance correlations from the factor copula models are shown in Figure 4B. Although all of the models generate asymmetries between positive and negative exceedance correlations, they poorly match the observed ones. The implied exceedance correlations for negative returns decrease as the threshold becomes more negative, sharing the same undesirable property as the GARCH models, consistent with what the authors mentioned in their paper: “The combination of time-varying conditional means and variance and a constant conditional copula makes this model similar in spirit to the CCC model of Bollerslev (1990).” So introducing asymmetries in the factor copula does not seem to be enough to capture the observed exceedance correlations. Furthermore, the use of additional factors that are region/country specific does not seem to improve the results. This confirms our findings about adding region-/country-specific jumps in our model.

Third, we consider the multivariate factor stochastic volatility models of Omori and Ishihara (2012) in which returns are driven by a set of factors and a vector of innovations with multivariate-t errors, and each of the factors and innovations follows a univariate stochastic volatility model. Figure 4C shows the implied exceedance correlations of models with one, three, and five

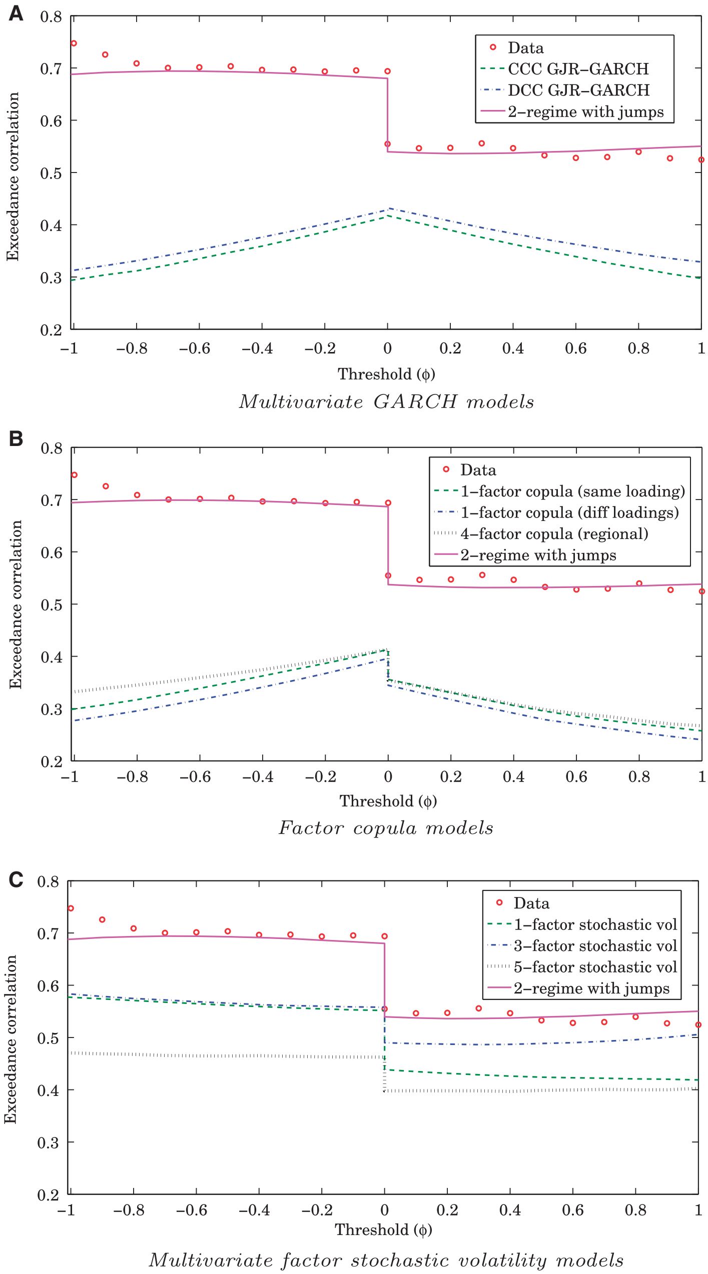

Figure 4
Exceedance correlations of GARCH, factor copula, and factor stochastic volatility models
This figure shows the average exceedance correlations from forty-five country pairs implied from data, our two-regime model with jumps, and the benchmark models: GARCH (*A*), factor copula (*B*), and factor stochastic volatility (*C*).

factors. The models generate asymmetries of exceedance correlations for positive and negative returns. Unlike the multivariate GARCH and factor copula models, the exceedance correlations increase as the threshold becomes more positive or more negative. However, all of the multivariate factor stochastic volatility models underestimate the exceedance correlations for both positive and negative returns.

