



WP 47_12

John M. Maheu

University of Toronto, Canada
The Rimini Centre for Economic Analysis (RCEA), Italy

Thomas H. McCurdy

University of Toronto, Canada
CIRANO, Canada

Xiaofei Zhao

University of Toronto, Canada

DO JUMPS CONTRIBUTE TO THE DYNAMICS OF THE EQUITY PREMIUM?

Copyright belongs to the author. Small sections of the text, not exceeding three paragraphs, can be used provided proper acknowledgement is given.

The *Rimini Centre for Economic Analysis* (RCEA) was established in March 2007. RCEA is a private, nonprofit organization dedicated to independent research in Applied and Theoretical Economics and related fields. RCEA organizes seminars and workshops, sponsors a general interest journal *The Review of Economic Analysis*, and organizes a biennial conference: *The Rimini Conference in Economics and Finance* (RCEF). The RCEA has a Canadian branch: *The Rimini Centre for Economic Analysis in Canada* (RCEA-Canada). Scientific work contributed by the RCEA Scholars is published in the RCEA Working Papers and Professional Report series.

The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Rimini Centre for Economic Analysis.

Do Jumps Contribute to the Dynamics of the Equity Premium?*

John M. Maheu[†] Thomas H. McCurdy[‡] Xiaofei Zhao[§]

This Draft, November 2011

Abstract

This paper investigates whether risks associated with time-varying arrival of jumps and their effect on the dynamics of higher moments of returns are priced in the conditional mean of daily market excess returns. We find that jumps and jump dynamics are significantly related to the market equity premium. The results from our time-series approach reinforce the importance of the skewness premium found in cross-sectional studies using lower-frequency data; and offer a potential resolution to sometimes conflicting results on the intertemporal risk-return relationship. We use a general utility specification, consistent with our pricing kernel, to evaluate the relative value of alternative risk premium models in an out-of-sample portfolio performance application.

*The authors are grateful to Peter Christoffersen, George Constantinides, Redouane Elkamhi, Massimo Guidolin, Neil Pearson, Lukasz Pomorski and Liyan Yang, as well as participants at the 22nd Annual Conference on Financial Economics and Accounting (CFEA) at Indiana University, the 2011 Desautels-HEC-Rotman Winter Finance Workshop, the Faculty of Business, Brock University, and the 2011 Workshop on Time Series and Financial Econometrics at Bocconi University, for valuable comments. We also thank the Social Sciences and Humanities Research Council of Canada for financial support.

[†]Department of Economics, University of Toronto and RCEA Italy, jmaheu@chass.utoronto.ca

[‡]Rotman School of Management, University of Toronto and CIRANO, tmccurdy@rotman.utoronto.ca

[§]Rotman School of Management, University of Toronto, Xiaofei.Zhao08@rotman.utoronto.ca

I. Introduction

This paper evaluates whether jumps contribute to the dynamics of the equity premium for a broadly diversified portfolio of U.S. stocks. Motivated by the preference-restricted nonlinear pricing kernel proposed by Dittmar (2002), we test whether the conditional variance, skewness and kurtosis risks are priced in aggregate stock returns. Our focus is the effect of jumps on the dynamics of those conditional moments and consequently, if priced, on the dynamics of expected excess returns (the equity premium) associated with the market portfolio. Building on Harvey and Siddique (2000) and Guidolin and Timmermann (2008), we use a general utility specification, consistent with the pricing kernel, to evaluate the relative utility of alternative risk premium models in an out-of-sample portfolio performance application.

A. Overview of Our Model and Empirical Results

Our model filters daily market excess returns into large versus smaller changes, simultaneously with estimation of all of the parameters of the conditional distribution. In our parameterization, large changes in daily returns (jumps) contribute to the dynamics of conditional variance, the dynamics of conditional skewness and kurtosis, and consequently the dynamics of expected return through pricing of the associated risks. This allows expected jumps to have an impact (whether or not they occur) on the shape *and* location of the distribution of market excess returns.

More specifically, we model innovations of the return process using a GARCH-jump mixture model. That is, the jump component of the innovation follows a compound Poisson-Normal distribution with an auto-regressive jump intensity and a normal jump size distribution. The diffusive component of the innovation is directed by an asymmetric 2-component GARCH process, and allows the persistence of jump effects on variance to be different than that of the diffusive component. These features are important for our pricing application since the 2nd GARCH component helps control for noise associated with daily returns and, as such, improves the sorting into jumps versus diffusive components.

Flexible modeling of the conditional variance, skewness and kurtosis dynamics will undoubtedly improve the explanatory power of the model for capturing the changing shape of the distribution. However, the focus of this paper is concerned with whether the dynamics of those higher moments of returns are associated with time-varying expected returns. In particular, are the risks associated with the arrival of jumps, and their effect on the higher moments of returns, priced in the mean? In other words, since our application is to excess returns on a market-wide index of stocks, are jumps related to the dynamics of the market equity premium?

Studies on jumps often assume that the compensation for jump risk is a linear function of the jump intensity, mostly to make risk neutral pricing (of options) tractable. In contrast,

using a pricing kernel associated with general preferences to derive our equity premium specification,¹ we allow the jump risk to be priced both linearly and nonlinearly. More specifically, following Guidolin and Timmermann (2008), we have a 4-moment specification for the equity premium. However, in our case, jump risk is priced *linearly* through the conditional dynamics of variance, and *nonlinearly* through conditional skewness and kurtosis. To the best of our knowledge, this is the first study to find significant pricing of both jump risk and diffusive risk, as well as realistic total equity premium estimates, using only a time series of equity return data.

Our empirical results show that higher-order moments are significantly priced in the equity premium. First of all, we find a positive risk-return tradeoff associated with the traditional risk for the market measured by the conditional variance.² The pricing of the conditional variance is robust across our proposed time-varying jump model specifications. Interestingly, when we restrict the model to have no jumps and only include one GARCH component, the variance dynamics are not significantly priced.

To see to what extent jumps are related to the equity premium dynamics, we analyze the marginal effect of jumps on the equity premium by fixing the GARCH component of volatility. We show that the marginal effect of jumps on the equity premium is positive at all levels of the GARCH volatility. The equity premium is increasing in the conditional jump frequency and this increase is greatest for low jump arrival rates and for low levels of the GARCH variance component. For our parameterization and sample, if the expected number of jumps increases by one per year, a representative investor will demand, on average, 0.1057% additional expected return for taking on the extra jump risk. This implies that the equity premium associated with jumps is about 3.96% per annum on average. All higher-order moments can be affected by jumps to returns. According to our parameter estimates, on average jumps contribute 1.12% to the equity premium through the variance dynamics and also add 2.84% to the equity premium through their contribution to skewness.

We find robust pricing of both the conditional variance and the conditional skewness in the market equity premium. For example, the equity premium associated with skewness risk is about 3.5% per annum. This is very close to the 3.6% per annum risk premium compensation for systematic skewness found by Harvey and Siddique (2000) who study the conditional skewness in a cross-section of monthly stock returns. Interestingly, when we impose the preference restriction of a non-negative price associated with kurtosis risk, our findings show that the price of kurtosis risk is close to zero. This is due to the high negative correlation between the conditional skewness and conditional kurtosis in the market index. Conditional kurtosis is significantly priced with a positive sign when the skewness factor is not included. In

¹See sections II.E and VII.C for details.

²Assuming power utility, our estimate of the price of conditional variance risk implies a relative risk aversion of 2.7 for a representative agent. However, our proposed specification generalizes the power utility assumption.

other words, at least at the market level, the skewness and kurtosis are likely to be representing the same source of risk.

Our results also offer a potential resolution to the conflicting results in the literature on risk and expected return for the market as a whole. We find a significantly positive equity premium but our results also reveal that a positive relationship between conditional variance and return only occurs when the GARCH variance component is at or above average levels. During calm times (low level of the GARCH variance component), the skewness effect dominates to generate a positive equity premium. For these periods if we were to omit conditional skewness from the equity premium specification, we could see a negative equity premium estimate (reflecting the negative relationship between the equity premium and the conditional variance at low levels of the latter), due to the missing skewness factor. In more volatile times, the variance premium effect dominates and we will be able to see a positive risk-variance tradeoff, whether we include conditional skewness in the equity premium specification or not.

Finally, we assess the value-added associated with pricing risk from jumps to returns that contribute to the dynamics of conditional variance, skewness and kurtosis. We do this by evaluating the realized utility and certainty-equivalent returns associated with a simple portfolio allocation application which requires out-of-sample equity premium forecasts. We conduct this out-of-sample evaluation for our model versus several special case benchmarks. Solving for the functional relationship between the parameters of our assumed general utility function and the prices of risk associated with the associated asset pricing model, we are able to calibrate the implied utility parameters to the empirical estimates for our equity premium specification. Our maintained *prudence* model generates higher realized utility and certainty-equivalent returns than special case benchmarks, including one that does not include jumps.

B. Related Literature and Our Model Features

There is a large literature extending the linear single-factor Capital Asset Pricing Model (CAPM).³ An alternative approach is a CAPM-type model that evaluates whether higher-order moments of returns are priced. Notably, Harvey and Siddique (2000) extend the Kraus and Litzenberger (1978) 3-moment CAPM to a conditional setting in an extensive evaluation of whether co-skewness is priced in the cross-section of equity returns.⁴

³Including multi-factor asset pricing models, such as, Fama and French (1996) and the Merton (1973) intertemporal CAPM (ICAPM) model which prices factors that hedge against changes in the intertemporal investment opportunity set; and nonparametric asset pricing models, for example, Bansal and Viswanathan (1993), Chapman (1997) and Rossi and Timmermann (2009), which approximate the pricing kernel using a flexible functional form.

⁴For additional development or application of 3-moment CAPM models see, for example, Friend and Westerfield (1980), Sears and Wei (1985), Lim (1989), Harvey and Siddique (1999), Hwang and Satchell (1999), and Smith (2007). Chang, Christoffersen, and Jacobs (2009) use the ICAPM to motivate their evaluation of whether market skewness risk is priced in the cross-section of stock returns.

Dittmar (2002) derives a pricing kernel that depends on higher-order moments by approximating the pricing kernel using a Taylor series expansion of the marginal utility of returns on aggregate wealth. Furthermore, Dittmar (2002) links preference-based asset pricing models with this nonparametric approach by using preference theory to restrict the sign of the coefficients. His cubic pricing kernel is consistent with a CAPM model in which investors have preferences that depend on the first four moments of returns as in Guidolin and Timmermann (2008). As in some of the pure nonparametric approaches, these parameterizations can capture nonlinearities in the relationship between risks and expected return.

Guidolin and Timmermann (2008) analyze the international asset allocation implications of time-variation in higher-order moments of stock returns using a 4-moment international CAPM allowing time-varying prices of covariance, co-skewness and co-kurtosis risk. Among other things, they explore the relative importance of modeling preferences for skewness and kurtosis versus incorporating time-varying moments within a typical mean-variance preference model.

Conrad, Dittmar, and Ghysels (2009) use option prices and the method of Bakshi, Kapadia, and Madan (2003) to estimate higher moments of the underlying security's risk-neutral return distribution which they then relate to subsequent returns. Duan and Zhang (2010) estimate investors' risk aversion using the Bakshi and Madan (2006) expression for volatility spread (between risk-neutral volatility and physical return volatility). Applying this methodology assuming power utility, Duan and Zhang (2010) derive an expression for the equity premium which is a function of the physical return variance, skewness and kurtosis with coefficients that are restricted to be a function of the risk aversion estimate.

In Guidolin and Timmermann (2008) variations in skewness and kurtosis are linked to uncertainty associated with shifts between bull and bear states. In our paper, the driver of conditional skewness and kurtosis dynamics is time-varying jump arrival and the associated multiple-component GARCH variance dynamics. Chernov, Gallant, Ghysels, and Tauchen (2003) suggest that either a 2-component parameterization of stochastic volatility (SV) or a SV-jump (SV-J) diffusion can capture the volatility dynamics.⁵ We allow for dependence in the jump arrival process in a discrete-time setting.⁶

Most of the extant literature on pricing jump risk has focused on option pricing issues. A few, for example, Pan (2002), Christoffersen, Jacobs, and Ornthanalai (2008) and Elkamhi and Ornthanalai (2009), have fit options and the underlying jointly but with the focus on solving option pricing puzzles. These papers adopt parametric models for the underlying assets.

⁵Early examples of SV-jump-diffusion specifications with time-varying jump intensities include Bates (2000), Andersen, Benzoni, and Lund (2002), Pan (2002), Chernov et al. (2003) and Eraker, Johannes, and Polson (2003).

⁶Examples with time-varying jump arrival include Pan (1997), Bekaert and Gray (1998), Bates and Craine (1999), Neely (1999) and Das (2002), who allow a volatility factor or financial/macroeconomic variables to affect the jump intensity. Johannes, Kumar, and Polson (1999) consider a state dependent jump model which allows past jumps and observables to affect the jump probability.

Another strand of the literature employs a nonparametric approach to estimate realized jumps from high-frequency data. For example, Bollerslev and Todorov (2011) estimate the market jump tail under the physical measure using high frequency intra-day data and estimate the risk-neutral counterpart from index options.

Our paper differs from these alternative approaches to pricing jump risk in several important dimensions. First of all, studies using both underlying asset returns and options are based on no-arbitrage (risk neutral) pricing whereas our specification of the equity premium is based on equilibrium asset pricing theory. Consequently, we only need the index return data to estimate the jump risk component of the equity premium. Studies which require options data or very high-frequency data will be restricted to shorter samples due to data availability. In some cases, the price of jump risk for the underlying is restricted to be zero, in some others it is imprecisely estimated, or implies an equity premium that is counterfactually large. Learning about tail events and jump dynamics could require a long span of calendar data such as used in our paper.

Secondly, our equity premium specification allows jumps to be priced linearly through the conditional variance and nonlinearly through the higher-order moments. In contrast, papers relying on options data typically assume that the compensation for jump risk is a linear function of the jump intensity, in order to have tractable option pricing formula. As we will demonstrate in our results section, not having the nonlinear pricing of jumps could be a potential source of mis-specification that makes many papers fail to find significant pricing of both jump risk and diffusive risk, especially when only equity data are available.

