



Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Do jumps contribute to the dynamics of the equity premium? ☆

John M. Maheu^{a,b}, Thomas H. McCurdy^{c,d,*}, Xiaofei Zhao^e^a DeGroote School of Business, McMaster University, Canada^b RCEA, Rimini, Italy^c Rotman School of Management, University of Toronto, Canada^d CIRANO, Montreal, Canada^e Jindal School of Management, University of Texas at Dallas, United States

ARTICLE INFO

Article history:

Received 27 February 2012

Received in revised form

26 March 2013

Accepted 21 April 2013

Available online 2 August 2013

JEL classification:

G11

G12

C22

Keywords:

Jumps

Higher-order moments

Skewness

Kurtosis

Equity premium

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.

© 2013 Elsevier B.V. All rights reserved.

☆ The authors are very grateful to the referee, David Chapman, for many suggestions that improved the paper. We also thank Peter Christoffersen, George Constantinides, Robert Dittmar, Redouane Elkamhi, Massimo Guidolin, Ranjini Jha, Neil Pearson, Lukasz Pomorski, Jay Shanken, and Liyan Yang, as well as participants at the 22nd Annual Conference on Financial Economics and Accounting (CFEA) at Indiana University, the 2011 Workshop on Time Series and Financial Econometrics at Bocconi University, the 2011 Desautels-HEC-Rotman Winter Finance Workshop, the 2012 Annual SoFiE Conference (Oxford), the 2012 Canadian Econometrics Study Group (Queen's University), the 2012 Southwestern and Midwest Finance Association Annual Meetings, the 2012 NFA Annual Meetings, the 2013 AFA Annual Meetings, and workshops at Brock University, Guelph University and Laval University for valuable comments. The Social Sciences and Humanities Research Council of Canada provided financial support.

* Corresponding author at: University of Toronto, Rotman School of Management, 105 St. George Street, Toronto, Ontario, Canada M5S 3E6. Tel.: +1 416 978 3425; fax: +1 416 971 3048.

E-mail address: tmccurdy@rotman.utoronto.ca (T.H. McCurdy).

1. 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 a generalized utility specification (Kimball, 1990) and nonlinear pricing kernel (Harvey and Siddique, 2000; Dittmar, 2002, Chabi-Yo, Ghysels, and Renault, 2007; Guidolin and Timmermann, 2008), we test whether risks due to dynamics of the conditional variance, skewness, and kurtosis are priced in aggregate stock returns. Our focus is the effect of jumps on the dynamics of the conditional moments and consequently, if priced, on the dynamics of expected excess returns (the equity premium) associated with the market portfolio. We derive a mapping between our estimated prices of risk and the generalized preferences to evaluate

the relative utility of alternative risk premium models in an out-of-sample portfolio performance application.

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.

We model innovations of the return process using a Generalized Autoregressive Conditional Heteroskedastic (GARCH)-jump mixture model. The jump component of the innovation follows a compound Poisson–Normal distribution with an autoregressive jump intensity and a normal jump size distribution. The diffusive component of the innovation is directed by an asymmetric two-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 second 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 the (standardized) higher moments of returns are associated with time-varying expected returns. Are the risks associated with the arrival of jumps, and their effect on the higher moments of returns, priced in the mean?

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 generalized preferences to derive our equity premium specification, prices of risk are not restricted by a single parameter of relative risk aversion and 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. The pricing of the conditional variance is robust across our proposed time-varying jump model specifications. When we restrict the model to have no jumps and only include one GARCH component, the variance dynamics are not significantly priced.