Exceedance correlation measures the dependency of the extreme returns based on conditional correlations. It depends on the choice of marginal and joint distributions of returns. Quantile dependence is an alternative dependence measure based on conditional probabilities. It depends only on the copula, but not on the marginal distribution. We provide results in Internet Appendix E. They lead to conclusions similar to those obtained with exceedance correlation.

Modeling and fitting multivariate models is a very challenging task. Most of multivariate models impose some structures or rely on factor models to make estimation possible, while having to sacrifice their flexibility. As we have seen, when certain structures are imposed, the multivariate GARCH, factor copula, and multivariate factor stochastic volatility models fail to capture the observed exceedance correlations and quantile dependence of returns. On the other hand, our two-regime model with jumps captures the observed values much better than do those models. At the same time, using the powerful EM-based estimation method we have derived, we find estimating our model is easy because our model relies on simple distributions, such as normal and Poisson. The results of this section clearly show that our model, which requires less compromise between estimation feasibility and model flexibility, provides a better solution for modeling and capturing extreme dependency in international equity markets than have previous models.

6. Asset Allocation Implications

In this section we formulate and solve *dynamic* portfolio choice problems when returns are subject to correlation asymmetries and breaks modeled by regime-switching and jumps. We assume that the underlying market regimes are *unobservable*, and investors dynamically update their beliefs of the market regimes based on the observed asset return history. We compare our optimal portfolio weights and those of the investors who ignore regime switching and/or jumps. We also test our portfolio model *out-of-sample* against various models, including the naïve $1/N$ model. The results show that our model outperforms the other models *out-of-sample*, especially during crisis periods.

6.1 Optimal portfolio choice

This section considers an optimal portfolio choice problem for an agent with constant relative risk aversion. Consider a U.S. investor who invests in the international equity markets for which the log-returns follow process (1) and a risk-free asset with continuously compounded rate of return r_f. Let $x_{i,t}$ denote the portfolio weight in market i at time t for $i = 1, \ldots, n$ and $t = 0, 1, \ldots, T-1$. The wealth process W_t satisfies:

$$W_{t+1} = W_t\left(e^{r_f} + x_t'(e^{R_{t+1}} - e^{r_f}\mathbf{1})\right), \tag{7}$$

where $x_t = [x_{1,t}, \ldots, x_{n,t}]'$ denotes the vector of portfolio weights, $R_t = [R_{1,t}, \ldots, R_{n,t}]'$ is the vector of market returns with $e^{R_t} \equiv [e^{R_{1,t}}, \ldots, e^{R_{n,t}}]'$, and $\mathbf{1}$ is the vector of ones. We assume that the objective is to maximize the investor's expected power utility of terminal wealth at time T:

$$\max_x \mathbb{E}[U(W_T)] \text{ where } U(W) = \begin{cases} \dfrac{W^{1-\gamma}}{1-\gamma} & W > 0 \\ -\infty & W \leq 0 \end{cases}$$

and $\gamma > 0$ is the relative risk aversion coefficient. For $\gamma = 1$, the utility is logarithmic. We assume further that the investor's information set at time t contains the history of asset prices and wealth, but *not* the market regimes. That is, the regime Y_t is *unobservable*. The investor uses the observed historical returns to derive a belief about the likelihood of the market regime $q_t = [q_{1,t}, \ldots, q_{K,t}]'$, where

$$q_{y,t} = P(Y_t = y \mid R_1, \ldots, R_t)$$

is the probability that the current regime is y conditional on all historical market returns. The investor updates belief q_{t+1} on observing new market returns r_{t+1} using the Bayes' rule. The following theorem provides the optimal weights, for which the derivation is given in Internet Appendix F.

Theorem 1. The optimal portfolio weight at time t when the regime probability vector is $q = [q_1 \ldots, q_K]'$ is the maximizer of the following problem:

$$x^*(t, q) = \arg\max_x \sum_{z=1}^{K}\sum_{y=1}^{K} q_z p_{z,y} \mathbb{E}\left[\frac{\left(e^{r_f} + x'\left(e^{R_{t+1}} - e^{r_f}1\right)\right)^{1-\gamma}}{1-\gamma} \times h(t+1, q_{t+1}(q, R_{t+1})) \mid Y_{t+1} = y\right], \tag{8}$$

where the functions $q_{t+1}(q, r)$ and $h(t, q)$ are given by (F.1) and (F.5) in Internet Appendix F, respectively.