Thirdly, our parametric model differs from those which employ nonparametric methods to estimate realized jumps from high-frequency data, in that we estimate both the jump intensity and jump size distribution at the same time, along with the other parameters of the conditional distribution. Our approach allows us to filter returns into large changes and smaller changes, based on their different dynamic properties, simultaneously with estimation of all of the parameters of the conditional distribution. In our parameterization, large changes in daily returns contribute to the dynamics of conditional variance, the dynamics of conditional skewness and kurtosis, and consequently the dynamics of expected return through pricing of the associated risks. This allows expected jumps to have an impact on the shape and location of the distribution, whether or not they occur (the peso problem).

Finally, in contrast to many existing papers focusing on option pricing, we allow jump arrival to be directed by a different process than squared-return innovations. By introducing this additional source of dynamics, our approach is more flexible than those that parameterize jump arrival as an affine function of the time-varying volatility.

The next section outlines our flexible model for the dynamics of market excess returns followed by the parameterization of the equity premium and its relation to the conditional variance, skewness and kurtosis dynamics. Implications for how these conditional moments

could be related to the equity premium are discussed with reference to equilibrium pricing theory. Section III summarizes data sources and the estimation method; followed by a summary of the estimation results in section IV. Those results include parameter estimates for our maintained model derived in section II and some model comparisons with special cases, as well as a summary of the components of the equity risk premium dynamics. Section V evaluates the out-of-sample performance of our maintained model relative to some benchmarks. Section VI concludes. In the appendix, section VII.A summarizes how we scaled the daily returns to facilitate estimation. Since our derivation in section II uses continuously compounded returns, section VII.B derives the analogous equity risk premium associated with simple returns, that is, returns compounded per period rather than continuously. Finally, section VII.C presents how we calibrate the general utility function used in the portfolio allocation exercise to the estimates of our maintained model.

II. Derivation of our Model

We follow Harvey and Siddique (2000), Dittmar (2002) and Guidolin and Timmermann (2008) in providing an asset pricing derivation for an empirical specification that tests whether or not higher-order moments of returns are priced as risk factors. In our case, the dynamics of the conditional third and fourth moments are driven by an autoregressive jump arrival process; and the conditional variance dynamics are driven by both jumps and a 2-component GARCH process. The focus of our paper is on the potential effects of jumps on the equity premium. Jumps affect the dynamics of conditional variance, skewness and kurtosis. Using results from equilibrium pricing theory summarized in sections II.E and VII.C, we explore whether those dynamics are priced as risk factors for the equity premium.

A. Dynamics of continuously compounded returns

We define the continuously compounded excess return on the market index as

$$r_t \equiv r_{m,t} - r_{f,t}, \tag{1}$$

in which $r_{m,t}$ is the continuously compounded return (including distributions) on the market index and $r_{f,t}$ is the continuously compounded riskfree rate. Henceforth, we will usually refer to r_t , the excess continuously compounded return on the market, as the log return. In the following the information set is $\Phi_t = \{r_1, \dots, r_t\}$.

We label the continuously compounded market equity premium expected for period t , given information at time $t - 1$, as m_t . This conditional equity premium is parameterized in

section II.E below. Assume that the dynamics of realized log returns are driven by

$$r_t = m_t + u_t, \quad (2)$$

where

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \epsilon_t. \quad (3)$$

That is, u_t has a predictable autoregressive component due to stale prices⁷, or missing pricing factors, and a mean zero return innovation ϵ_t .

Combining equations (2) and (3), and decomposing the total return innovation ϵ_t into two components, we can rewrite realized log returns as

$$r_t = m_t + \rho_1(r_{t-1} - m_{t-1}) + \rho_2(r_{t-2} - m_{t-2}) + \epsilon_{1,t} + \epsilon_{2,t}, \quad (4)$$

in which m_t is the continuously compounded market equity premium expected for period t , given information, Φ_{t-1} , available at time $t - 1$. In addition, we assume that log returns are driven by two stochastic innovations $\epsilon_{1,t}$ and $\epsilon_{2,t}$. In particular: $\epsilon_{2,t}$ is a jump innovation to returns, compensated so that it is mean zero; $\epsilon_{1,t}$ is a mean-zero normal innovation to returns directed by a conditional normal process; $\epsilon_{1,t}$ and $\epsilon_{2,t}$ are *contemporaneously* independent.

Note that the conditional mean of the log return process is

$$E[r_t | \Phi_{t-1}] = m_t + \rho_1(r_{t-1} - m_{t-1}) + \rho_2(r_{t-2} - m_{t-2}), \quad (5)$$

but it is the market equity premium m_t (the conditional mean net of any remaining serial correlation due to, for example, stale prices) that we evaluate below using an asset pricing framework.

B. Parameterization of the jump component

The mean-zero (compensated) innovation to returns from jumps is labelled $\epsilon_{2,t}$. This innovation is directed by a conditional compound Poisson-Normal distribution reflecting a conditional Poisson jump arrival process combined with a conditional normal jump-size distribution.

⁷As described in section III.A below, our data are daily index returns, the components of which do not all trade every day. See Campbell, Lo, and MacKinlay (1997) for a derivation of the resulting autoregressive structure in realized returns. In Table IV we show that our results with respect to the equity premium are robust to whether or not we include this autoregressive structure.

B.1. Time-varying arrival of jumps

Define the discrete-valued number of jumps over the interval $(t - 1, t)$ as $n_t \in 0, 1, 2, \dots$. The conditional distribution of n_t is Poisson with parameter λ_t , that is,

$$P(n_t = j | \Phi_{t-1}) = \frac{\exp(-\lambda_t) \lambda_t^j}{j!}, j = 0, 1, 2, \dots \quad (6)$$

The conditional arrival rate of jumps, λ_t , is the expected number of jumps for period t given information at time $t - 1$, that is,

$$E[n_t | \Phi_{t-1}] = \lambda_t. \quad (7)$$

In other words, the number of jumps in period t , n_t , is directed by a conditional Poisson process with a time-varying jump arrival rate λ_t .

As in Chan and Maheu (2002) and Maheu and McCurdy (2004), we parameterize the time-series dynamics of λ_t as an autoregressive process (labelled ARJI for autoregressive jump intensity):

$$\lambda_t = \gamma_0 + \gamma_1 \lambda_{t-1} + \gamma_2 \zeta_{t-1}, \quad (8)$$

in which the jump arrival *innovation* for period $t - 1$ is defined as

$$\zeta_{t-1} = E[n_{t-1} | \Phi_{t-1}] - \lambda_{t-1} \quad (9)$$

$$= E[n_{t-1} | \Phi_{t-1}] - E[n_{t-1} | \Phi_{t-2}]. \quad (10)$$

Jumps are latent; the expected number of jumps is computed using the estimation filter. The jump arrival innovation is the update in the expected number of jumps, for example, for period $t - 1$, when information is updated from period $t - 2$ to period $t - 1$. The time series parameterization for the expected number of jumps, λ_t , given by equation (8), has a smoothing or persistence coefficient, γ_1 , associated with the expected number of jumps for the previous period, as well as a news-impact coefficient, γ_2 , associated with the jump arrival innovation.

It is important to note that we allow jump arrival to be directed by a different process than squared return innovations. Instead, the autoregressive jump frequency is directed by measurable jump-arrival innovations. This allows the impact and persistence of time-varying jump arrival on expected variance dynamics to be different than that captured by the GARCH component of variance.

B.2. The jump innovation to returns

The compensated jump innovation to returns, $\epsilon_{2,t}$, is given by

$$\epsilon_{2,t} = J_t - \theta \lambda_t \quad (11)$$

where the total size of jumps in period t , J_t , is

$$J_t = \sum_{k=1}^{n_t} Y_{t,k}, \quad (12)$$

in which $Y_{t,k}$ is the size of jump k in period t which is drawn from a normal distribution with mean θ and variance δ^2 as in:

$$Y_{t,k} \sim N(\theta, \delta^2). \quad (13)$$

Note that we estimate the moments of this jump size distribution which, for example, contributes to skewness of the return distribution by allowing the average jump size to be different from zero.

Since

$$E[J_t | \Phi_{t-1}] = \theta \lambda_t, \quad (14)$$

the compensated jump innovation is mean zero, that is,

$$E[\epsilon_{2,t} | \Phi_{t-1}] = E[J_t | \Phi_{t-1}] - \theta \lambda_t = 0. \quad (15)$$

C. Parameterization of the normal innovation component

The normal innovation to returns, $\epsilon_{1,t}$, is assumed to be directed by a 2-component GARCH specification⁸ with feedback from jumps. This specification, which generalizes that in Maheu and McCurdy (2004), is parameterized as follows:

$$\epsilon_{1,t} = \sigma_t z_t, \quad z_t \sim NID(0, 1), \quad \epsilon_{1,t} | \Phi_{t-1} \sim N(0, \sigma_t^2) \quad (16)$$

where σ_t^2 is directed by a two-component GARCH process specified as:

$$\sigma_t^2 = \sigma_{1,t}^2 + \sigma_{2,t}^2, \quad (17)$$

$$\sigma_{1,t}^2 = \omega + g_1(A_1, \Phi_{t-1}) \epsilon_{t-1}^2 + \beta_1 \sigma_{1,t-1}^2, \quad (18)$$

$$\sigma_{2,t}^2 = g_2(A_2, \Phi_{t-1}) \epsilon_{t-1}^2 + \beta_2 \sigma_{2,t-1}^2, \quad (19)$$

$$\epsilon_{t-1} = \epsilon_{1,t-1} + \epsilon_{2,t-1}, \quad (20)$$

The long run component is captured by $\sigma_{1,t}^2$ while the transitory moves in the GARCH conditional variance are modeled by $\sigma_{2,t}^2$. To help visualize their properties using our maintained model estimates, we plot the paths of these two components for 2007 in Figure 1. Having the 2nd component helps capture the diffusive volatility better. More importantly, for our purpose, the 2nd component helps control the noisy or transitory part of the diffusive volatility, without

⁸Other component GARCH-type models include Engle and Lee (1999), Maheu and McCurdy (2007) and Chan and Feng (2009).

which the noise could potentially be sorted as jumps, making jumps less precisely estimated.

[Insert Figure 1 about here.]

The generalized news impact coefficient $g_i(A_i, \Phi_{t-1})$ for the i^{th} GARCH component, $i = 1, 2$, allows asymmetric impact from good versus bad news, as well as from jump versus normal innovations. That is,

$$\begin{aligned} g_i(A_i, \Phi_{t-1}) &= \exp(\alpha_i + I(\epsilon_{t-1})(\alpha_{a,j,i}E[n_{t-1}|\Phi_{t-1}] + \alpha_{a,i})), i = 1, 2, \\ I(\epsilon_{t-1}) &= 1 \text{ if } \epsilon_{t-1} < 0, \text{ otherwise } 0. \end{aligned} \quad (21)$$

Recall that $E[n_{t-1}|\Phi_{t-1}]$ is the expected number of jumps at time $t - 1$ given Φ_{t-1} , provided by our estimation filter.

Therefore, jumps are allowed to affect variance dynamics in several ways. The direct effect originates from dynamics associated with the autoregressive jump frequency which is directed by measurable jump-arrival innovations. There is also a feedback effect of jumps through the GARCH parameterization of the impact of squared return innovations on future variance. In particular, we allow return innovations to have asymmetric impact and persistence effects on future variance depending on whether the source of the innovations was from jumps or normal innovations and whether or not it was associated with good or bad news.

D. Dynamics of higher-order conditional moments

The conditional variance (v_t), conditional skewness (s_t), and conditional kurtosis (k_t) are:

$$v_t = \sigma_t^2 + \lambda_t(\theta^2 + \delta^2), \quad (22)$$

$$s_t = \frac{\lambda_t(\theta^3 + 3\theta\delta^2)}{(\sigma_t^2 + \lambda_t\theta^2 + \lambda_t\delta^2)^{3/2}}, \quad (23)$$

$$k_t = 3 + \frac{\lambda_t(\theta^4 + 6\theta^2\delta^2 + 3\delta^4)}{(\sigma_t^2 + \lambda_t\delta^2 + \lambda_t\theta^2)^2}. \quad (24)$$

Clearly jumps affect all of the conditional moments. As indicated in equation (22), the moments of the jump size distribution and jump arrival have a direct impact, measured by $\lambda_t(\theta^2 + \delta^2)$, on the conditional variance. However, as equation (21) shows, jumps also contribute to volatility clustering. That is, jumps are allowed to affect the future conditional variance through σ_t^2 .

Time-varying jump arrival will be the source of time-variation in the conditional 3rd and 4th moments (numerators of equations (23) and (24) respectively), whereas the conditional skewness and kurtosis statistics, s_t and k_t , will also be affected by time variation in the variance clustering component σ_t^2 which is the total GARCH effect augmented by any persistence in the jump impacts.

Note that if jump arrival was constant but non-zero, $\lambda_t \equiv \lambda > 0$, we would still have time-varying and non-normal levels of skewness and kurtosis, as measured by s_t and k_t . On the other hand if there were no jumps expected, $\lambda_t = 0$ for all t , conditional skewness and conditional kurtosis would be the same as that for a conditional normal distribution. We estimate both of these special cases as part of our robustness analyses.

E. Parameterization of the equity premium

Applying the 3-moment CAPM of Harvey and Siddique (2000) to the market as a whole gives the following pricing relationship

$$E[r_t|\Phi_{t-1}] = \eta_v E[\epsilon_t^2|\Phi_{t-1}] + \eta_s E[\epsilon_t^3|\Phi_{t-1}], \quad (25)$$

in which $E[\epsilon_t^2|\Phi_{t-1}]$ is the conditional variance and $E[\epsilon_t^3|\Phi_{t-1}]$ is the third central moment of market excess returns.

Building on Harvey and Siddique (2000) and Guidolin and Timmermann (2008), we propose a 4-moment application. We begin by taking an N^{th} -order Taylor-series expansion of a general utility function, setting $N = 4$, so that

$$U(W_{t+1}) \approx \sum_{n=0}^4 \frac{U^{(n)}(W_t)}{n!} (W_{t+1} - W_t)^n = \sum_{n=0}^4 \frac{U^{(n)}(W_t)}{n!} (W_t R_{t+1}^W)^n, \quad (26)$$

in which W_t is the known initial wealth level about which the utility is expanded, $W_{t+1} = W_t(1 + R_{t+1}^W)$, and R^W is the net return on aggregate wealth. Correspondingly, the Taylor-series expansion of marginal utility is

$$U'(W_{t+1}) \approx \sum_{n=0}^3 \frac{U^{(n+1)}(W_t)}{n!} (W_{t+1} - W_t)^n = \sum_{n=0}^3 \frac{U^{(n+1)}(W_t)}{n!} (W_t R_{t+1}^W)^n. \quad (27)$$

Without loss of generality, we normalize the constant initial wealth level so that $W_t = 1$.