By fixing the GARCH component of volatility, we are able to analyze the marginal effect of jumps on the equity premium. We show that the latter 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, 0.1062% additional expected return for taking on the extra jump risk. This implies that the equity premium associated with jumps is about 3.61% per annum on average. All higher-order moments can be affected by jumps to returns. According to our parameter estimates, on average, jumps contribute 1.06% to the equity premium through the variance dynamics and also add 2.55% 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. The equity premium associated with skewness is about 3.4% 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. When we impose the preference restriction of a nonnegative price associated with risk due to dynamics of kurtosis, our findings show that this price is close to zero; although, conditional kurtosis is significantly priced with a positive sign when the skewness factor is not included. At least at the market level, any contribution of kurtosis to the equity premium has already been largely captured by dynamics of the conditional skewness.

Our results offer an explanation for the conflicting results in the literature on market risk and market expected return. We find a significantly positive equity premium but the positive relationship between conditional *variance* and return only occurs when the GARCH variance component is at or above average levels. An increase in GARCH variance increases both the conditional variance and the conditional skewness (s_t has a smaller negative value), leading to offsetting effects on the equity premium. During calm times (low level of the GARCH variance component), the skewness effect dominates. 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.

Solving for the functional relationship between the parameters of our assumed general utility function and the prices of risk associated with the asset pricing model, we are able to calibrate the implied utility parameters to the empirical estimates for our equity premium specification. We then evaluate the out-of-sample realized utility and certainty-equivalent returns associated with a simple portfolio allocation application. Compared to several special case benchmarks, including one that does not include jumps, our maintained *prudence* model generates higher realized utility and certainty-equivalent returns.

Finally, we check the robustness of our results by extending the model to include a variance risk term as defined by [Chabi-Yo \(2012\)](#). The variance risk is not significantly priced in our maintained model which includes the premium of conditional skewness. When we exclude the contribution of conditional skewness to equity premium, the price of variance risk becomes significant.

This is similar to the cross-sectional result in [Boguth and Kuehn \(2013\)](#) who use a recursive utility derivation to evaluate the premium associated with variance risk of the consumption growth rate. We appeal to the assumptions in [Campbell \(1993\)](#) and derive a time-series application for the market portfolio, and discuss our results in the context of a recursive utility setting.

Our paper differs from the existing literature on pricing jump risk in several important dimensions. Firstly, our parametric model differs from those which employ non-parametric methods to estimate realized jumps from high-frequency data, for example, [Bollerslev and Todorov \(2011\)](#), 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. 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).

Secondly, studies using both underlying asset returns and options are generally based on no-arbitrage (risk-neutral) pricing (for example, [Pan, 2002](#); [Conrad, Dittmar, and Ghysels, 2013](#); [Christoffersen, Jacobs, and Ornthanalai, 2012](#)); 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 will require a long span of calendar data such as used in our paper.

Thirdly, our equity premium specification allows jumps to be priced linearly through the conditional variance and nonlinearly through the higher-order (standardized) 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 misspecification that makes many papers fail to find significant pricing of both jump risk and diffusive risk, especially when only equity data are available.

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.

2. Derivation of our model

The focus of our paper is on the potential effects of large changes in return (jumps) on the equity premium. We build on [Harvey and Siddique \(2000\)](#), [Dittmar \(2002\)](#), and [Guidolin and Timmermann \(2008\)](#) who provide an asset pricing derivation for an empirical specification that tests whether or not higher-order moments of returns are priced as risk factors.¹ In our case, the dynamics of the conditional 3rd and 4th central moments are driven by an autoregressive jump-arrival process; and the conditional variance dynamics are driven by both jumps and a two-component GARCH process. Further, standardizing the conditional 3rd and 4th central moments by functions of the conditional variance (that is, pricing conditional skewness and conditional kurtosis) allows a separation of variance effects from asymmetry effects and allows jump risk to be priced both linearly and nonlinearly. Using results from equilibrium pricing theory, prices of risk are associated with a more general utility function rather than being restricted by the single parameter of relative risk aversion.