Optimal portfolio weights obtained from Equation (8) are used to study the diversification effects of asymmetries in systemic jumps on the portfolio choice problem. Observe that the optimal portfolio weights depend on time t and the probability q, but not on wealth W.

6.2 Optimal portfolio weights

We first illustrate optimal portfolio weights in-sample. A well-known fact is that simple estimates of mean returns are subject to relatively large estimation errors and are poor estimators of expected returns. Our contribution is on risk modeling, not on better models of expected returns. Hence, we constrain the means to be equal for each country. In this section, we simply illustrate how different risk models would affect portfolio decisions in the presence of crises, and this *constrained-mean* model serves our purposes. When studying performance out-of-sample, we introduce a Bayesian shrinkage estimator for expected returns. We restrict our study to the case in which short-selling is not allowed. This is consistent with most mutual funds that are not allowed to short sell, but are allowed to borrow (Almazan et al., 2004). This restriction rules out some hedge fund strategies arbitraging across markets, but these strategies are primarily based on country market valuation (expected return), which is not our focus.

The optimal portfolio weights of models with one regime are constant. For the two-regime models, the portfolio weights, conditional on a bad-regime probability q, depend on the investment horizon. We find that those weights are almost identical for horizons over a year, so we select one year as the investment horizon. The optimal weights obtained from the four models for different probabilities of bad regime (q) are provided in Table 7 for a risk aversion coefficient of five.[22] The weights of the ten countries are grouped and reported by regions: Asia-Pacific (AU, HK, and JP), Europe (FR, GE, SP, SW, and UK), and North America (CA and US). These regional weights are the weights within the risky-asset portfolio and hence sum to one. These weights allow us to study diversification across regions, while the total risky weights and the risk-free weights allow us to study the leverage positions. We use the Federal funds rate as the risk-free rate. The top panel of Table 7 reports portfolio weights of the models without jumps, and the bottom panel reports the weights of the models with jumps. Figure 5 provides the same regional weights in a more detailed fashion.

Let's first look at the portfolio weights of the models with **one regime**, without jumps (mean-variance) and with jumps. Interestingly, modeling jumps leads investors to take larger positions in the risky assets. The risky weight of the one-regime model with jumps is 58.7%, while the risky weight of the one-regime model without jump is only 39.3%. This seems to suggest that precise risk modeling allows investors to take more aggressive positions for a similar risk level. The regional weights of the equity portfolio composition differ slightly from around 3% – 8%.

Next, we compare the **two-regime** models without and with jumps. Unlike the one-regime models, the two-regime models suggest investors change the

[22] Results for other levels of risk aversion are available from the authors. The composition of the risky portfolios does not change much for different levels of risk aversion, but the amount of leverage does as expected.

Table 7
Optimal portfolio weights

	One regime without jumps	Two regimes without jumps										
		$q=0$	$q=0.1$	$q=0.2$	$q=0.3$	$q=0.4$	$q=0.5$	$q=0.6$	$q=0.7$	$q=0.8$	$q=0.9$	$q=1$
Risky-asset port ($x/x'\mathbf{1}$)												
Asia-Pacific	0.560	0.362	0.442	0.506	0.570	0.317	0.000	–	–	–	–	–
Europe	0.182	0.176	0.150	0.143	0.220	0.684	1.000	–	–	–	–	–
North America	0.258	0.462	0.408	0.351	0.211	0.000	0.000	–	–	–	–	–
Total risky wgt ($x'\mathbf{1}$)	0.393	1.739	1.103	0.670	0.329	0.093	0.012	0.000	0.000	0.000	0.000	0.000
Risk-free wgt ($1-x'\mathbf{1}$)	0.607	−0.739	−0.103	0.330	0.671	0.907	0.988	1.000	1.000	1.000	1.000	1.000
	One regime with jumps	**Two regimes with jumps**										
		$q=0$	$q=0.1$	$q=0.2$	$q=0.3$	$q=0.4$	$q=0.5$	$q=0.6$	$q=0.7$	$q=0.8$	$q=0.9$	$q=1$
Risky-Asset port ($x/x'\mathbf{1}$)												
Asia-Pacific	0.522	0.368	0.450	0.502	0.549	0.606	0.511	0.329	0.000	–	–	–
Europe	0.155	0.180	0.160	0.155	0.172	0.214	0.489	0.671	1.000	–	–	–
North America	0.322	0.452	0.391	0.343	0.279	0.180	0.000	0.000	0.000	–	–	–
Total risky wgt ($x'\mathbf{1}$)	0.587	2.419	1.463	0.952	0.599	0.331	0.138	0.052	0.001	0.000	0.000	0.000
Risk-free wgt ($1-x'\mathbf{1}$)	0.413	−1.419	−0.463	0.048	0.401	0.669	0.862	0.948	0.999	1.000	1.000	1.000