This implies that the pricing kernel, $M_{t+1} \equiv \frac{U'(W_{t+1})}{U'(W_t)}$, is

$$M_{t+1} \approx g_{0t} + g_{1t}R_{t+1}^W + g_{2t}(R_{t+1}^W)^2 + g_{3t}(R_{t+1}^W)^3. \quad (28)$$

in which $g_{nt} = \frac{U^{(n+1)}(W_t)}{U'(W_t)} \frac{1}{n!}$ for $n = 0, 1, 2, 3$.

As in Dittmar (2002), $U' > 0, U'' < 0, U''' > 0$, and $U'''' < 0$, that is, positive marginal utility, risk aversion, decreasing absolute risk aversion, and decreasing absolute prudence,⁹ respectively, imply that $g_{1t} < 0, g_{2t} > 0, g_{3t} < 0$ while $g_{0t} = 1$. Since the utility approximation is truncated at $N = 4$, $g_{nt} = 0$ for $n > 3$.

⁹As derived in Kimball (1990), and Kimball (1993), decreasing absolute prudence implies that as wealth increases the precautionary savings motive declines.

Using the fundamental pricing equation:

$$E_t[(1 + R_{t+1}^W)M_{t+1}] = 1, \quad (29)$$

with Equation (28), and assuming the existence of a riskfree asset with net return R_t^f , we have a four-moment asset pricing model, as in Guidolin and Timmermann (2008). For our application,

$$E_t[R_{t+1}^W] - R_t^f = \kappa_{1t}Cov_t(R_{t+1}^W, R_{t+1}^W) + \kappa_{2t}Cov_t(R_{t+1}^W, (R_{t+1}^W)^2) + \kappa_{3t}Cov_t(R_{t+1}^W, (R_{t+1}^W)^3) \quad (30)$$

where $\kappa_{nt} = -g_{nt}(1 + R_t^f)$. Our preference assumptions imply that $\kappa_{1t} > 0, \kappa_{2t} < 0, \kappa_{3t} > 0$, that is, covariance and co-kurtosis earn positive premia while co-skewness earns a negative risk premium.

As derived further in section VII.C below, for our parameterization the conditional market equity premium m_t is

$$m_t = \psi_v v_t + \psi_s s_t + \psi_k k_t, \quad \psi_v \geq 0, \psi_s \leq 0, \psi_k \geq 0, \quad (31)$$

in which v_t, s_t and k_t are the conditional variance, and the conditional skewness and kurtosis statistics, respectively, all measurable with respect to information at time $t-1$.¹⁰ We estimate this conditional equity premium with and without an intercept. The restrictions on the parameters, notably that $\psi_s \leq 0$ and $\psi_k \geq 0$, follow from the preference specification and the associated approximation to the pricing kernel given by equation (28). Section VII.C below also shows how we use our empirical estimates to calibrate the utility parameters for our out-of-sample portfolio application.

The special case of power utility, as in Duan and Zhang (2010), is also consistent with the restrictions $\psi_s \leq 0, \psi_k \geq 0$. This parameterization is:

$$m_t = \psi_v v_t + \psi_s v_t^{\frac{3}{2}} s_t + \psi_k v_t^2 k_t, \quad \text{where} \quad (32)$$

$$\psi_v = (\gamma - 0.5), \psi_s = -(3\gamma^2 - 3\gamma + 1)/6, \psi_k = (4\gamma^3 - 6\gamma^2 + 4\gamma - 1)/24. \quad (33)$$

For robustness, we also estimate an unrestricted version with no restrictions on the coefficients ψ_v, ψ_s, ψ_k associated with the conditional market equity premium parameterization given by equation (31). However, our focus is to evaluate whether jumps contribute to the dynamics of the premium as parameterized in equation (31) or special cases thereof. Note that we include the effects of jumps on the conditional variance v_t , as well as on the higher-order (normalized) conditional moments s_t and k_t .

¹⁰Note that the 3rd and 4th central moments of market excess returns will be $v_t^{\frac{3}{2}} s_t$ and $v_t^2 k_t$, respectively.

III. Data and Estimation

A. Data

Our data are returns including distributions from a broadly diversified equity index, that is, the CRSP (Center for Research in Security Prices) NYSE/AMEX/NASDAQ value-weighted index (vwretd from dsix) for the period January 2, 1926 to December 31, 2007. These returns are converted to continuously compounded daily returns. For the risk-free rates, we convert 30-day Treasury Bill returns (t30ret from mcti) to continuously compounded monthly returns and divide by 22 to approximate daily risk-free continuously compounded rates. These are subtracted from the continuously compounded daily index returns resulting in our full sample of 21775 daily excess log returns. Descriptive statistics are reported in Table I.

[Insert Table I about here.]

B. Estimation method

As in Maheu and McCurdy (2004), analytical filtering allows one to infer probabilities associated with the unobservable jumps. The filter can be constructed as,

$$P(n_t = j | \Phi_t, \Theta) = \frac{f(r_t | n_t = j, \Phi_{t-1}, \Theta) P(n_t = j | \Phi_{t-1}, \Theta)}{f(r_t | \Phi_{t-1}, \Theta)} \quad j = 0, 1, 2, \dots \quad (34)$$

This filter provides an *ex post* distribution for the number of jumps, n_t . For example, one method to assess whether or not a jump occurred in a particular period would be to use the filter to find the probability that at least one jump occurred. This is, $P(n_t \geq 1 | \Phi_t) = 1 - P(n_t = 0 | \Phi_t)$, which is directly available from model estimation.

The model can be estimated by maximum likelihood. This involves integrating out the number of unobserved jumps. Given the number of jumps j and the parameter vector Θ the conditional density of returns $f(r_t | \Phi_{t-1}, \Theta, n_t = j)$ is

$$\frac{1}{\sqrt{2\pi(\sigma_t^2 + j\delta^2)}} \exp\left(-\frac{1}{2} \frac{(r_t - m_t - \rho_1(r_{t-1} - m_{t-1}) - \rho_2(r_{t-2} - m_{t-2}) - (j - \lambda_t)\theta)^2}{\sigma_t^2 + j\delta^2}\right), \quad (35)$$

where $j \in \{0, 1, 2, \dots\}$. The full likelihood contribution in terms of r_t is then

$$f(r_t | \Phi_{t-1}, \Theta) = \sum_{j=0}^{\infty} f(r_t | \Phi_{t-1}, \Theta, n_t = j) P(n_t = j | \Phi_{t-1}, \Theta) \quad (36)$$

where the second term in the summation is the p.d.f. of the time-varying Poisson distribution in (6). Finally, the full sample loglikelihood is $l = \sum_{t=1}^T \log f(r_t | \Phi_{t-1}, \Theta)$ which is maximized with respect to Θ by a quasi-Newton routine.

The terms in the likelihood and filter involve an infinite summation. To make estimation feasible, we truncated this summation at 25. In practice, for our model estimates, we found that the conditional Poisson distribution had zero probability in the tail for values of $n_t \geq 10$.

IV. Results

A. Parameter Estimates

Table II provides parameter estimates for the full model with alternative specifications for the conditional market equity premium, m_t . Column 2, labelled ‘prudence’, reports estimates for our maintained parameterization of m_t given in equation (31). Column 3, labelled ‘intercept’, provides estimates for the same model including an intercept, μ , in the conditional mean. Column 4, labelled ‘unrestricted’, removes the sign restrictions on skewness and kurtosis implied by decreasing absolute prudence. Column 5, labelled ‘variance’, just prices the conditional variance, v_t ; column 6, labelled ‘skewness’, prices both the conditional variance and conditional skewness; and column 7, labelled ‘kurtosis’ prices both the conditional variance and conditional kurtosis. As described below, we try pricing these components separately due to the strong negative correlation between skewness and kurtosis.

[Insert Table II about here.]

In order to focus on the alternative specifications for the equity premium, all of the models reported in Table II include an AR(2) dynamic in the conditional mean to capture remaining serial correlation, have an identical GARCH specification for volatility clustering, and have the same parameterization for jump dynamics.

The estimation results in Table II reveal significant risk pricing associated with the market equity premium. In our maintained specification (column 2), the coefficient on the conditional variance is 0.027 (ψ_v) with a significant t-stat of 3.3. The skewness coefficient ψ_s is -0.028 with a significant t-stat of -3.4, implying a significant *negative* price of skewness risk. If the skewness statistic, s_t , is also negative, which it is in our sample, this implies that investors will be compensated with extra expected return for being exposed to negative skewness. The kurtosis risk, however, does not seem to be priced in the presence of the skewness risk premium, suggesting that market-wide skewness and kurtosis are capturing the same source of higher-order risk. For the ‘intercept’ case (column 3), the intercept μ , which is used to capture missing pricing factors, is not statistically significant (t-stat -1.4). This result supports our maintained specification given by equation (31) with results reported in column 2.¹¹

The GARCH components of the variance are very similar across models. The first GARCH component captures the long-run dynamics (highly persistent with a β_1 estimate of 0.98);

¹¹Note that our results are not driven by the potential structural break at the end of the 1930s. Estimation using the sub-sample of 1940 to 2007 gives similar results.

while the second GARCH component has a β_2 of about 0.77 indicating lower persistence. As reported in Table V, a 1-component GARCH specification is rejected (p-value 4.13e-52) in favor of the maintained 2-component GARCH parameterization.

The next set of parameters (3rd panel Table II) capture volatility asymmetry with respect to the sign of the return innovation and the inferred number of jumps as parameterized in equation (21). These asymmetries enter both the short and long-run component of the GARCH specification. Again, parameter estimates are very similar across models. In each case, a negative return innovation (bad news) results in a significant increase in the conditional variance ($\alpha_{a,1}$ and $\alpha_{a,2}$ are positive) while an inferred jump contributes to a drop in the GARCH variance ($\alpha_{a,j,1}$ and $\alpha_{a,j,2}$ are negative). That is, jump innovations to returns get incorporated into prices more quickly than normal innovations. This effect ($\alpha_{a,j,1}$) is strongest for the long-run component.

Despite a rich 2-component GARCH specification, there remains strong evidence of jump dynamics in daily returns. The arrival of jumps is autocorrelated with a γ_1 of 0.95 which is highly significant. Jumps tend to arrive in clusters and this will have important implications for the dynamics of the higher-order conditional moments of returns as summarized in equations (22) to (24). The LR test for no autocorrelation in jump arrival is decisively rejected with a p-value of 3.7e-23. The arrival of jumps is infrequent on average. That is, according to the parameter estimates, the unconditional jump arrival rate, $E[\lambda_t] = \gamma_0/(1 - \gamma_1)$, is 0.14 which implies about 35 jumps on average per year in the long run. Also, on average, jumps result in a drop in the market price, that is, the mean of the jump size distribution, θ , is significantly negative.

A.1. Alternative Equity Premium Specifications

We explore additional pricing specifications in columns 5 to 7 of Table II. For example, in column 5 we report results from a parameterization that imposes the restriction that skewness and kurtosis are not priced, that is, $\psi_s = \psi_k = 0$. In this special case, the conditional variance (which includes jump effects) is still significantly priced. However, the likelihood ratio test presented in Table V Panel A rejects $\psi_s = \psi_k = 0$.¹²

It is interesting that if we estimate the power utility special case of our maintained model, that is, equations (32) and (33), the results are very similar to our special case of pricing risk captured by the conditional variance. That is, estimating a risk premium specification under the power utility assumption gives essentially the same results as in column 5 of Table II since ψ_s and ψ_k associated with equations (32) and (33) are estimated to be close to zero and

¹²The test statistic is 8.72 and the reported p-value in Table V Panel A corresponds to two degrees of freedom (two restrictions). However, as shown in Table II, the skewness and the kurtosis factors are both capturing the same risk from a pricing perspective so this is effectively a one degree-of-freedom test in which case the p-value would be lower than that reported in the table.

statistically insignificant in our sample. Note that this special case implies a coefficient of relative risk aversion of about 2.7.¹³

Column 4 of Table II reveals that if we remove the restriction that $\psi_s \leq 0, \psi_k \geq 0$ implied by the preference theory in Dittmar (2002) and others, both skewness and kurtosis are significantly priced in the equity premium, but the price, ψ_k , associated with kurtosis is negative. This is counterintuitive and at odds with the preference-based restrictions.

If we only price the conditional variance and conditional skewness (column 6), the results are essentially the same as in column 3 with the difference due to numerical approximation errors. If we only price conditional variance and conditional kurtosis (column 7), then both these risk factors are also significantly priced in the equity premium. More importantly, the coefficient on the kurtosis becomes significantly positive. Taken together, our results suggest that skewness and kurtosis at the market level seem to pick up the same source of risk.

Investigating the correlation of the estimated skewness and kurtosis (from both our model and sample counterpart), we find a high negative correlation between skewness and kurtosis (-0.96 from our model and -0.93 from the sample estimate computed using a 15-year moving window of the historical returns). A negative correlation between risk-neutral skewness and kurtosis is also documented in Bakshi, Kapadia, and Madan (2003) in a study of the effect of skewness and kurtosis on the slope of the volatility surface. Chang, Christoffersen, and Jacobs (2009) also find a large negative correlation between risk-neutral skewness and kurtosis, a price of skewness risk robust to different specifications, but no robust pricing result on kurtosis in a cross-sectional study. This supports our conclusion gleaned from estimating the alternative special cases and also supports the preference-based restrictions associated with our maintained specification in column 2.

A.2. Importance of Nonlinear Pricing of Jumps

To demonstrate the importance of having a nonlinear pricing structure for jumps in the equity premium, we estimate a model with the full dynamics as in our ‘prudence’ model but with the equity premium specified as follows:

$$m_t = \psi_v \sigma_t^2 + \psi_j \lambda_t. \quad (37)$$

This linear pricing specification is typically assumed in the studies that use both equity and options to identify the pricing of jump risk (e.g., Pan (2002)). We label this specification of the equity premium as the ‘linear’ model.