2.1. Parameterization of the equity premium

We begin by taking a fourth-order Taylor-series expansion of a general utility function $U(W_{t+1})$ in which W_{t+1} is aggregate wealth at time $t+1$. Defining R_{t+1}^W as the simple net return on wealth and using the equality $W_{t+1} = W_t(1 + R_{t+1}^W)$, we expand $U(W_{t+1})$ around $W_t(1 + C_t)$, where C_t is an arbitrary return:

$$\begin{aligned}
 U(W_{t+1}) &\approx \sum_{n=0}^4 \frac{U^{(n)}(W_t(1 + C_t))}{n!} (W_{t+1} - W_t(1 + C_t))^n \\
 &= \sum_{n=0}^4 \frac{U^{(n)}(W_t(1 + C_t))}{n!} (W_t(R_{t+1}^W - C_t))^n.
 \end{aligned}
 \tag{1}$$

Without loss of generality, assuming the known initial wealth $W_t=1$, the Taylor-series expansion of marginal utility is

$$U'(W_{t+1}) \approx \sum_{n=0}^3 \frac{U^{(n+1)}(1 + C_t)}{n!} (R_{t+1}^W - C_t)^n.
 \tag{2}$$

This implies that the pricing kernel, $M_{t+1} \equiv U'(W_{t+1})/U'(W_t)$, can be approximated by

$$\begin{aligned}
 M_{t+1} &\approx \sum_{n=0}^3 \frac{U^{(n+1)}(1 + C_t)}{U'(1)n!} (R_{t+1}^W - C_t)^n \\
 &= g_{0t} + g_{1t}(R_{t+1}^W - C_t) + g_{2t}(R_{t+1}^W - C_t)^2 + g_{3t}(R_{t+1}^W - C_t)^3,
 \end{aligned}
 \tag{3}$$

¹ Early examples of three-moment Capital Asset Pricing Model (CAPM) applications include [Kraus and Litzenberger \(1976\)](#), [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 \(2013\)](#) use the Intertemporal Capital Asset Pricing Model (ICAPM) to motivate their evaluation of whether market skewness is priced in the cross-section of stock returns. Nonparametric asset pricing models, for example, [Bansal and Viswanathan \(1993\)](#), [Chapman \(1997\)](#), and [Rossi and Timmermann \(2010\)](#) approximate the pricing kernel using a flexible functional form.

in which $g_{nt} = (U^{(n+1)}(1 + C_t)/U'(1))1/n! = (U^{(n+1)}(1 + C_t)/U'(1 + C_t)n!)U'(1 + C_t)/U'(1)$ for $n=0, 1, 2, 3$. The corresponding four-moment asset pricing model becomes

$$E_t[R_{t+1}^W] - R_t^f = \kappa_{1t} \text{Cov}_t(R_{t+1}^W, R_{t+1}^W - C_t) + \kappa_{2t} \text{Cov}_t(R_{t+1}^W, (R_{t+1}^W - C_t)^2) + \kappa_{3t} \text{Cov}_t(R_{t+1}^W, (R_{t+1}^W - C_t)^3), \tag{4}$$

$\kappa_{nt} = -g_{nt}(1 + R_t^f)$ and R_t^f is the net return on the riskfree asset.

As in [Dittmar \(2002\)](#), $U' > 0, U^{(2)} < 0, U^{(3)} > 0$, and $U^{(4)} < 0$, that is, positive marginal utility, risk aversion, decreasing absolute risk aversion, and decreasing absolute prudence,² respectively, imply that $g_{1t} < 0, g_{2t} > 0, g_{3t} < 0$ while $g_{0t} = 1$ for $C_t=0$. Since the utility approximation is truncated at $N=4$, $g_{nt} = 0$ for $n > 3$. With these preference restrictions, $\kappa_{1t} > 0, \kappa_{2t} < 0$, and $\kappa_{3t} > 0$ in Eq. (4).