This table provides optimal portfolio weights under four different models: one regime without jumps, one regime with jumps, two regimes without jumps, and two regimes with jumps. The investment universe consists of the ten country indices and the risk-free asset. The weights are presented by regions (Asia-Pacific, Europe, and North America). These regional weights are weights within the risky-asset portfolio and hence sum to one. The total weight for the risky-asset portfolio (Total risky wgt) and the risk-free weights (Risk-free wgt) are also provided. Means of diffusion and jumps are constrained to be equal across countries for each regime. The investment horizon is one year, and the relative risk aversion coefficient (γ) is 5. The weights are provided for the probability of starting in the bad regime (q) of 0, 0.1, …, 1. Shortselling is not allowed.

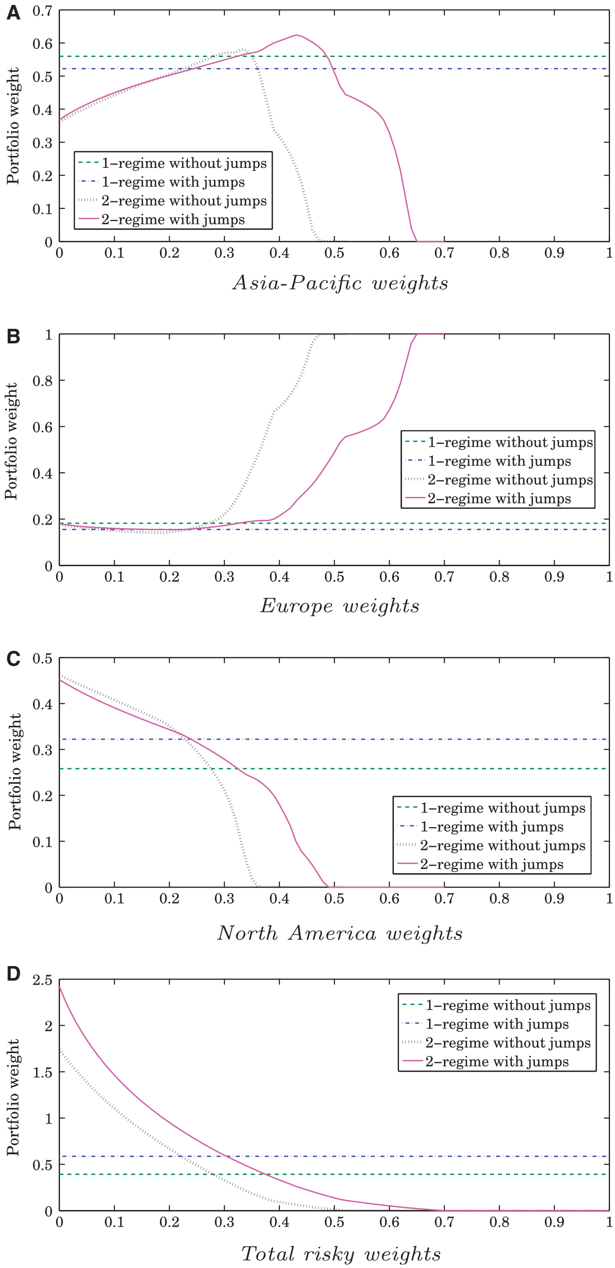

Figure 5
Optimal portfolio weights
This figure shows the optimal weights of the constrained-mean portfolios as functions of the bad regime probability of the four models: one-regime without jumps (dashed line), one-regime with jumps (dot-dashed line), two-regime without jumps (dotted line), and two-regime with jumps (solid line). The weights in the Asia-Pacific, Europe, and North America regions are provided in (*A*), (*B*), and (*C*), respectively. They are the weights in the risky portfolio. The total risky weights are in (*D*). The risk aversion coefficient (γ) is 5 and the investment horizon is one year.