Table III presents a comparison of the results for the linear versus our nonlinear specification of the equity premium. The parameter estimates for the GARCH and jump dynamics

¹³As derived in our appendix, the coefficient on the variance for the risk premium, associated with simple as opposed to continuously compounded returns, is $100 \times \psi_v + 0.5$.

of the ‘linear’ model are quite similar to those in the ‘prudence’ model. Therefore we only report the estimates of the pricing coefficients to facilitate comparison. In the ‘linear’ model, we have $\hat{\psi}_j = 0.185$ (t-stat=2.57) and $\hat{\psi}_v = 0.017$ (t-stat=1.29), i.e., jump risk is significantly priced whereas diffusive risk is not. Furthermore, the likelihood (-24582) of the ‘linear’ model is much smaller than that (-24577) of the ‘prudence’ model, despite the rich dynamics in the ‘linear’ model. The pricing result of the ‘linear’ model is quite similar to what Pan (2002) finds in a continuous time framework with stochastic volatility and time-varying jumps. In addition, both equity and options are used in Pan (2002).

[Insert Table III about here.]

This comparison highlights the importance of having a nonlinear pricing structure for jumps in the equity premium. In particular, having jumps priced through higher-order moments not only works well empirically but, theoretically, is also consistent with more general preferences as derived in section VII.C.

A.3. Alternative Specifications of Dynamics

Table IV provides estimates for alternative specifications of the volatility and jump dynamics. For example, removing jumps from our parameterization results in the special cases of either a one-component or a two-component GARCH-in-Mean specification (columns 2 and 3 respectively).¹⁴ As evident from the resulting likelihoods in Table IV, compared with our maintained model in columns 2 or 3 of Table II, including jumps results in a much superior fit. As reported in Panel B of Table V, the likelihood ratio test (LRT) statistics of 1872.4 and 1459.2 reject the no-jump 1-component and 2-component GARCH-in-Mean special cases with extremely small p-values.

[Insert Table IV about here.]

Nevertheless, even with our most general specification of jumps, the 2nd GARCH component is still very important; the LRT reported in Panel B of Table V comparing column 3 of Table II with a special case without the 2nd GARCH component has a p-value of 4.13e-52.

The asymmetric feedback from jumps to diffusive volatility is also very important as revealed by the test of the restriction $\alpha_{a,j,1} = \alpha_{a,j,2} = 0$ reported in the first line of Panel B of Table V.

Column 4 of Table IV shows the results of assuming constant jump arrival. The LRT reported in Panel B of Table V decisively rejects constant jump arrival in favor of our autoregressive parameterization of the jump arrival process (p-value of 3.74e-23).

¹⁴As shown in equations (22) to (24), if λ_t is zero for all t the conditional skewness and kurtosis terms drop out.

[Insert Table V about here.]

Finally, column 5 of Table IV confirms that the equity premium pricing structure is robust to the specification without an AR(2) structure in the conditional mean. The pricing coefficients are similar to those of our maintained specification in columns 2 or 3 of Table II: both ψ_v and ψ_s are significantly priced and ψ_k is close to 0.

B. Risk and the Equity Premium

Jumps contribute to the dynamics of the total conditional variance and also drive the dynamics of conditional skewness and conditional kurtosis in our maintained model. The results for that model, reported in column 2 of Table II, reveal a positive coefficient on the variance and negative coefficient on skewness. Apparently jumps contribute to the pricing of the equity premium. However, since both jumps and the GARCH volatility enter our parameterizations of the conditional variance, conditional skewness and conditional kurtosis, the net contribution of jumps versus GARCH volatility to the dynamics of equity premium is difficult to disentangle. To this end, in this subsection we focus on these two contributions to risk and attempt to isolate their net effects on the equity premium.

Using the variance forecast as a measure of risk for the market as a whole follows the long tradition from Merton (1980). Given that we estimate constant moments for the jump-size distribution, the contribution of jumps to the dynamics of the equity premium is driven by λ_t . A larger λ_t indicates a larger probability of a jump event. Jump events are generally realizations in the tails of the distribution and, according to our estimate of the jump-size distribution, they are more likely to be the left tail.

We compute the marginal effect of these two components using the partial derivative of the equity premium m_t with respect to λ_t and σ_t respectively. Figure 2 displays the partial derivative of m_t with respect to λ_t for a range of empirically realistic values of λ_t . This is done for three different levels of the GARCH volatility component. Note that the equity premium always increases in response to an increase in jump risk (λ_t). However, the size of the effect differs depending on the level of the GARCH volatility and the current value of λ_t . A unit increase in λ_t yields the largest increase in the premium when the GARCH volatility is low and when jump activity is expected to be low. In more volatile times, as measured by larger σ_t and/or larger λ_t , the effect of an *increase* in jump risk is still positive but much smaller.

[Insert Figure 2 about here.]

In order to get some idea of the magnitude of the effect of jump risk on the equity premium m_t , we compute the partial derivative of m_t with respect to λ_t at each t . The sample mean of this derivative is around 0.1057, suggesting that one more jump in a year increases the equity premium by 0.1057%. On average we estimate that there are about 37.5 jumps a year.

Therefore, based on our maintained model reported in column 2 of Table II, the contribution of jumps to the market equity premium is about 3.96% per annum. Note that Pan (2002) and Elkamhi and Ornathanalai (2009) report jump risk premia of approximately 3.5% and 3.18% respectively by jointly estimating return dynamics and option dynamics. Bollerslev and Todorov (2011) obtain a median jump risk premium of 5.2% using options data and a nonparametric approach to estimate realized jumps from high-frequency data.

Figure 3 reports the partial derivative of the equity premium with respect to σ_t . Interestingly, in contrast to λ_t , an increase in GARCH volatility has a negative effect on the premium for small values of σ_t and a positive effect for average to larger values of GARCH volatility. The negative effect originates from the fact that an increase in σ_t increases both the conditional variance and the skewness (decreases the negative skewness).¹⁵ Therefore, our results suggest that in relatively calm times (small σ_t), an increase in σ_t has a stronger effect on the equity premium from our (normalized) skewness measure than from the variance. Intuitively, investors are more excited about the increase in skewness (s_t has a smaller negative value) and the potential increase in upside, therefore they demand a smaller equity premium. In more volatile times, the equity premium effect of σ_t through the variance channel dominates that from the skewness and investors demand a higher equity premium for an increase in σ_t . Note that average σ_t for our sample is 0.75 which is close to where the partial derivative changes from negative to positive.

[Insert Figure 3 about here.]

We also present a numerical example in Table VI to help capture the intuition. To illustrate, we set the jump intensity at the sample average. This table presents the results on two low volatility days. Suppose on one day the diffusive volatility is 0.07 (annualized), then according to our ‘prudence’ model estimation, the equity premium from conditional variance is 2.1% (annualized) and the equity premium from the conditional skewness is 7.5% (annualized), resulting in a total equity premium of 9.6% (annualized) for this particular day. It is clear that the skewness component dominates in this case because the magnitude of the skewness is quite high (-1.1). Skewness at the market level illustrates the left-tail risk as measured by the distributional asymmetry relative to the dispersion of the distribution (where the latter is

¹⁵This can be verified from equations (22) to (24) and the fact that the estimated jump-size mean θ is negative. That is:

$$\begin{aligned} \partial m / \partial \sigma &= \partial \left(\hat{\psi}_v \times \hat{v} + \hat{\psi}_s \times \hat{s} + \hat{\psi}_k \times \hat{k} \right) / \partial \sigma \\ &\approx \partial \left(\hat{\psi}_v \times \hat{v} + \hat{\psi}_s \times \hat{s} \right) / \partial \sigma \\ &= \underbrace{\hat{\psi}_v}_{>0} \times \underbrace{\partial \hat{v} / \partial \sigma}_{>0} + \underbrace{\hat{\psi}_s}_{<0} \times \underbrace{\partial \hat{s} / \partial \sigma}_{>0} \end{aligned}$$

Note that $\hat{s} < 0$ since $\hat{\theta} < 0$; so $\partial \hat{s} / \partial \sigma > 0$ refers to a decrease in negative skewness.

measured by volatility). When the dispersion of the distribution is low, a moderate amount of jump risk (to the left tail on average) could generate a significant asymmetry in the distribution. When the dispersion increases, e.g., σ increases from 0.07 to 0.1, the magnitude of the skewness drops significantly to -0.5 , resulting in a total equity premium of 7.1%, despite an increase in the equity premium due to the variance (from 2.1% to 3.5%). Therefore, if we omit the skewness component of the equity premium, an increase of σ from 0.07 to 0.1 corresponds to a decrease in the equity premium, m , from 9.6% to 7.1%. In contrast, in more volatile times an increase in σ will correspond to an increase in m because the equity premium will then be dominated by the variance. Even if the jump risk were higher in more volatile times, the return distribution could actually be less asymmetric than in less volatile times so that the variance component of the equity premium dominates.

[Insert Table VI about here.]

These results offer a potential resolution to the conflicting results in the literature on risk and expected return for the market as a whole. In this literature, higher-order moments are not considered as part of the risk. A positive relationship between conditional variance and return only occurs when the GARCH variance component is at or above average levels. During calm times (low level of the GARCH variance component), the skewness premium effect dominates. For these periods if we were to omit conditional skewness from the equity premium specification, due to the missing skewness factor we could, inappropriately, estimate a negative relationship between the equity premium and the conditional variance at low levels of the latter. In more volatile times, the variance premium effect dominates and we will be able to see a positive risk-variance tradeoff, whether we include conditional skewness in the equity premium specification or not.

C. Equity Premium Size and Dynamics

Table VII presents the descriptive statistics of the equity premium from our maintained model in column 2 of Table II, as well as those for the unrestricted model in column 4 of the same table. The median (average) equity premium estimate is about 7.8% (9.8%) per annum for the maintained prudence model and 7.1% (9.6%) for the unrestricted parameterization.¹⁶ Figures 4 and 5 illustrate the time-series dynamics of the equity premium for the maintained and unrestricted specifications respectively. Although the time-series pattern of the equity premium is quite similar for the two specifications, the unrestricted parameterization has larger outliers and some negative values. In other words, the restriction implied from preference theory for the prudence model ensures that the (expected) equity premium is always positive.

¹⁶Daily returns were scaled by 100. Therefore, to annualize the median or average daily premiums reported in Table VII we scale by $(252/100)$.

Notice that the average premium will be affected by a few large outliers associated particularly with 1987 and the 1930s. Note though, as mentioned above, our results with respect to the significance of the variance and skewness components of the equity premium are robust for the post-1930s (1940-2007) subsample.

[Insert Table VII about here.]

According to our parameter estimates, the average expected number of jumps per year is 37.5. Combining this with the average impact of a change in jump risk implies that the equity premium associated with jumps is about 3.96% per annum on average.

In our parameterization all higher-order moments can incorporate jumps. According to our parameter estimates for the maintained model (column 2 of Table II), using the average of the estimated jump arrival rate $\bar{\lambda}$, jumps contribute 1.12% to the equity premium through the variance dynamics¹⁷ and 2.84% associated with the skewness premium.

In addition to the significant pricing of variance with respect to the market equity premium, we find robust pricing of skewness for the market equity premium. The equity premium contributed by the premium for skewness is on average 3.5%, which contributes about 36% of the overall equity premium for our sample.¹⁸ This is very close to the 3.6% per annum risk premium compensation for systematic skewness found by Harvey and Siddique (2000) who study the conditional skewness in a cross-section of monthly stock returns. In our parameterization of the time-series of daily market excess returns, jumps account for about 81% of that skewness premium.

As noted above, when we impose the preference restriction of a non-negative price associated with kurtosis risk, our findings show that the price of kurtosis risk is close to zero. We find that this is due to the high negative correlation between the conditional skewness and conditional kurtosis in the market index. Conditional kurtosis is significantly priced with a positive sign when the skewness factor is not included. In other words, at least at the market level, the skewness and kurtosis are likely to be representing the same source of risk, that of the crash risk.

[Insert Figure 4 and Figure 5 about here.]

Table VIII reports summary statistics for our estimated higher-order moments for both the maintained prudence model and the unrestricted model (columns 2 and 4 of Table II, respectively) which are plotted in Figures 6 and 7.

[Insert Table VIII about here.]

[Insert Figure 6 and Figure 7 about here.]

¹⁷ $\psi_v \times (\bar{\lambda} \times (\theta^2 + \delta^2)) \times 2.52 = 0.027 \times (0.149 \times ((-0.467)^2 + 0.942^2)) \times 2.52.$

¹⁸We have $\psi_s = -0.028$, average $s_t = -0.496$. Therefore the skewness premium is $(-0.028) \times (-0.496) \times \frac{252}{100} = .035$ or 3.5% per annum.

V. Out-of-Sample Asset Allocation Performance

In this section, we evaluate the value-added associated with out-of-sample forecasts that incorporate priced risks associated with time-varying arrival of jumps through their effect on the dynamics of conditional variance, skewness and kurtosis. We do this by evaluating the realized utility and certainty-equivalent returns associated with a simple portfolio allocation application. As our model is on the market index, we assume that an investor is making investment decisions between the market portfolio and the risk-free asset. We derive optimal portfolio weights using the forecasts associated with our maintained model versus several alternative benchmarks.

Building on Harvey and Siddique (2000) and Guidolin and Timmermann (2008), we use a Taylor-series approximation to a general utility function which is consistent with the pricing kernel used to derive our maintained risk premium specification. Determining the additional parameters associated with a more general utility function is challenging. However, the equity premium associated with our maintained *prudence* model, summarized in section II.E, provides additional parameters (prices of risk) associated with the higher-order terms in the risk premium specification. As shown in section VII.C (in the appendix), we obtain the functional relationship between these parameters and the coefficients for the Taylor-series expansion of a general utility function. We can then solve for the coefficients of that approximation to general utility and calibrate the implied utility coefficients to the empirical estimates associated with our equity premium specification. This provides an approximation to general utility for our performance application that is consistent with our asset pricing model and empirical estimates.