One frequently used expansion point is $C_t=0$. This has been used by [Harvey and Siddique \(2000\)](#), [Dittmar \(2002\)](#), [Guidolin and Timmermann \(2008\)](#), and others. In what follows, we will discuss an alternative expansion point, $C_t = E_t[R_{t+1}^W]$, also used by [Chabi-Yo, Ghysels, and Renault \(2007\)](#) and [Chabi-Yo \(2012\)](#). As noted in the latter, this is equivalent to Samuelson's small noise expansion. To see this connection, write R_{t+1}^W as

$$R_{t+1}^W - E_t[R_{t+1}^W] = \epsilon Y_{t+1}, \tag{5}$$

then we can see that as ϵ approaches zero, R_{t+1}^W will approach $E_t[R_{t+1}^W]$.

Simplifying notation to

$$R_{t+1} \equiv R_{t+1}^W - R_t^f,$$

$$\epsilon_{t+1} \equiv R_{t+1}^W - E_t[R_{t+1}^W] = (R_{t+1}^W - R_t^f) - (E_t[R_{t+1}^W] - R_t^f) = R_{t+1} - E_t[R_{t+1}],$$

and setting $C_t = E_t[R_{t+1}^W]$, the expected excess return will be a function of the centralized moments, that is,

$$E_t[R_{t+1}] = a_t \text{Var}_t(R_{t+1}) + b_t \text{Cov}_t(R_{t+1}, \epsilon_{t+1}^2) + c_t \text{Cov}_t(R_{t+1}, \epsilon_{t+1}^3) = a_t \text{Var}_t(\epsilon_{t+1}) + b_t E_t[\epsilon_{t+1}^3] + c_t E_t[\epsilon_{t+1}^4], \tag{6}$$

where

$$a_t \equiv -\frac{U^{(2)}(1 + C_t)U'(1 + C_t)(1 + R_t^f)}{U'(1 + C_t)U'(1)}, \tag{7}$$

$$b_t \equiv -\frac{U^{(3)}(1 + C_t)U'(1 + C_t)(1 + R_t^f)}{2U'(1 + C_t)U'(1)}, \tag{8}$$

$$c_t \equiv -\frac{U^{(4)}(1 + C_t)U'(1 + C_t)(1 + R_t^f)}{6U'(1 + C_t)U'(1)}. \tag{9}$$

Expanding around $C_t=0$ and abstracting from changes in R_t^f , these coefficients are constant; but, in that case, the test Eq. (6) will not be a function of the central moments.

For our empirical model, we label the continuously compounded market equity premium expected for period $t+1$, given information at time t , as m_t . Defining $v_t \equiv \text{Var}_t(\epsilon_{t+1})$, $s_t \equiv E_t[\epsilon_{t+1}^3]/v_t^{3/2}$, $k_t \equiv E_t[\epsilon_{t+1}^4]/v_t^2$, that is,

the conditional variance, conditional skewness, and conditional kurtosis, Eq. (6) becomes

$$m_t = \psi_{v,t}v_t + \psi_{s,t}s_t + \psi_{k,t}k_t, \quad \psi_{v,t} \geq 0, \psi_{s,t} \leq 0, \psi_{k,t} \geq 0, \tag{10}$$

and

$$\begin{aligned} \psi_{v,t} &= a_t, \\ \psi_{s,t} &= b_tv_t^{3/2}, \\ \psi_{k,t} &= c_tv_t^2. \end{aligned} \tag{11}$$

Following [Guidolin and Timmermann \(2008\)](#) and [Chabi-Yo \(2012\)](#), we make an additional assumption to estimate an empirical version of the model. In our case,

$$m_t = \psi_v v_t + \psi_s s_t + \psi_k k_t, \quad \psi_v \geq 0, \psi_s \leq 0, \psi_k \geq 0. \tag{12}$$

The sample coefficient estimates in Eq. (12) can be viewed as the unconditional mean of the corresponding time-varying coefficients on the right-hand side (RHS) of Eqs. (11).