compositions within the risky portfolio, as well as the leverage position based on the current probability of the bad market regime (q). We start with a summary of the asset allocation implications of our model with jumps. An investor, who is certain that the current regime is good ($q = 0$), will hold a leveraged position in equity. The equity allocation is 46% to America, 18% to Europe, and 36% to Asia-Pacific. As q increases, the investment in the risky assets drops, and Asia-Pacific replaces America, while Europe is stable within the risky portfolio. Investors keep holding risky assets until $q = 0.7$. As the expected return decreases with the higher probability of a crisis, risk focus becomes more important. That is achieved both by a reduced leverage and a higher allocation to the region providing the best diversification benefit (lower correlation). But when the probability of a bad regime looms large (say q over 0.4), the proportion of the risky assets gets smaller and a correlation break becomes more likely. This leads to a drastic reduction in the Asia-Pacific allocation as the correlation break is most pronounced for this region. The allocation to Europe goes up as some European countries are less sensitive to correlation breaks, while the America allocation keeps dropping. For $q = 0.5$, the portfolio only has a 14% investment in risky assets with roughly equal weights between Asia-Pacific and Europe. For higher q's, the weight of risky assets is slowly reduced, with a concentration in Europe until $q = 0.7$, where the portfolio becomes fully invested in the risk-free asset. In comparison, for the two-regime model without jumps, investors take less leverage for all probabilities q and the model without jumps stops investing in equity for q around 0.5 (compared to 0.7 for our model). The equity allocations of both models are fairly similar when there is a large probability of a good regime (low q), but quite different when q is above 0.3.

This suggests that correlation asymmetries between the good and bad regimes have substantial impacts on the composition within the equity portfolio. Improved risk modeling (including jumps) allows a better differentiation between the regimes. It also allows for taking more aggressive positions for a similar perceived risk level.

6.3 Out-of-sample tests

Our model with regime switching and jumps has a large number of parameters that may be subject to overfitting and estimation risk, so testing the model out-of-sample is important. DeMiguel, Martin-Utrera, and Nogales (2009) analyze out-of-sample portfolio performance of numerous mean-variance models, with and without taking into account estimation error, against the naïve $1/N$ model.[23] They find that none of the models consistently outperforms the $1/N$ model and conclude that the gain from optimal diversification is more than offset by estimation error. In this section our model is tested

[23] The $1/N$ model is an equally weighted portfolio in which each risky asset has the same portfolio weight and no weight is given to the risk-free asset.

Table 8
Out-of-sample portfolio performance

	Mean	SD	Sharpe ratio	p–value
$1/N$	2.24%	25.00%	0.090	0.1094
One regime without jumps	–8.59%	20.28%	–0.424	0.0092
One regime with jumps	–10.15%	27.16%	–0.374	0.0126
Two regimes without jumps	1.74%	18.87%	0.092	0.0474
Two regimes with jumps	9.82%	18.71%	0.525	

This table provides out-of-sample portfolio performance under five different models: $1/N$, one regime without jumps, one regime with jumps, two regimes without jumps, and two regimes with jumps. The results are based on the shrinkage-mean models. The entire sample data are from January 2001 to December 2013, and the out-of-sample data are from January 2008 to December 2013. The models are refitted at the beginning of each year. The trading frequency is weekly. The performances are measured by the annualized mean of excess log-returns (Mean), the annualized standard deviation of excess log-returns (SD), and the annualized Sharpe ratio (Sharpe ratio). The last column (p–value) provides the one-sided p–value for the null hypothesis of equal Sharpe ratio between a given model and the two-regime model with jumps, against the alternative hypothesis of higher Sharpe ratio for the two-regime model with jumps. The p–values are computed based on the method of Ledoit and Wolf (2008). The risk aversion coefficient (γ) is 5, and no shortselling is allowed.

out-of-sample against simpler models, including the robust $1/N$ model. Our out-of-sample period runs from January 2008 to December 2013, a total of six years. This period covers various market conditions, including the crisis in 2008, global stock rally in 2009, and other up and down years from 2010 to 2013.

As both expected returns and risk play an important role in portfolio performance, we take care of the estimation errors in the mean parameters using the Bayesian shrinkage method of DeMiguel, Martin-Utrera, and Nogales (2013) in which the estimate of the mean return of each asset is shrunk toward its grand mean (average of the means across assets). They show that out-of-sample Sharpe ratios for various mean-variance portfolios are improved under their shrinkage method. We apply their method to each mean parameter (mean of diffusion and mean of jump size) in each regime and refer to this case as a *shrinkage-mean* model. Our analyses in this section rely on the shrinkage-mean models, but the results are qualitatively similar for constrained-mean and unconstrained-mean models.

At the beginning of each year, the investors are allowed to re-estimate their models based on the data from 2001, and with the new estimates, they solve dynamic portfolio optimization problems to obtain the optimal weights, which are functions of the market regime probability and the investment horizon.