We compare the out-of-sample portfolio performance based on four different models of the market index return. The four models are the *prudence* model (our maintained model), the *variance* model, the *GIM-1* model, and a *constant* model which assumes a constant equity premium. The *prudence*, *variance*, and *constant* models share the same jump dynamics and associated time-varying variance, skewness and kurtosis. The *prudence* and *variance* models are reported in columns 2 and 5 of Table II. The *GIM-1* model, which is the traditional single-component GARCH-in-mean model, is reported in column 2 of Table IV. For each of the models, we estimate the parameters using the data up to the end of 2001; using data from 2002 to the end of 2007 to evaluate the out-of-sample performance.

The out-of-sample portfolio performance results are reported in Table IX. For comparison with other studies, such as Guidolin and Timmermann (2008), the asset allocation performance results are performed using daily returns scaled to a monthly equivalent. As is clear from Table IX, the *prudence* model dominates the other three benchmark models. It results in higher average realized utility, lower standard deviation of realized utility, and a higher

annualized certainty equivalent return (CEQ).¹⁹ There are clear benefits to pricing jumps (*prudence* versus *GIM-1*), and to pricing higher-order moments (*prudence* versus *variance*).

VI. Concluding Comments

In this paper we demonstrate that jump risk is priced in the market index and contributes to the equity premium. Jump risk potentially gets priced through the time-varying conditional variance, skewness and kurtosis. The time-varying conditional moments of excess returns are generated by the time-varying jump arrival process and two-component GARCH dynamics.

Empirically, we find that the conditional variance and skewness are priced in the market equity premium. Our results highlight the importance of having a nonlinear pricing structure for jumps in the equity premium. In particular, having jumps priced through higher-order moments not only works well empirically but, theoretically, is also consistent with more general preferences. The results are robust to different specifications of the equity premium and/or the volatility and jump dynamics.

We compute the marginal effect of jumps on the equity premium and find that on average 1 more jump per year would increase the equity premium by 0.1057%. The average number of jumps per year according to our parameterization and sample is 37.5 which implies an overall jump risk premium of 3.96% per annum. Jump risk is priced in the market equity premium both through the jump component of variance dynamics and skewness contributing 1.12% and 2.84%, respectively, to those premiums.

The total skewness premium is about 3.5% per annum. Our maintained model imposes a nonnegative coefficient on conditional kurtosis as implied by decreasing absolute prudence preferences. In this case, we find the price of the kurtosis to be zero in the presence of the skewness risk. We explore different specifications of the equity premium by pricing skewness and kurtosis separately and find them to have the expected signs. We further investigate the correlation between kurtosis and skewness and find a strong negative correlation. These findings suggest that at least at the market level, the skewness and kurtosis seem to represent the same source of market crash risk.

Our study has some important implications for the literature on estimation of the risk-return relationship with high frequency data. Identifying jump risk and the associated intertemporal risk-return tradeoff has been found to be difficult in prior studies unless one also uses options data. Our results show that it is possible to identify jump risk and its contribution to the market risk-return tradeoff using equity data only.

¹⁹As in Guidolin and Timmermann (2008), the certainty equivalent return is calculated based on the average realized utility.

VII. Appendix

A. Scaling Returns

Empirically, we find that we need to scale the daily returns r_t in order to get numerically stable estimates. Assuming that $\rho_1 = \rho_2 = 0$ and suppressing the time index for convenience of notation, recall from section II that

$$r = m + \epsilon_1 + \epsilon_2, \quad (38)$$

$$m = \mu + \psi_v v + \psi_s s + \psi_k k, \quad (39)$$

where $\epsilon_1 \sim N(0, \sigma^2)$, $\epsilon_2 = J - \lambda\theta$, and $J = \sum_{k=1}^n Y_k$ follows a compound Poisson distribution with parameters $(\lambda, \theta, \delta)$. λ is the arrival rate for the Poisson distributed variate n and the jump size is distributed normally as $Y_k \sim N(\theta, \delta^2)$.

Scaling r by 100 and denoting the scaled return by r_{100} ,

$$r_{100} = m_{100} + \epsilon_{1,100} + \epsilon_{2,100}, \quad (40)$$

in which

$$m_{100} \equiv 100m = 100(\mu + \psi_v v + \psi_s s + \psi_k k), \quad (41)$$

$$\epsilon_{1,100} \equiv 100\epsilon_1, \quad (42)$$

$$\epsilon_{2,100} \equiv 100\epsilon_2. \quad (43)$$

Estimating equation (40) is equivalent to estimating equation (38) using MLE.

One can verify that

$$100\epsilon_1 \sim N(0, 100^2\sigma^2) \quad (44)$$

Also, $100\epsilon_2$ follows a compound Poisson distribution with parameters $(\lambda, 100\theta, 100\delta)$. In particular, the Poisson distribution that governs the number of jumps does not change whereas the jump size is scaled by 100. That is,

$$100Y_k \sim N(100\theta, 100^2\delta^2) \quad (45)$$

First we define the parameters of $\epsilon_{1,100}$ and $\epsilon_{2,100}$ in the following way:

$$\sigma_{100} = 100\sigma \quad (46)$$

$$\lambda_{100} = \lambda \quad (47)$$

$$\theta_{100} = 100\theta \quad (48)$$

$$\delta_{100} = 100\delta, \quad (49)$$

where the parameters with subscript 100 are for the scaled returns r_{100} . After the reparametrization, $\epsilon_{1,100} \sim N(0, \sigma_{100}^2)$ and $\epsilon_{2,100}$ follows a compound Poisson distribution with parameters $(\lambda_{100}, \theta_{100}, \delta_{100})$ and $Y_{k,100} \sim N(\theta_{100}, \delta_{100}^2)$.

Using equations (46) to (49), it is straightforward to see that

$$v_{100} = \sigma_{100}^2 + \lambda_{100}(\theta_{100}^2 + \delta_{100}^2) \quad (50)$$

$$= 100^2\sigma^2 + \lambda(100^2\theta^2 + 100^2\delta^2) = 10000v, \quad (51)$$

$$s_{100} = \frac{\lambda_{100}(\theta_{100}^3 + 3\theta_{100}\delta_{100}^2)}{(\sigma_{100}^2 + \lambda_{100}\theta_{100}^2 + \lambda_{100}\delta_{100}^2)^{3/2}} = s, \quad (52)$$

$$k_{100} = 3 + \frac{\lambda_{100}(\theta_{100}^4 + 6\theta_{100}^2\delta_{100}^2 + 3\delta_{100}^4)}{(\sigma_{100}^2 + \lambda_{100}\delta_{100}^2 + \lambda_{100}\theta_{100}^2)^2} = k. \quad (53)$$

Note that in the estimation we only use the scaled return r_{100} , therefore in equation (41)

$$m_{100} = 100(\mu + \psi_v v + \psi_s s + \psi_k k)$$

we need to replace v with v_{100} using $v_{100} = 10000v$, replace s with s_{100} using $s_{100} = s$, and replace k with k_{100} using $k_{100} = k$. After this operation we have

$$m_{100} = 100\mu + 0.01\psi_v v_{100} + 100\psi_s s_{100} + 100\psi_k k_{100}. \quad (54)$$

Therefore, we scale the parameters in m_{100} as

$$\mu_{100} = 100\mu, \quad (55)$$

$$\psi_{v,100} = 0.01\psi_v, \quad (56)$$

$$\psi_{s,100} = 100\psi_s, \quad (57)$$

$$\psi_{k,100} = 100\psi_k. \quad (58)$$

and obtain

$$m_{100} = \mu_{100} + \psi_{v,100}v_{100} + \psi_{s,100}s_{100} + \psi_{k,100}k_{100}. \quad (59)$$

Combining all of the rescaled terms:

$$r_{100} = \mu_{100} + \psi_{v,100}v_{100} + \psi_{s,100}s_{100} + \psi_{k,100}k_{100} + \epsilon_{1,100} + \epsilon_{2,100} \quad (60)$$

B. Simple Risk Premium

Equation (31) specifies the conditional equity risk premium associated with excess continuously compounded returns for a market-wide index. Often the equity premium is computed using simple returns, that is, compounded per period rather than continuously.

For our specification, the conditional simple equity risk premium, labelled as $m_{s,t}$ is:

$$m_{s,t} = E[\exp(r_t - \rho_1(r_{t-1} - m_{t-1}) - \rho_2(r_{t-2} - m_{t-2}))|\Phi_{t-1}] - 1 \quad (61)$$

$$= E[\exp(m_t + \epsilon_{1,t} + \epsilon_{2,t})|\Phi_{t-1}] - 1. \quad (62)$$

To derive this conditional simple equity premium $m_{s,t}$, we need to find $E[\exp(\epsilon_{1,t})|\Phi_{t-1}]$ and $E[\exp(\epsilon_{2,t})|\Phi_{t-1}]$. Using the characteristic function of the normal distribution and the fact that $\epsilon_{1,t}|\Phi_{t-1} \sim N(0, \sigma_t^2)$,

$$E[\exp(\epsilon_{1,t})|\Phi_{t-1}] = \exp(0.5\sigma_t^2). \quad (63)$$

Recall from section II.B.1 that the compensated jump innovation is

$$\epsilon_{2,t} = J_t - \theta\lambda_t = \sum_{k=1}^{n_t} Y_{t,k} - \theta\lambda_t,$$

and $J_t = \sum_{k=1}^{n_t} Y_{t,k}$ is directed by a compound Poisson distribution with parameters $(\lambda_t, \theta, \delta)$. Note that

$$\begin{aligned} E[\exp(\epsilon_{2,t})|\Phi_{t-1}] &= E[\exp(J_t)|\Phi_{t-1}] \exp(-\theta\lambda_t) \\ &= \exp(\lambda_t(E[\exp(Y)|\Phi_{t-1}] - 1) - \theta\lambda_t) \end{aligned}$$

As noted in equation (13), the jump-size distribution is normal, in which case

$$E[\exp(Y)|\Phi_{t-1}] = \exp(\theta + \frac{1}{2}\delta^2),$$

so that

$$E[\exp(\epsilon_{2,t})|\Phi_{t-1}] = \exp(\lambda_t(\exp(\theta + \frac{1}{2}\delta^2) - 1 - \theta)) = \exp(\lambda_t\xi). \quad (64)$$

Therefore, the conditional risk premium $m_{s,t}$ is

$$E[\exp(r_t - \rho_1(r_{t-1} - m_{t-1}) - \rho_2(r_{t-2} - m_{t-2}))|\Phi_{t-1}] - 1 \quad (65)$$

$$\begin{aligned} &= (\exp(m_t)E[\exp(\epsilon_{1,t})|\Phi_{t-1}]E[\exp(\epsilon_{2,t})|\Phi_{t-1}]) - 1 \\ &= (\exp(\mu + \psi_v v_t + \psi_s s_t + \psi_k k_t + 0.5\sigma_t^2 + \lambda_t\xi) - 1) \end{aligned} \quad (66)$$

in which

$$\xi = \exp(\theta + \frac{\delta^2}{2}) - 1 - \theta.$$

Written in terms of returns scaled by 100, this is the daily risk premium expressed as a percentage.

C. General Utility Approximation

This appendix shows how we obtain the coefficients of an approximation to a general utility function implied by our estimates of the parameters of the associated risk premium specification.

Building on Harvey and Siddique (2000) and Guidolin and Timmermann (2008), we begin by taking an N^{th} -order Taylor-series expansion of a general utility function, setting $N = 4$, so that

$$U(W_{t+1}) \approx \sum_{n=0}^4 \frac{U^{(n)}(W_t)}{n!} (W_{t+1} - W_t)^n = \sum_{n=0}^4 \frac{U^{(n)}(W_t)}{n!} (W_t R_{t+1}^W)^n, \quad (67)$$

in which W_t is the known initial wealth level about which the utility is expanded, $W_{t+1} = W_t(1 + R_{t+1}^W)$, and R^W is the net return on aggregate wealth. Correspondingly the Taylor-series expansion of marginal utility is

$$U'(W_{t+1}) \approx \sum_{n=0}^3 \frac{U^{(n+1)}(W_t)}{n!} (W_{t+1} - W_t)^n = \sum_{n=0}^3 \frac{U^{(n+1)}(W_t)}{n!} (W_t R_{t+1}^W)^n. \quad (68)$$

Without loss of generality, we normalize the constant initial wealth level so that $W_t = 1$.

This implies that the pricing kernel, $M_{t+1} \equiv \frac{U'(W_{t+1})}{U'(W_t)}$, is

$$M_{t+1} \approx g_{0t} + g_{1t}R_{t+1}^W + g_{2t}(R_{t+1}^W)^2 + g_{3t}(R_{t+1}^W)^3. \quad (69)$$

in which $g_{nt} = \frac{U^{(n+1)}}{U'} \frac{1}{n!}$ for $n = 0, 1, 2, 3$.

As in Dittmar (2002), $U' > 0$, $U'' < 0$, $U''' > 0$, and $U'''' < 0$, that is, positive marginal utility, risk aversion, decreasing absolute risk aversion, and decreasing absolute prudence, respectively, imply that $g_{1t} < 0$, $g_{2t} > 0$, $g_{3t} < 0$ while $g_{0t} = 1$. Since the utility approximation is truncated at $N = 4$, $g_{nt} = 0$ for $n > 3$.

Using the fundamental pricing equation:

$$E_t[(1 + R_{t+1}^W)M_{t+1}] = 1, \quad (70)$$

with Equation (69), and assuming the existence of a riskfree rate, we have a four-moment asset pricing model, as in Guidolin and Timmermann (2008),

$$E_t[R_{t+1}^W] - R_t^f = \kappa_{1t}Cov_t(R_{t+1}^W, R_{t+1}^W) + \kappa_{2t}Cov_t(R_{t+1}^W, (R_{t+1}^W)^2) + \kappa_{3t}Cov_t(R_{t+1}^W, (R_{t+1}^W)^3), \quad (71)$$

where $\kappa_{nt} = -g_{nt}(1 + R_t^f)$ and R_t^f is the net return on the risk free asset. The above preference assumptions imply that $\kappa_{1t} > 0$, $\kappa_{2t} < 0$, $\kappa_{3t} > 0$.