We estimate this conditional equity premium with and without an intercept and the restriction on ψ_v . The restrictions on the parameters in Eq. (12), notably that $\psi_s \leq 0$ and $\psi_k \geq 0$, follow from the preference specification and the associated approximation to the pricing kernel described above. [Appendix A.3](#) 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

$$m_t = \psi_v v_t + \psi_s v_t^{3/2} s_t + \psi_k v_t^2 k_t \quad \text{where} \tag{13}$$

$$\psi_v = (\gamma - 0.5), \quad \psi_s = -(3\gamma^2 - 3\gamma + 1)/6, \quad \psi_k = (4\gamma^3 - 6\gamma^2 + 4\gamma - 1)/24. \tag{14}$$

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 Eq. (12). However, our focus is to evaluate whether jumps contribute to the dynamics of the premium as parameterized in Eq. (12) 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 standardized conditional moments s_t and k_t .

2.2. Dynamics of continuously compounded returns

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

$$r_t \equiv r_{m,t} - r_{f,t}, \tag{15}$$

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 $\phi_t = \{r_1, \dots, r_t\}$.

Assume that the dynamics of realized log returns are driven by

$$r_{t+1} = m_t + u_{t+1}, \tag{16}$$

² As derived in [Kimball \(1990, 1993\)](#), decreasing absolute prudence implies that as wealth increases, the precautionary savings motive declines.

where

$$u_{t+1} = \rho_1 u_t + \rho_2 u_{t-1} + \epsilon_{t+1}. \quad (17)$$

That is, u_{t+1} has a predictable autoregressive component due to stale prices,³ or missing pricing factors, and a mean-zero return innovation ϵ_{t+1} .

Combining Eqs. (16) and (17), and decomposing the total return innovation ϵ_{t+1} into two components, we can rewrite realized log returns as

$$r_{t+1} = m_t + \rho_1(r_t - m_{t-1}) + \rho_2(r_{t-1} - m_{t-2}) + \epsilon_{1,t+1} + \epsilon_{2,t+1}, \quad (18)$$

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

Note that the conditional mean of the log return process is

$$E[r_{t+1}|\Phi_t] = m_t + \rho_1(r_t - m_{t-1}) + \rho_2(r_{t-1} - m_{t-2}), \quad (19)$$

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.

2.3. Parameterization of the jump component

The mean-zero (compensated) innovation to returns from jumps is labeled $\epsilon_{2,t+1}$. 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.

2.3.1. Time-varying arrival of jumps

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

$$P(n_{t+1} = j|\Phi_t) = \frac{\exp(-\lambda_t)\lambda_t^j}{j!}, \quad j = 0, 1, 2, \dots \quad (20)$$

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

$$E[n_{t+1}|\Phi_t] = \lambda_t. \quad (21)$$

In other words, the number of jumps in period $t+1$, n_{t+1} , is directed by a conditional Poisson process with a time-varying jump-arrival rate λ_t .⁴

³ As described in Section 3.1 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 5 we show that our results with respect to the equity premium are robust to whether or not we include this autoregressive structure.

⁴ Early examples of time-varying jump arrival include Johannes, Kumar, and Polson (1999), Bates (2000), Andersen, Benzoni, and Lund (2002), Pan (2002), and Eraker (2004).

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

$$\lambda_t = \gamma_0 + \gamma_1 \lambda_{t-1} + \gamma_2 \zeta_t, \quad (22)$$

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

$$\zeta_t = E[n_t|\Phi_t] - \lambda_{t-1} \quad (23)$$

$$\zeta_t = E[n_t|\Phi_t] - E[n_t|\Phi_{t-1}]. \quad (24)$$

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 period t , when information is updated from period $t-1$ to period t . The time-series parameterization for the expected number of jumps, λ_t , given by Eq. (22), 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.