Table 8 provides the portfolio performance of five models: $1/N$, one regime without jumps, one regime with jumps, two regimes without jumps, and two regimes with jumps. We report the annualized mean of excess returns (Mean), the annualized standard deviation of excess returns (SD), and the annualized Sharpe ratio (Sharpe ratio). The values in Table 8 are based on $\gamma = 5$. The results for $\gamma = 3$ and 10 are similar and are not reported here.

We find that our model with regime switching and jumps strongly outperforms the other four models in many aspects. That is, it provides the highest

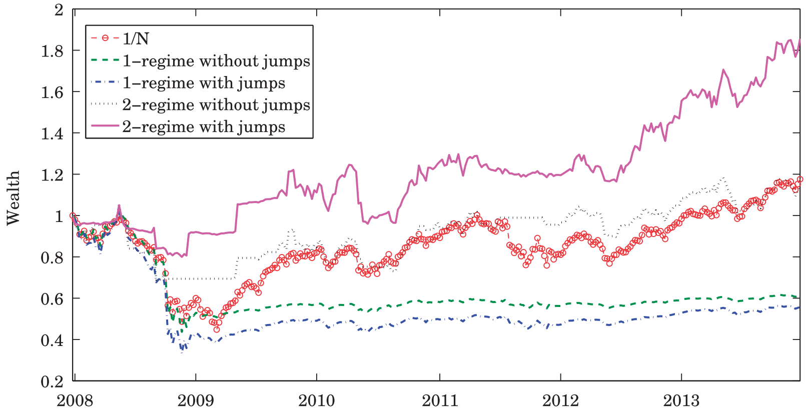

Figure 6
Out-of-sample wealth processes
This figure shows the out-of-sample wealth processes corresponding to five different trading models: $1/N$ (circles with dashed line), one-regime model without jumps (dashed line), one-regime model with jumps (dot-dashed line), two-regime model without jumps (dotted line), and two-regime model with jumps (solid line). The results are based on the shrinkage-mean models. The entire sample data are from January 2001 to December 2013, and the out-of-sample data are from January 2008 to December 2013. The models are refitted at the beginning of each year. The trading frequency is weekly. The risk aversion coefficient (γ) is 5, and no shortselling is allowed.

excess return (9.82% per year), the lowest standard deviation (18.71% per year), and the highest Sharpe ratio (0.525). The one-regime models perform poorly with negative excess returns and higher standard deviations. The two-regime model without jumps has a Sharpe ratio comparable to the $1/N$ strategy (0.092 vs. 0.090). Interestingly, our model strongly outperforms the $1/N$ strategy with much higher excess return (9.82% vs. 2.24%), lower standard deviation (18.71% vs. 25.00%) and, thus, a much higher Sharpe ratio (0.525 vs. 0.090). We further test the statistical significance of the differences in the Sharpe ratios of each model compared with our model. We use the test of Ledoit and Wolf (2008) that requires no assumptions of normal distribution and i.i.d. property of returns, to test the null hypothesis of equal Sharpe ratio against the alternative hypothesis that our model has a higher Sharpe ratio.[24] The one-sided p–values of the tests against our model are given in the last column of Table 8. As we can see, we can reject the null hypothesis for both one-regime models with p–values around 1%. The p–value for the two-regime model without jumps is below 5%. The p–value for the $1/N$ strategy is around 10%. The out-of-sample period is only six years, so it is not surprising that the testing power is a bit limited, even if the difference in point estimates is large in economic terms.

Now we investigate in more detail how our model provides superior out-of-sample performance. Figure 6 shows the wealth processes during the out-of-

[24] This test relies on fewer distribution assumptions than does the test of Jobson and Korkie (1981) with correction by Memmel (2003), or the JKM test, used in DeMiguel, Garlappi, and Uppal (2009). We also used the JKM test, but results are quite similar on our data and not reported here.

sample period for each model. We end up with a total wealth increase of slightly more than 85% over the six-year period. The $1/N$ strategy and the two-regime model without jumps end up with slightly less than 18% wealth increase. This is a large financial difference, and our model also has lower volatility. The difference mostly comes from the early detection of crisis. Our model is able to recognize the bad regimes and suggests increasing money into the risk-free asset and altering the allocation within the remaining equity portfolio as the probability of a correlation break rises. Interestingly, after observing the start of the 2008 global market crash in August, detected in our model by the high correlation of jumps in bad regimes, investors relying on our model are very confident that the markets have entered into the bad regime, and hence put all or most of their wealth in the risk-free asset, freezing the loss from the beginning of the year to about 20%. On the other hand, investors relying on the other models still invest in the risky assets, and face much more loss during 2008. For those models, the maximum losses range from 31% to 67%. A similar situation happens in the crises of the first quarter of 2009 or of late 2011.