For notational convenience, in the following derivation we define the excess return on market index as $R_{t+1} \equiv R_{t+1}^W - R_t^f$ and abstract from the autocorrelation terms in our empirical

implementation, so that:

$$\begin{aligned} m_t &\equiv E_t[R_{t+1}] = \kappa_{1t}\text{Var}_t(R_{t+1} + R_t^f) + \kappa_{2t}\text{Cov}_t(R_{t+1} + R_t^f, (R_{t+1} + R_t^f)^2) \\ &\quad + \kappa_{3t}\text{Cov}_t(R_{t+1} + R_t^f, (R_{t+1} + R_t^f)^3). \end{aligned} \quad (72)$$

Rearranging the above equation, and given that R^f is conditionally non random, we have

$$m_t = E_t[R_{t+1}] = a_t\text{Var}_t(R_{t+1}) + b_t\text{Cov}_t(R_{t+1}, R_{t+1}^2) + c_t\text{Cov}_t(R_{t+1}, R_{t+1}^3), \quad (73)$$

where

$$\begin{aligned} a_t &\equiv \kappa_{1t} + 2\kappa_{2t}R_t^f + 3\kappa_{3t}(R_t^f)^2, \\ b_t &\equiv \kappa_{2t} + 3\kappa_{3t}R_t^f, \\ c_t &\equiv \kappa_{3t}. \end{aligned}$$

One can verify that

$$\text{Cov}_t(R_{t+1}, R_{t+1}^2) = E_t[R_{t+1}^3] - E_t[R_{t+1}]E_t[R_{t+1}^2], \quad (74)$$

$$c_{3t} \equiv E_t[(R_{t+1} - E_t[R_{t+1}])^3] = E_t[R_{t+1}^3] - E_t[R_{t+1}]E_t[R_{t+1}^2] - 2E_t[R_{t+1}]\text{Var}_t(R_{t+1}). \quad (75)$$

Therefore we have

$$\text{Cov}_t(R_{t+1}, R_{t+1}^2) = c_{3t} + 2E_t[R_{t+1}]\text{Var}_t(R_{t+1}). \quad (76)$$

Similarly we can get

$$\text{Cov}_t(R_{t+1}, R_{t+1}^3) = E_t[R_{t+1}^4] - E_t[R_{t+1}]E_t[R_{t+1}^3], \quad (77)$$

$$c_{4t} \equiv E_t[(R_{t+1} - E_t[R_{t+1}])^4] = E_t[R_{t+1}^4] - E_t[R_{t+1}]E_t[R_{t+1}^3] \quad (78)$$

$$- \frac{3}{2}E_t[R_{t+1}][c_{3t} + \text{Cov}_t(R_{t+1}, R_{t+1}^2)], \quad (79)$$

so that

$$\text{Cov}_t(R_{t+1}, R_{t+1}^3) = c_{4t} + \frac{3}{2}E_t[R_{t+1}][c_{3t} + \text{Cov}_t(R_{t+1}, R_{t+1}^2)]. \quad (80)$$

Plugging Equations (76) and (80) into Equation (73) we have

$$m_t = a_t\text{Var}_t(R_{t+1}) + b_t(c_{3t} + 2m_t\text{Var}_t(R_{t+1})) + c_t \left[c_{4t} + \frac{3}{2}m_t(c_{3t} + \text{Cov}_t(R_{t+1}, R_{t+1}^2)) \right]. \quad (81)$$

Solving the above equation for m_t gives

$$m_t = \frac{a_t}{d_t}v_t + \frac{b_t}{d_t}v_t^{3/2}s_t + \frac{c_t}{d_t}v_t^2k_t, \quad (82)$$

where

$$\begin{aligned}
v_t &\equiv \text{Var}_t(R_{t+1}) \\
d_t &\equiv 1 - 2b_tv_t - \frac{3}{2}c_t[c_{3t} + \text{Cov}_t(R_{t+1}, R_{t+1}^2)], \\
s_t &\equiv \frac{c_{3t}}{v_t^{3/2}}, \\
k_t &\equiv \frac{c_{4t}}{v_t^2}.
\end{aligned}$$

Note that v_t , s_t , and k_t are the conditional variance, conditional skewness, and conditional kurtosis, respectively. Let

$$\psi_v \equiv \frac{a_t}{d_t}, \quad (83)$$

$$\psi_s \equiv \frac{b_tv_t^{3/2}}{d_t}, \quad (84)$$

$$\psi_k \equiv \frac{c_tv_t^2}{d_t} \quad (85)$$

then we have

$$m_t = \psi_v v_t + \psi_s s_t + \psi_k k_t. \quad (86)$$

Our in-sample estimation of the coefficient ψ_k associated with kurtosis is close to 0 (as it is for the full-sample results for our maintained prudence model reported in the column titled ‘prudence’ in Table II). As $d_t \neq 0$ we must have $c_t = 0$, in other words, $\kappa_{3t} = 0$. Therefore to calibrate the coefficients $\frac{U^{(n+1)}}{U'}$ from the risk premium estimates, we assume $N = 3$ for the utility approximation. With $\kappa_{3t} = 0$, some of the parameters can be simplified as follows:

$$a_t = \kappa_{1t} + 2\kappa_{2t}R_t^f, \quad (87)$$

$$b_t = \kappa_{2t}, \quad (88)$$

$$\psi_v = \frac{a_t}{1 - 2b_tv_t}, \quad (89)$$

$$\psi_s = \frac{b_tv_t^{3/2}}{1 - 2b_tv_t}. \quad (90)$$

Solving the above equations we have

$$\kappa_{2t} = \frac{\psi_s}{2\psi_s v_t + v_t^{3/2}}, \quad (91)$$

$$\kappa_{1t} = \psi_v(1 - 2\kappa_{2t}v_t) - 2\kappa_{2t}R_t^f, \quad (92)$$

and, given the relationship, $\kappa_{nt} = -g_{nt}(1 + R_t^f) = -\frac{U^{(n+1)}}{U'} \frac{1}{n!}(1 + R_t^f)$, implied by the connection

between the risk premium specification and the pricing kernel,

$$\frac{U^{(2)}}{U'} = -\frac{\kappa_{1t}}{1 + R_t^f}, \quad (93)$$

$$\frac{U^{(3)}}{U'} = -\frac{2\kappa_{2t}}{1 + R_t^f}. \quad (94)$$

Therefore, we can calibrate the utility coefficients on the LHS of the Equations (93) and (94) using the functional relationship with the estimated parameters given by the RHS of Equations (91) and (92). In our out-of-sample portfolio allocation evaluation, we fix v_t and R_t^f on the RHS of Equations (91) and (92) at their sample averages.

Note that the utility coefficients given by Equations (93) and (94) are underdetermined for the coefficients required for the approximation to the level of utility as expressed in Equation (67). However, without loss of generality, since initial wealth is fixed and normalized to $W_t = 1$, we can set $U'(W_t)$ to a constant and solve for $U^{(2)}(W_t)$ and $U^{(3)}(W_t)$ using Equations (93) and (94).

REFERENCES

- Andersen, Torben G., Luca Benzoni, and Jesper Lund, 2002, An empirical investigation of continuous-time equity return models, *Journal of Finance* 62, 1239–1284.
- Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.
- Bakshi, Gurdip, and Dilip Madan, 2006, A theory of volatility spreads, *Management Science* 52, 1945–1956.
- Bansal, Ravi, and S. Viswanathan, 1993, No-arbitrage and arbitrage pricing: A new approach, *Journal of Finance* 48, 1231–1262.
- Bates, David S., 2000, Post-'87 crash fears in the s&p 500 futures option market, *Journal of Econometrics* 94, 181–238.
- , and Roger Craine, 1999, Valuing the futures market clearinghouse's default exposure during the 1987 crash, *Journal of Money, Credit, and Banking* 31, 248–272.
- Bekaert, Geert, and Stephen F. Gray, 1998, Target zones and exchange rates: An empirical investigation, *Journal of International Economics* 45, 1–35.
- Bollerslev, Tim, and Viktor Todorov, 2011, Tails, fears and risk premia, *Journal of Finance* forthcoming.
- Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press).

- Chan, Wing Hong, and Liling Feng, 2009, Extreme news events, long-memory volatility, and time varying risk premia in stock market returns, Manuscript.
- Chan, Wing Hong, and John M. Maheu, 2002, Conditional jump dynamics in stock market returns, *Journal of Business & Economic Statistics* 20, 377–389.
- Chang, Bo Young, Peter Christoffersen, and Kris Jacobs, 2009, Market skewness risk and the cross-section of stock returns, Manuscript, McGill University.
- Chapman, David, 1997, Approximating the asset pricing kernel, *Journal of Finance* 52, 1383–1410.
- Chernov, Mikhail, Ronald A. Gallant, Eric Ghysels, and George Tauchen, 2003, Alternative models for stock price dynamics, *Journal of Econometrics* 116, 225–257.
- Christoffersen, Peter, Kris Jacobs, and Chayawat Ornthanalai, 2008, Exploring time-varying jump intensities: Evidence from s&p500 returns and options, Manuscript, McGill University.
- Conrad, Jennifer, Robert F. Dittmar, and Eric Ghysels, 2009, Ex ante skewness and expected stock returns, Manuscript.
- Das, Sanjiv R., 2002, The surprise element: Jumps in interest rate, *Journal of Econometrics* 106, 27–65.
- Dittmar, Robert F., 2002, Nonlinear pricing kernels, kurtosis preference, and evidence from the cross-section of equity returns, *Journal of Finance* 57, 369–403.
- Duan, Jin-Chuan, and Weiqi Zhang, 2010, Forward-looking market risk premium, Manuscript.
- Elkamhi, Redouane, and Chayawat Ornthanalai, 2009, The market jump risk and the price structure of individual equity options, Manuscript.
- Engle, Robert F., and Gary G. J. Lee, 1999, A long-run and short-run component model of stock return volatility, in Engle, and White, ed.: *Cointegration, Causality, and Forecasting*. Oxford University Press.
- Eraker, Bjorn, Michael S. Johannes, and Nicholas G. Polson, 2003, The impact of jumps in volatility and returns, *Journal of Finance* 58, 1269–1300.
- Fama, Eugene, and Kenneth French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 55, 55–84.
- Friend, Irwin, and Randolph Westerfield, 1980, Co-skewness and capital asset pricing, *Journal of Finance* 35, 897–914.
- Guidolin, Massimo, and Allan Timmermann, 2008, International asset allocation under regime switching, skew, and kurtosis preferences, *Review of Financial Studies* 21, 889–935.
- Harvey, Campbell R., and Akhtar Siddique, 1999, Autoregressive conditional skewness, *Journal of Financial and Quantitative Analysis* 34, 465–487.
- , 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263–1295.

- Hwang, Soosung, and Stephen E. Satchell, 1999, Modeling emerging market risk premia using higher moments, *International Journal of Finance and Economics* 4, 271–296.
- Johannes, Michael, Rohit Kumar, and Nicholas G. Polson, 1999, State dependent jump models: How do U.S. equity indices jump?, Working Paper, Graduate School of Business, University of Chicago.
- Kimball, Miles S., 1990, Precautionary saving in the small and in the large, *Econometrica* 58, pp. 53–73.
- , 1993, Standard risk aversion, *Econometrica* 61, pp. 589–611.
- Kraus, Alan, and Robert Litzenberger, 1978, Skewness preference and the valuation of risk assets, *Journal of Finance* 31, 1085–1100.
- Lim, Kian-Guan, 1989, A new test of the three-moment capital asset pricing model, *Journal of Financial and Quantitative Analysis* 24, 205–216.
- Maheu, John M., and Thomas H. McCurdy, 2004, News arrival, jump dynamics, and volatility components for individual stock returns, *Journal of Finance* 59, 755–794.
- , 2007, Components of market risk and return, *Journal of Financial Econometrics* 5, 560–590.
- Merton, Robert C., 1973, An intertemporal asset pricing model, *Econometrica* 41, 867–888.
- Neely, Christopher J., 1999, Target zones and conditional volatility: The role of realignments, *Journal of Empirical Finance* 6, 177–192.
- Pan, Jun, 1997, Stochastic volatility with resets at jumps, Working Paper, Graduate School of Business, Stanford University.
- , 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.
- Rossi, Alberto, and Allan Timmermann, 2009, What is the shape of the risk-return relation?, Manuscript.
- Sears, Stephen R., and K-C John Wei, 1985, Asset pricing, higher moments, and the market risk premium, *Journal of Finance* 40, 1251–1253.
- Smith, Daniel R., 2007, Conditional coskewness and asset pricing, *Journal of Empirical Finance* 14, 91–119.

Table I
Summary Statistics for Daily Excess Returns r_t

Obs	Mean	StDev	Skewness	Kurtosis	Min	Max
21775	0.023	1.023	-0.439	21.943	-18.823	14.412

The daily excess returns r_t are scaled by 100.