2.3.2. The jump innovation to returns

The compensated jump innovation to returns, $\epsilon_{2,t+1}$, is given by

$$\epsilon_{2,t+1} = J_{t+1} - \theta \lambda_t, \quad (25)$$

where the total size of jumps in period $t+1$, J_{t+1} , is

$$J_{t+1} = \sum_{k=1}^{n_{t+1}} Y_{t+1,k}, \quad (26)$$

in which $Y_{t+1,k}$ is the size of jump k in period $t+1$ which is drawn from a normal distribution with mean θ and variance δ^2 as in

$$Y_{t+1,k} \sim N(\theta, \delta^2). \quad (27)$$

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

$$E[J_{t+1}|\Phi_t] = \theta \lambda_t, \quad (28)$$

the compensated jump innovation is mean zero, that is,

$$E[\epsilon_{2,t+1}|\Phi_t] = E[J_{t+1}|\Phi_t] - \theta \lambda_t = 0. \quad (29)$$

2.4. Parameterization of the normal innovation component

The normal innovation to returns, $\epsilon_{1,t+1}$, is assumed to be directed by a two-component GARCH specification⁵ with feedback from jumps. This specification, which

⁵ Other component GARCH-type models include Engle and Lee (1999), Maheu and McCurdy (2007), and Chan and Feng (2012). Chernov, Gallant, Ghysels, and Tauchen (2003) suggest that either a

generalizes that in Maheu and McCurdy (2004), is parameterized as follows:

$$\epsilon_{1,t+1} = \sigma_t z_{t+1}, \quad z_{t+1} \sim NID(0, 1), \quad \epsilon_{1,t+1} | \Phi_t \sim N(0, \sigma_t^2), \quad (30)$$

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

$$\sigma_t^2 = \sigma_{1,t}^2 + \sigma_{2,t}^2, \quad (31)$$

$$\sigma_{1,t}^2 = \omega + g_1(A_1, \Phi_t) \epsilon_t^2 + \beta_1 \sigma_{1,t-1}^2, \quad (32)$$

$$\sigma_{2,t}^2 = g_2(A_2, \Phi_t) \epsilon_t^2 + \beta_2 \sigma_{2,t-1}^2, \quad (33)$$

$$\epsilon_t = \epsilon_{1,t} + \epsilon_{2,t}. \quad (34)$$

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 Fig. 1. Having the second component helps capture the diffusive volatility better. More importantly, for our purpose, the second component helps control the noisy or transitory part of the diffusive volatility, without which the noise could potentially be sorted as jumps, making jumps less precisely estimated.

The generalized news-impact coefficient $g_i(A_i, \Phi_t)$ 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,

$$g_i(A_i, \Phi_t) = \exp(\alpha_i + I(\epsilon_t)(\alpha_{a,j,i} E[n_t | \Phi_t] + \alpha_{a,i})), \quad i = 1, 2, \\ I(\epsilon_t) = 1 \quad \text{if } \epsilon_t < 0, \quad \text{otherwise } 0. \quad (35)$$

Recall that $E[n_t | \Phi_t]$ is the expected number of jumps at time t given Φ_t , 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. 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.

2.5. Dynamics of higher-order conditional moments

Extending Das and Sundaram (1997) and Maheu and McCurdy (2004) for the ARJI-GARCH specification, the conditional variance (v_t), conditional skewness (s_t), and conditional kurtosis (k_t) are

$$v_t = \sigma_{1,t}^2 + \sigma_{2,t}^2 + \lambda_t(\theta^2 + \delta^2), \quad (36)$$

(footnote continued)

two-component parameterization of stochastic volatility (SV) or an SV-jump (SV-J) diffusion can capture the volatility dynamics.