As mentioned above, our model may have more difficulty detecting the return to a good regime, characterized by a succession of positive diffusion returns. This appears after April 2009. As the good regimes are characterized by jumps of smaller magnitude, the return to the good regime is harder to detect. But once confirmed, investors invest more in risky assets, yielding a higher growth (e.g., from mid-2012 to late 2013). The contribution is not only a different allocation between risk-free and risky assets. The composition of the equity portfolio also changes with the bad-regime probability. As the international correlation structure drastically changes in bad regimes, due to breaks in jump correlation, the asset allocation adopts an effective diversification strategy. Our out-of-sample period is representative of periods with booming markets and crises, albeit being somewhat short, and clearly illustrates the benefits of our risk modeling.

7. Conclusion

There is extensive evidence on asymmetries in international equity returns. In particular, extreme negative returns are of much larger magnitude than are extreme positive returns and tend to be much more correlated across countries. At the beginning of crises, return correlations and volatility increase rapidly, and large negative returns are observed repeatedly during the crisis periods. This cannot be captured by a smooth return distribution (such as normal or skewed-t) with slowly changing volatility. Discontinuous shocks and sudden shifts in parameters have to be considered. We propose a model that comprises diffusion *and* jump processes with regime switching and develop a powerful application of the EM algorithm that allows a multivariate estimation for a large number of countries. Our model yields results that can replicate the stylized facts, while classes of the models we reviewed cannot (such as regime-switching without jumps, DCC GARCH, factor copula, and multivariate factor stochastic volatility).

Our model uses simple distributions (normal and Poisson) with well-known properties. We model systemic jumps for which country jump sizes are random but correlated. We find evidence for bad/crisis regimes (high volatility, high correlation and negative mean return) and good regimes (low volatility, low correlation and positive mean return), and evidence for periodic jumps (extreme returns or shocks), especially in the bad regime. Bad regimes are characterized by large negative shocks with high correlation. Regimes are persistent and can be detected early; when we observe large negative returns across all markets, persistent bad regimes can be detected early.

Bad regimes display a big increase in international correlation ("correlation break"). This is caused by the much higher correlation of negative jumps, while the correlation of the diffusion component of returns remains stable. For example, the average jump correlation for Hong Kong increases from 0.47 in good markets to 0.92 in bad markets, reducing the diversification benefit when it is most needed. Correlation is crucial in asset management. Asset managers rely on diversification, but diversification can fail during a crisis because of a break in market correlation. The LTCM's failure taught us that investments that appeared well diversified, due to low correlation during good times, could become extremely risky when volatility and correlation shoot up during extreme bad markets.

Our findings have important asset allocation implications. We solve dynamic optimal portfolio by assuming that the underlying market regimes are unobservable, and investors dynamically update their beliefs about the market regimes based on the observed asset return history. We conduct out-of-sample performance tests and show that our model outperforms various models, including the well-known robust "$1/N$" portfolio. When the bad-regime probability increases, investors reduce their exposure to risky assets or even move their entire investment to the risk-free asset. Investors also take advantage of the improved estimation of time-varying correlation to select better-diversified asset allocations with higher Sharpe ratio.

A regime-switching model with jumps is a fairly simple and intuitive model that should help improve managing risk in international portfolios. Risk models that feature correlation asymmetries are especially appealing when the occurrence of worldwide bad markets seem more frequent.

References

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

Ait-Sahalia, Y., and D. Xiu. 2016. Increased correlation among asset classes: Are volatility or jumps to blame, or both? *Journal of Econometrics* Advance Access published May 28, 2016, 10.1016/j.jeconom.2016.05.002.

Alexander, C., and E. Lazar. 2006. Normal mixture GARCH(1,1): Applications to exchange rate modelling. *Journal of Applied Econometrics* 21:307–36.

Almazan, A., K. C. Brown, M. Carlson, and D. A. Chapman. 2004. Why constrain your mutual fund manager? *Journal of Financial Economics* 73:289–321.

Ang, A., and G. Bekaert. 2002. International asset allocation with regime shifts. *Review of Financial Studies* 15:1137–87.

Ang, A., and J. Chen. 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63:443–94.

Ball, C., and W. Torous. 2000. Stochastic correlation across international stock markets. *Journal of Empirical Finance* 7:373–88.

Barndorff-Nielsen, O. E., and N. Shephard. 2006. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4:1–30.