Model specification for parameter estimates in Table II and Table IV:

$$r_t = m_t + \rho_1(r_{t-1} - m_{t-1}) + \rho_2(r_{t-2} - m_{t-2}) + \epsilon_{1,t} + \epsilon_{2,t}$$

$$\lambda_t = E[n_t | \Phi_{t-1}], \quad \lambda_t = \gamma_0 + \gamma_1 \lambda_{t-1} + \gamma_2 \zeta_{t-1},$$

$$\epsilon_{2,t} = \sum_{k=1}^{n_t} Y_{t,k} - \theta \lambda_t, \quad Y_{t,k} \sim N(\theta, \delta^2)$$

$$\epsilon_{1,t} = \sigma_t z_t, \quad z_t \sim NID(0, 1)$$

$$\sigma_t^2 = \sigma_{1,t}^2 + \sigma_{2,t}^2,$$

$$\sigma_{1,t}^2 = \omega + g_1(A_1, \Phi_{t-1}) \epsilon_{t-1}^2 + \beta_1 \sigma_{1,t-1}^2$$

$$\sigma_{2,t}^2 = g_2(A_2, \Phi_{t-1}) \epsilon_{t-1}^2 + \beta_2 \sigma_{2,t-1}^2$$

$$\epsilon_{t-1} = \epsilon_{1,t-1} + \epsilon_{2,t-1}$$

$$g_i(A_i, \Phi_{t-1}) = \exp(\alpha_i + I(\epsilon_{t-1})(\alpha_{a,j,i} E[n_{t-1} | \Phi_{t-1}] + \alpha_{a,i})), \quad \text{for } i = 1, 2;$$

$$I(\epsilon_{t-1}) = 1 \text{ if } \epsilon_{t-1} < 0, \text{ otherwise } 0$$

$$v_t = \sigma_t^2 + \lambda_t(\theta^2 + \delta^2)$$

$$s_t = \frac{\lambda_t(\theta^3 + 3\theta\delta^2)}{(\sigma_t^2 + \lambda_t\theta^2 + \lambda_t\delta^2)^{3/2}}$$

$$k_t = 3 + \frac{\lambda_t(\theta^4 + 6\theta^2\delta^2 + 3\delta^4)}{(\sigma_t^2 + \lambda_t\delta^2 + \lambda_t\theta^2)^2}$$

Column 2 (labelled ‘prudence’):

$$m_t = \psi_v v_t + \psi_s s_t + \psi_k k_t, \quad \psi_s \leq 0, \psi_k \geq 0.$$

Column 3 (labelled ‘intercept’):

$$m_t = \mu + \psi_v v_t + \psi_s s_t + \psi_k k_t, \quad \psi_s \leq 0, \psi_k \geq 0.$$

Column 4 (labelled ‘unrestricted’):

$$m_t = \mu + \psi_v v_t + \psi_s s_t + \psi_k k_t.$$

Table II
Parameter Estimates: Alternative Specifications of the Equity Premium

Parameter	Pricing Models					
	prudence	intercept	unrestricted	variance	skewness	kurtosis
μ		-0.026 (-1.377)	-0.015 (-0.866)	0.022 (3.562)	-0.027 (-1.386)	-0.013 (-0.691)
ψ_v	0.027 (3.270)	0.040 (3.266)	0.051 (4.453)	0.022 (1.991)	0.040 (3.274)	0.031 (2.757)
ψ_s	-0.028 (-3.438)	-0.054 (-2.395)	-0.256 (-3.337)		-0.055 (-2.402)	
ψ_k	0.000 (0.000)	0.000 (0.000)	-0.021 (-2.777)			0.004 (1.931)
ρ_1	0.157 (21.909)	0.157 (21.879)	0.154 (21.382)	0.158 (22.054)	0.157 (21.878)	0.158 (22.030)
ρ_2	-0.048 (-6.958)	-0.049 (-7.000)	-0.051 (-7.325)	-0.047 (-6.788)	-0.049 (-7.003)	-0.048 (-6.872)
ω	0.000 (-0.700)	0.000 (-0.308)	0.000 (0.447)	0.000 (-0.383)	0.000 (-0.296)	0.000 (-0.585)
α_1	-4.798 (-26.895)	-4.829 (-26.577)	-4.874 (-26.840)	-4.780 (-26.905)	-4.827 (-26.620)	-4.798 (-26.878)
β_1	0.974 (357.003)	0.976 (333.603)	0.978 (380.423)	0.973 (355.490)	0.976 (333.507)	0.974 (352.002)
α_2	-8.254 (-3.515)	-8.197 (-3.918)	-8.139 (-3.616)	-8.296 (-3.429)	-8.213 (-3.684)	-8.253 (-3.426)
β_2	0.760 (35.475)	0.767 (38.291)	0.780 (48.587)	0.753 (32.052)	0.767 (38.363)	0.759 (34.476)
$\alpha_{a,j,1}$	-2.131 (-5.956)	-2.357 (-5.000)	-3.220 (-4.339)	-1.935 (-6.601)	-2.368 (-4.965)	-2.077 (-5.882)
$\alpha_{a,1}$	1.603 (8.083)	1.519 (6.693)	1.373 (6.818)	1.656 (9.124)	1.514 (6.657)	1.613 (8.136)
$\alpha_{a,j,2}$	-0.050 (-0.511)	-0.119 (-1.028)	-0.263 (-2.617)	0.002 (0.020)	-0.121 (-1.044)	-0.042 (-0.428)
$\alpha_{a,2}$	6.222 (2.648)	6.266 (2.990)	6.369 (2.830)	6.179 (2.552)	6.284 (2.815)	6.207 (2.573)
γ_0	0.007 (3.458)	0.006 (2.590)	0.003 (2.897)	0.009 (4.103)	0.006 (2.578)	0.008 (3.524)
γ_1	0.949 (89.851)	0.957 (80.449)	0.970 (122.384)	0.941 (88.914)	0.957 (80.488)	0.947 (87.323)
γ_2	0.182 (4.392)	0.152 (3.567)	0.091 (3.831)	0.212 (4.887)	0.151 (3.555)	0.190 (4.337)
θ	-0.467 (-9.529)	-0.460 (-9.107)	-0.446 (-8.148)	-0.490 (-10.050)	-0.460 (-9.095)	-0.476 (-9.719)
δ	0.942 (16.540)	0.976 (14.028)	1.040 (14.747)	0.920 (17.584)	0.978 (13.992)	0.940 (16.291)
lgl	-24577.290	-24576.271	-24571.976	-24580.631	-24576.268	-24578.235

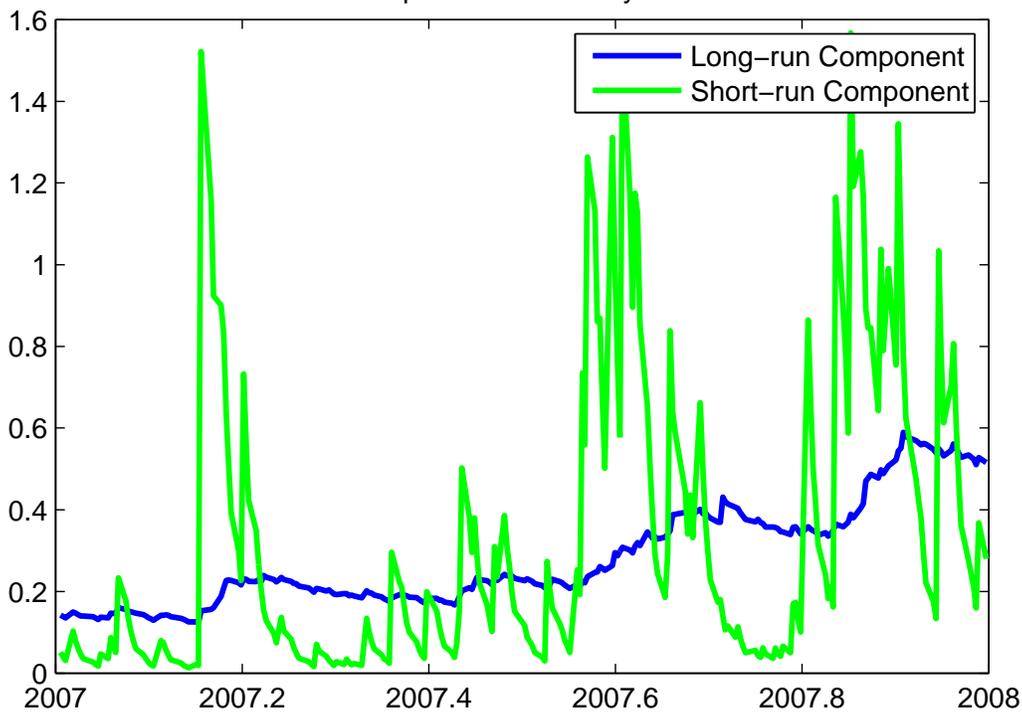
t-stats are in parenthesis. lgl is the loglikelihood. Model specifications are summarized following Table I.

Table III
Importance of Pricing Jumps Nonlinearly

Parameter	Pricing Models	
	prudence	linear
ψ_v	0.027 (3.27)	0.017 (1.29)
ψ_s	-0.028 (-3.44)	
ψ_k	0.000 (0.00)	
ψ_j		0.185 (2.57)
lgl	-24577.29	-24582.35

t-stats are in parenthesis. lgl is the loglikelihood. Model specifications are summarized following Table I. The ‘linear’ model assumes that $m_t = \psi_v \sigma_t^2 + \psi_j \lambda_t$, with other dynamics the same as in the ‘prudence’ model.

Figure 1
Long-run and short-run components of the diffusive volatility
2-component GARCH dynamics



This figure plots the paths of the long-run and short-run components of the diffusive volatility from our maintained model estimation.

Table IV
Parameter Estimates: Alternative Specifications of Dynamics

Parameter	Pricing Models			
	GIM-1	GIM-2	constant λ	no $AR(2)$
μ	0.028 (4.146)	0.019 (2.785)	0.010 (0.731)	-0.051 (-2.255)
ψ_v	0.001 (0.111)	0.026 (2.711)	0.019 (1.678)	0.041 (3.380)
ψ_s			-0.014 (-1.029)	-0.112 (-2.923)
ψ_k			0.000 (0.000)	0.000 (0.000)
ρ_1	0.168 (22.515)	0.168 (22.216)	0.159 (22.180)	
ρ_2	-0.029 (-3.889)	-0.032 (-4.350)	-0.048 (-6.864)	
ω	0.013 (12.524)	0.001 (6.512)	0.000 (1.614)	0.000 (-0.825)
α_1	-3.366 (-33.124)	-4.526 (-38.321)	-4.393 (-32.291)	-4.862 (-23.191)
β_1	0.895 (181.624)	0.984 (595.498)	0.974 (276.056)	0.978 (308.368)
α_2		-8.210 (-1.622)	-8.287 (-2.956)	-8.198 (-3.727)
β_2		0.798 (74.803)	0.793 (62.991)	0.776 (41.041)
γ_0			0.117 (5.858)	0.003 (2.195)
γ_1				0.974 (105.193)
γ_2				0.121 (4.348)
θ			-0.380 (-7.593)	-0.435 (-8.655)
δ			0.919 (14.265)	1.031 (15.330)
$\alpha_{a,j,1}$			-6.439 (-1.568)	-2.601 (-5.959)
$\alpha_{a,1}$	1.432 (14.832)	-1.654 (-3.755)	0.691 (0.897)	1.395 (6.909)
$\alpha_{a,j,2}$			-0.588 (-6.310)	-0.194 (-1.745)
$\alpha_{a,2}$		6.604 (1.305)	6.834 (2.438)	6.302 (2.864)
lgl	-25512.470	-25305.875	-24627.911	-24825.400

t-stats are in parenthesis. lgl is the loglikelihood. Model specifications are summarized following Table I.

Table V
Likelihood Ratio Tests
Panel A: Equity Premium Specification

H_1	Test statistic (p-value)
$\psi_v = \psi_s = \psi_k = 0$	12.62 (0.0056)
$\psi_s = \psi_k = 0$	8.720 (0.0128)
Panel B: Specification of Dynamics	
$\alpha_{a,j,1} = \alpha_{a,j,2} = 0$	37.006 (9.21e-09)
$\gamma_1 = \gamma_2 = 0$	103.28 (3.74e-23)
$g_2(\Lambda, \Phi_{t-1}) = \beta_2 = 0$	246.28 (4.13e-52)
$\lambda_t = 0$	1459.2 (1.2e-308)
$\lambda_t = g_2(\Lambda, \Phi_{t-1}) = \beta_2 = 0$	1872.4 (1.6e-394)
$\rho_1 = \rho_2 = 0$	498.26 (6.4e-109)

Null specification: Table II Column 3

Alternative special cases:

$\psi_v = \psi_s = \psi_k = 0$: no pricing of higher-order moments;

$\psi_s = \psi_k = 0$: no pricing of variance, skewness or kurtosis;

$\alpha_{a,j,1} = \alpha_{a,j,2} = 0$: no asymmetry associated with jump innovations;

$\gamma_1 = \gamma_2 = 0$: constant jump arrival;

$g_2(\Lambda, \Phi_{t-1}) = \beta_2 = 0$: no 2nd GARCH component;

$\lambda_t = 0$: no jumps;

$\lambda_t = g_2(\Lambda, \Phi_{t-1}) = \beta_2 = 0$: no jumps and no 2nd GARCH component;

$\rho_1 = \rho_2 = 0$ no AR(2) structure for return innovations (stale pricing).

p-values (computed in Mathematica at high levels of precision) are reported in brackets.

Table VI
Risk-Return Tradeoff with Higher-order Moments
At Low Level of σ

σ	s	$\psi_v \times v$	$\psi_s \times s$	m
0.07	-1.1	2.1%	7.5%	9.6%
0.10	-0.5	3.5%	3.6%	7.1%

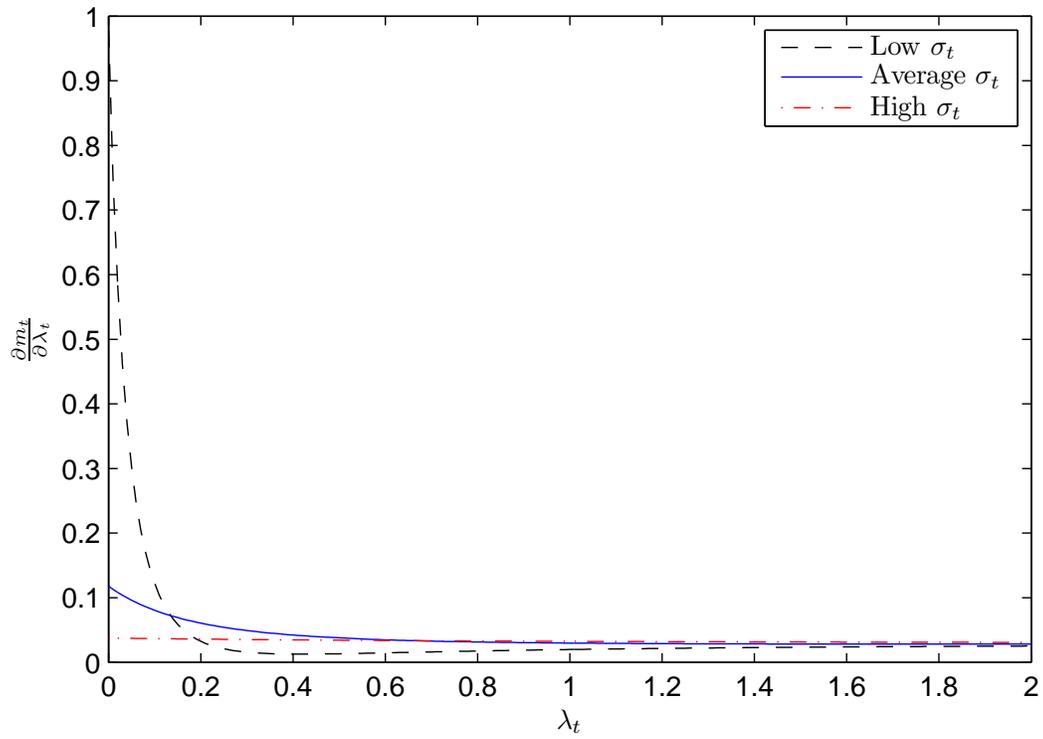
This example presents the equity premium at two different daily diffusive volatility levels. In this example, the parameter estimates are all taken from the ‘prudence’ model. We fix the jump intensity at the sample average to facilitate comparison. All the numbers are annualized.