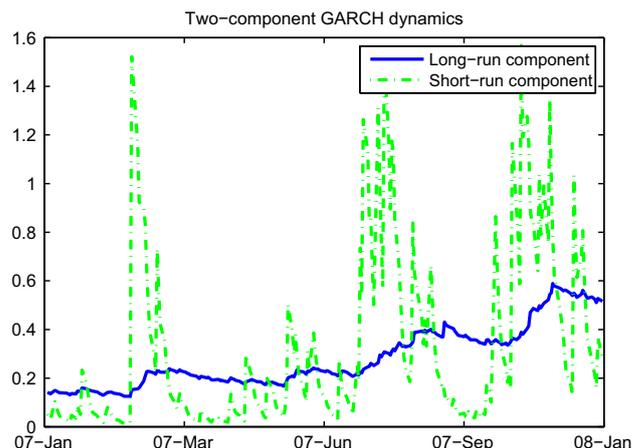


Fig. 1. Long-run and short-run components of the diffusive volatility. This figure plots the paths of the long-run and short-run components of the diffusive volatility from our maintained model estimation.

$$s_t = \frac{\lambda_t(\theta^3 + 3\theta\delta^2)}{(\sigma_{1,t}^2 + \sigma_{2,t}^2 + \lambda_t\theta^2 + \lambda_t\delta^2)^{3/2}}, \quad (37)$$

$$k_t = 3 + \frac{\lambda_t(\theta^4 + 6\theta^2\delta^2 + 3\delta^4)}{(\sigma_{1,t}^2 + \sigma_{2,t}^2 + \lambda_t\theta^2 + \lambda_t\delta^2)^2}. \quad (38)$$

Clearly, jumps affect all of the conditional moments. As indicated in Eq. (36), 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. Clustering of jump arrival, as parameterized in Eq. (22), contributes to volatility clustering. As Eq. (35) shows, jumps also contribute to volatility clustering through the news-impact coefficient in the GARCH specification.

Time-varying jump arrival will be the source of time-variation in the conditional 3rd and 4th moments (numerators of Eqs. (37) and (38), 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 nonzero, $\lambda_t \equiv \lambda > 0$, 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.

3. Data and estimation

3.1. 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, 2011. These returns are converted to continuously compounded daily returns. For the riskfree rates, we convert 30-day Treasury bill returns (t30ret from mcti) to continuously compounded monthly

Table 1

Summary statistics for daily excess returns r_t . The daily excess returns r_t are scaled by 100.

Obs	Mean	StDev	Skewness	Kurtosis	Min	Max
22,785	0.021	1.071	-0.436	20.447	-18.823	14.412

returns and divide by 22 to approximate daily riskfree continuously compounded rates. These are subtracted from the continuously compounded daily index returns resulting in our full sample of 22,785 daily excess log returns. Descriptive statistics are reported in Table 1.

3.2. 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

$$P(n_{t+1} = j | \Phi_{t+1}, \theta) = \frac{f(r_{t+1} | n_{t+1} = j, \Phi_t, \theta) P(n_{t+1} = j | \Phi_t, \theta)}{f(r_{t+1} | \Phi_t, \theta)},$$

$$j = 0, 1, 2, \dots \quad (39)$$

This filter provides an ex post distribution for the number of jumps, n_{t+1} . 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, $P(n_{t+1} \geq 1 | \Phi_{t+1}) = 1 - P(n_{t+1} = 0 | \Phi_{t+1})$, 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 $f(r_{t+1} | \Phi_t, \theta, n_{t+1} = j)$ is

$$\frac{1}{\sqrt{2\pi(\sigma_t^2 + j\delta^2)}} \times \exp\left(-\frac{1}{2} \frac{(r_{t+1} - m_t - \rho_1(r_t - m_{t-1}) - \rho_2(r_{t-1} - m_{t-2}) - (j - \lambda_t)\theta)^2}{\sigma_t^2 + j\delta^2}\right), \quad (40)$$

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

$$f(r_{t+1} | \Phi_t, \theta) = \sum_{j=0}^{\infty} f(r_{t+1} | \Phi_t, \theta, n_{t+1} = j) P(n_{t+1} = j | \Phi_t, \theta), \quad (41)$$

where the second term in the summation is the probability density function (p.d.f.) of the time-varying Poisson distribution in (20). Finally, the full sample loglikelihood is $l = \sum_{t=1}^T \log f(r_{t+1} | \Phi_t, \theta)$ 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 $n_t \geq 10$.