Bollerslev, T. 1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics* 72:498–505.

Boyer, B. H., M. S. Gibson, and M. Loretan. 1999. Pitfalls in tests for changes in correlations. Working Paper 597R, Federal Reserve Board International Finance Division.

Campbell, R., K. Koedijk, and P. Kofman. 2002. Increased correlation in bad markets. *Financial Analysts Journal* 58:87–94.

Cappiello, L., R. F. Engle, and K. Sheppard. 2006. Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics* 4:537–72.

Christoffersen, P., V. Errunza, K. Jacobs, and H. Langlois. 2012. Is the potential for international diversification disappearing? A dynamic copula approach. *Review of Financial Studies* 25:3711–51.

Das, S., and R. Uppal. 2004. Systemic risk and international portfolio choice. *Journal of Finance* 59:2809–34.

Davies, R. 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 74:33–43.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22:1915–53.

DeMiguel, V., A. Martin-Utrera, and F. J. Nogales. 2013. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking and Finance* 37:3018–34.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society: Series B* 39:1–38.

Duncan, J., J. Randal, and P. Thomson. 2009. Fitting jump diffusion processes using the EM algorithm. Working Paper, Victoria University of Wellington.

Engle, R. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20:339–50.

Erb, C. B., C. R. Harvey, and T. E. Viskanta. 1994. Forecasting international correlation. *Financial Analysts Journal* 50:32–45.

Forbes, K. J., and R. Rigobon. 2002. No contagion, only interdependence: Measuring stock market co-movements. *Journal of Finance* 57:2223–61.

Galambos, J. 1978. The asymptotic theory of extreme order statistics. New York: John Wiley and Sons.

Glosten, L. R., R. Jagannathan, and D. E. Runkle. 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48:1779–801.

Guidolin, M., and A. Timmermann. 2007. Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control* 31:3503–44.

Hansen, B. E. 1992. The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP. *Journal of Applied Econometrics* 7:S61–S82.

Honda, T. 2003. Optimal portfolio choice for unobservable and regime-switching mean returns. *Journal of Economic Dynamics and Control* 28:45–78.

Jobson, J. D., and B. M. Korkie. 1981. Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance* 36:889–908.

Kim, S., N. Shephard, and S. Chib. 1998. Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65:361–93.

Ledoit, O., and M. Wolf. 2008. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance* 15:850–9.

Liu, J., F. Longstaff, and J. Pan. 2003. Dynamic asset allocation with event risk. *Journal of Finance* 58:231–59.

Longin, F., and B. Solnik. 1995. Is the correlation in international equity returns constant: 1960–1990? *Journal of International Money and Finance* 14:3–26.

———. 2001. Extreme correlation of international equity markets. *Journal of Finance* 56:649–76.

Memmel, C. 2003. Performance hypothesis testing with the Sharpe ratio. *Finance Letters* 1:21–23.

Merton, R. C. 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3:125–44.

Oh, D. H., and A. J. Patton. 2012. Modelling dependence in high dimensions with factor copulas. Working Paper, Duke University.

Okimoto, T. 2008. New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis* 43:787–816.

Omori, Y., and T. Ishihara. 2012. Multivariate stochastic volatility models. In Handbook of volatility models and their applications, pp. 221–48. Eds. L. Bauwens , C. M. Hafner, and S. Laurent. New York: John Wiley and Sons.

Pickard, D. K., P. J. Kempthorne, and A. R. Zakaria. 1986. Inference for jump diffusion stock prices. Proceedings of the Business and Economic Statistics Section of American Statistical Association, 107–111.

Pukthuanthong, K., and R. Roll. 2015. Internationally correlated jumps. *Review of Asset Pricing Studies* 5:92–111.

Ramchand, L., and R. Susmel. 1998. Volatility and cross correlation across major stock markets. *Journal of Empirical Finance* 5:397–416.

Rivers, D., and Q. Vuong. 2002. Model selection tests for nonlinear dynamic models. *Econometrics Journal* 5:1–39.

Ryden, T., T. Terasvirta, and S. Asbrink. 1998. Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics* 13:217–44.

Stambaugh, R. F. 1995. Unpublished discussion of Karolyi and Stulz (1996). National Bureau of Economic Research, Universities Research Conference on Risk Management.

Wu, C. F. J. 1983. On the convergence properties of the EM algorithm. *Annals of Statistics* 11:95–103.

Yu, J. 2002. Forecasting volatility in the New Zealand stock market. *Applied Financial Economics* 12:193–202.