Table VII
Summary Statistics: Daily Equity Premium m_t

	Median	Mean	StDev
m_{prudence}	0.031	0.039	0.036
$m_{\text{unrestricted}}$	0.028	0.038	0.057

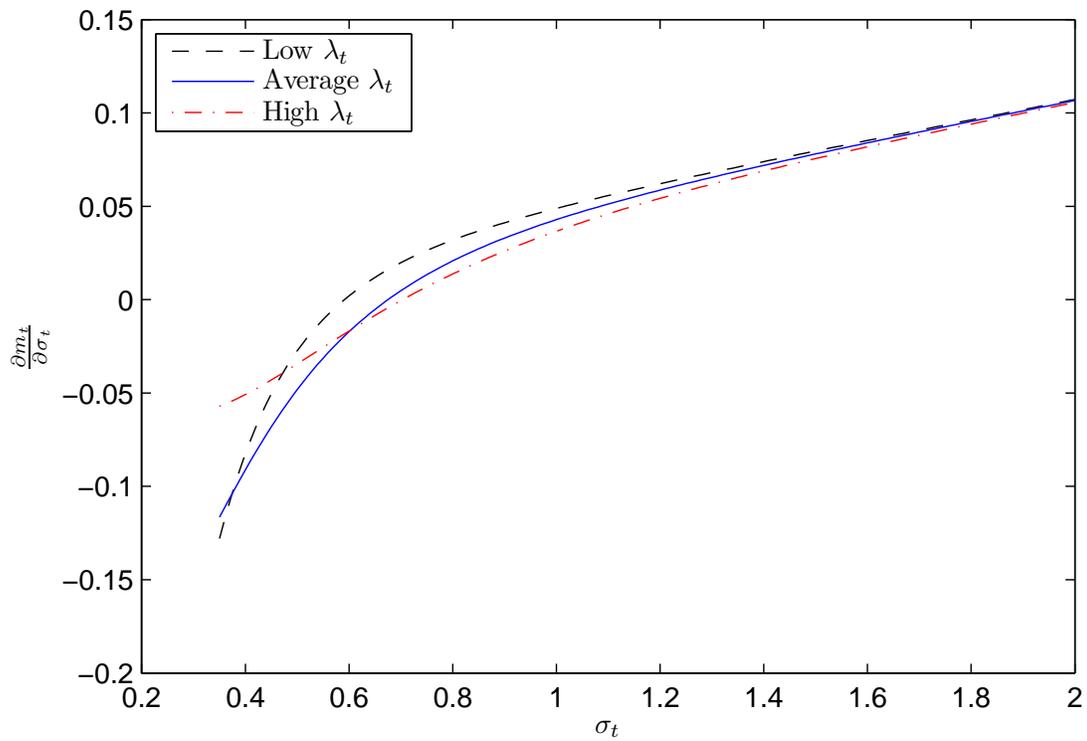
‘Prudence’ and ‘unrestricted’ models in column 2 and 4 of Table II.

Figure 2
Sensitivity of m_t to λ_t at different levels of σ_t



Low $\sigma_t = 0.34$ is the 5th percentile of the estimated σ_t , corresponding to an annualized 5.40%. Average $\sigma_t = 0.75$ is the mean of the estimated σ_t , corresponding to an annualized 11.91%. High $\sigma_t = 1.68$ is the 95th percentile of the estimated σ_t , corresponding to an annualized 26.67%.

Figure 3
Sensitivity of m_t to σ_t at different levels of λ_t



Low $\lambda_t = 0.05$ is the 5th percentile of the estimated λ_t , corresponding to an annualized 12 jumps per year. Average $\lambda_t = 0.149$ is the mean of the estimated λ_t , corresponding to an annualized 37.5 jumps per year. High $\lambda_t = 0.34$ is the 95th percentile of the estimated λ_t , corresponding to an annualized 85 jumps per year.

Table VIII
Summary Statistics: Higher-Order Moments

	Median	Mean	StDev
‘Prudence model’			
Variance	0.518	0.947	1.465
Skewness	-0.451	-0.496	0.320
Kurtosis	4.669	5.238	2.117
‘Unrestricted model’			
Variance	0.512	0.954	1.512
Skewness	-0.422	-0.466	0.311
Kurtosis	4.825	5.447	2.324

‘Prudence’ and ‘unrestricted’ models in column 2 and 4 of Table II.

Table IX
Out-of-Sample Portfolio Performance

Model	Realized Utility		CEQ	Portfolio Weight	
	Mean	Std		Mean	Std
Prudence	4.724E-03	4.46E-02	5.707	0.236	0.040
Variance	4.505E-03	5.29E-02	5.441	0.294	0.070
GIM-1	4.450E-03	5.33E-02	5.374	0.303	0.094
Constant	4.381E-03	5.44E-02	5.291	0.307	0.094

Figure 4
Prudence model: Dynamics of the conditional equity premium m_t

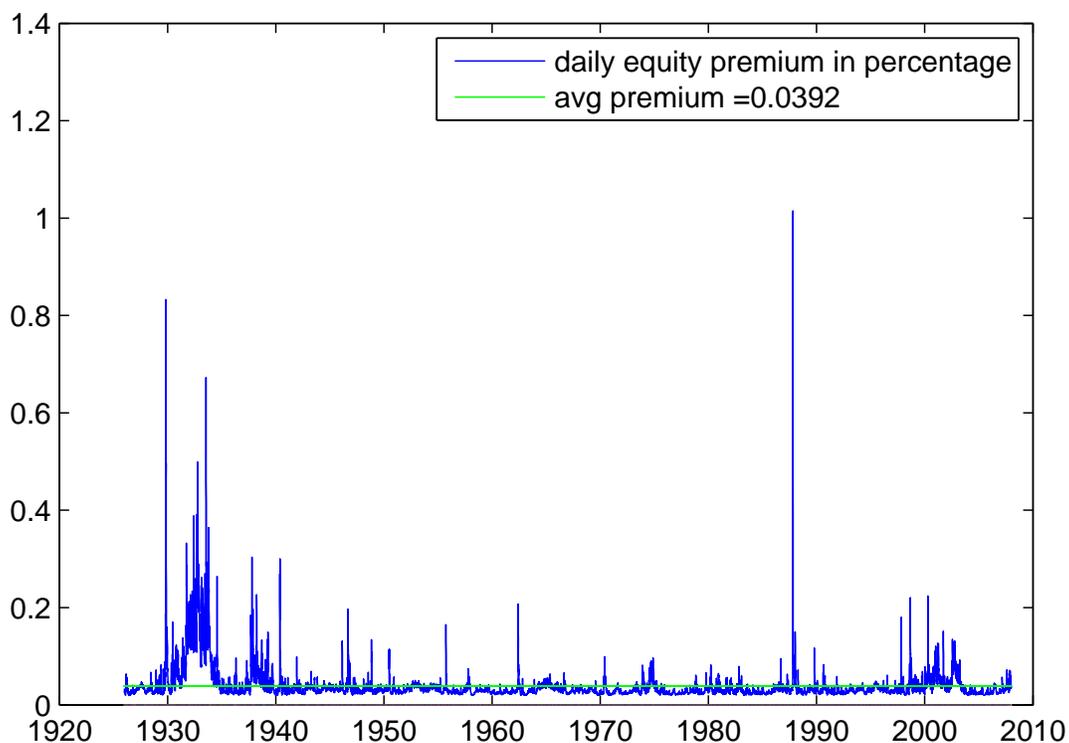


Figure 5
Unrestricted model: Dynamics of the conditional equity premium m_t

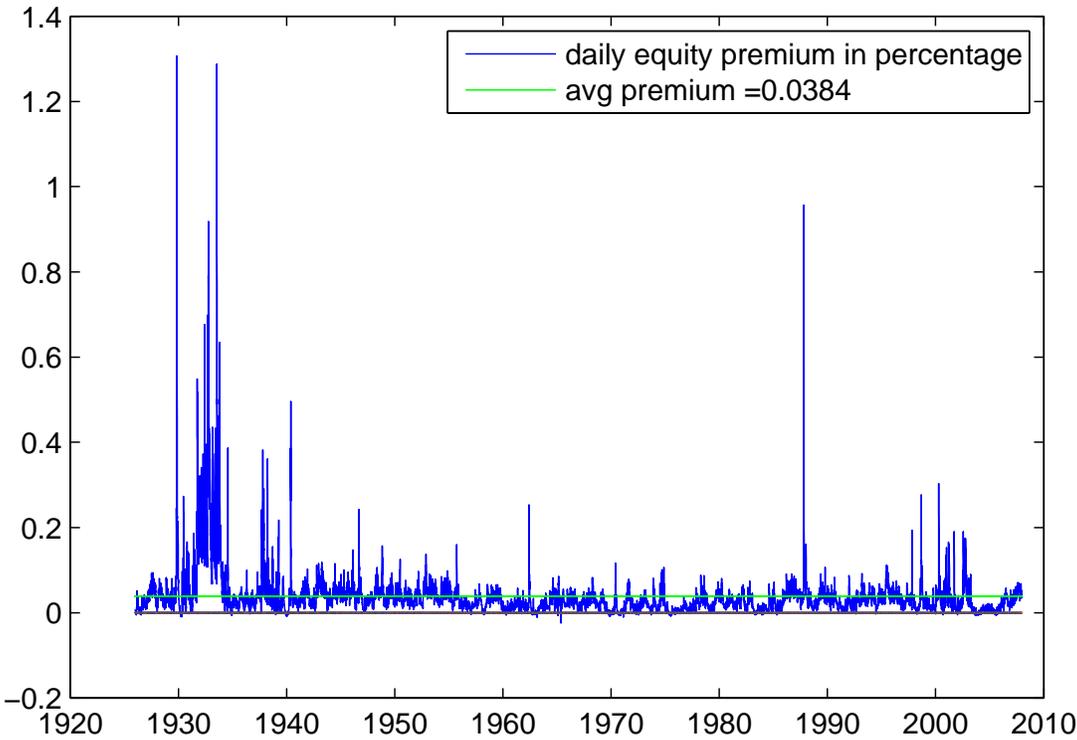


Figure 6
Prudence model: Dynamics of the conditional moments

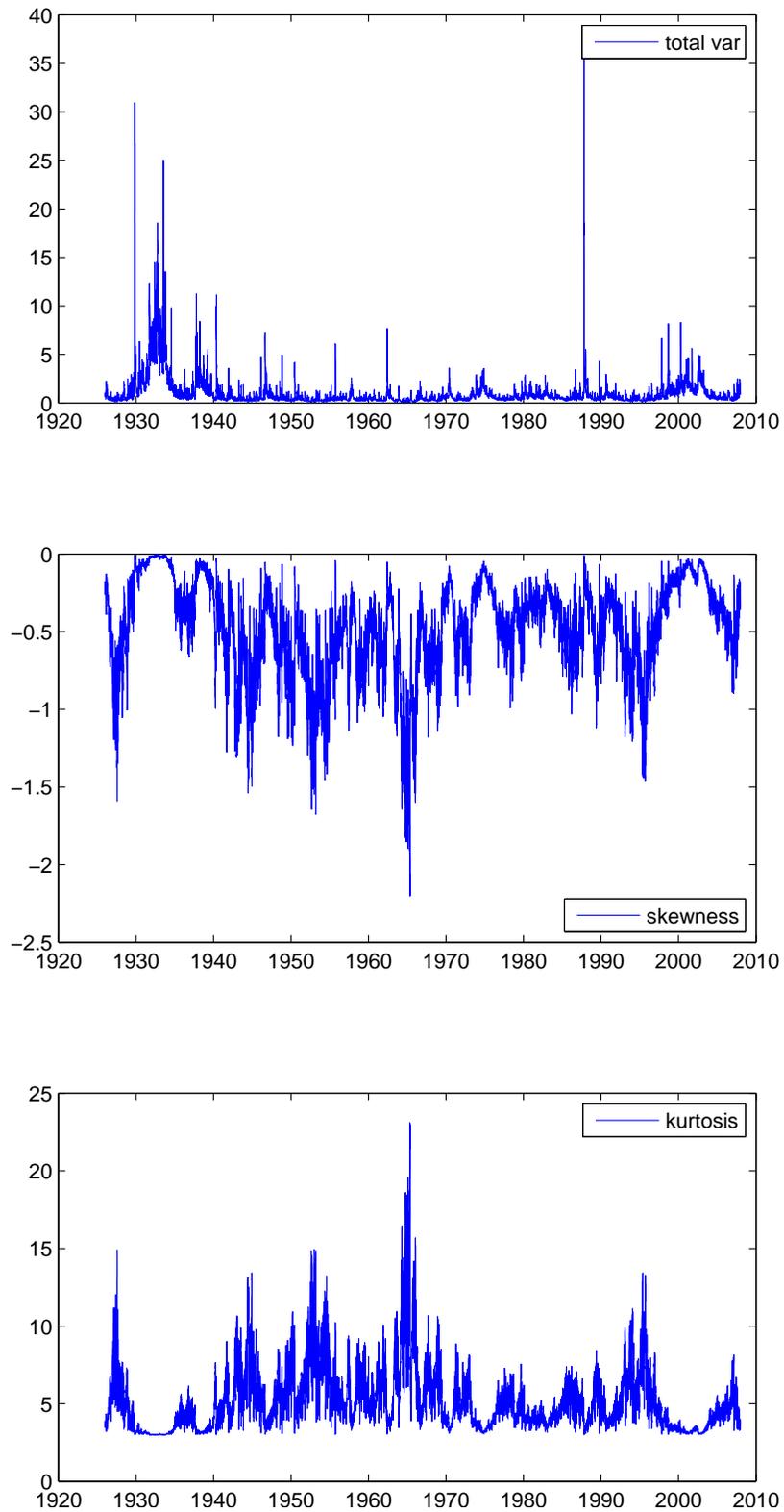


Figure 7
Unrestricted model: Dynamics of the conditional moments