Table 2

Model summary.

$r_{t+1} = m_t + \rho_1(r_t - m_{t-1}) + \rho_2(r_{t-1} - m_{t-2}) + \epsilon_{1,t+1} + \epsilon_{2,t+1}$,
 'Prudence' model: $m_t \equiv \psi_v v_t + \psi_s s_t + \psi_k k_t$, $\psi_s \leq 0, \psi_k \geq 0$,
 'Intercept' model: $m_t \equiv \mu + \psi_v v_t + \psi_s s_t + \psi_k k_t$, $\psi_s \leq 0, \psi_k \geq 0$,
 'Unrestricted' model: $m_t \equiv \mu + \psi_v v_t + \psi_s s_t + \psi_k k_t$,

$$\begin{aligned} \epsilon_{1,t+1} &= \sigma_t z_{t+1}, \quad z_{t+1} \sim NID(0, 1), \\ \sigma_t^2 &= \sigma_{1,t}^2 + \sigma_{2,t}^2, \\ \sigma_{1,t}^2 &= \omega + g_1(A_1, \Phi_t) \epsilon_t^2 + \beta_1 \sigma_{1,t-1}^2, \\ \sigma_{2,t}^2 &= g_2(A_2, \Phi_t) \epsilon_t^2 + \beta_2 \sigma_{2,t-1}^2, \\ \epsilon_t &= \epsilon_{1,t} + \epsilon_{2,t}, \\ g_i(A_i, \Phi_t) &= \exp(\alpha_i + I(\epsilon_t)(\alpha_{a,j,i} E[n_t | \Phi_t] + \alpha_{a,i})) \quad \text{for } i=1,2;, \\ I(\epsilon_t) &= 1 \quad \text{if } \epsilon_t < 0, \quad \text{otherwise } 0, \\ \epsilon_{2,t+1} &= \sum_{k=1}^{n_{t+1}} Y_{t+1,k} - \theta \lambda_t, \quad Y_{t+1,k} \sim N(\theta, \delta^2), \\ \lambda_t &= E[n_{t+1} | \Phi_t], \quad \lambda_t = \gamma_0 + \gamma_1 \lambda_{t-1} + \gamma_2 \epsilon_t, \\ v_t &= \sigma_t^2 + \lambda_t(\theta^2 + \delta^2), \\ s_t &= \frac{\lambda_t(\theta^3 + 3\theta\delta^2)}{(\sigma_t^2 + \lambda_t\theta^2 + \lambda_t\delta^2)^{3/2}}, \\ k_t &= 3 + \frac{\lambda_t(\theta^4 + 6\theta^2\delta^2 + 3\delta^4)}{(\sigma_t^2 + \lambda_t\delta^2 + \lambda_t\theta^2)^2}. \end{aligned}$$

4. Results

4.1. Parameter estimates

Table 3 provides parameter estimates for the full model with alternative specifications for the conditional market equity premium, m_t . Column 2, labeled 'prudence', reports estimates for our maintained parameterization of m_t given in Eq. (12). Column 3, labeled 'intercept', provides estimates for the same model including an intercept, μ , in the conditional mean. Column 4, labeled 'unrestricted', removes the sign restrictions on skewness and kurtosis implied by decreasing absolute prudence. Column 5, labeled 'variance', just prices the conditional variance, v_t ; column 6, labeled 'skewness', prices both the conditional variance and conditional skewness; and column 7, labeled '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.

In order to focus on the alternative specifications for the equity premium, all of the models reported in Table 3 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 3 reveal significant risk pricing associated with the market equity premium. In our maintained specification (column 2), the coefficient on the conditional variance is 0.026 (ψ_v) with a significant t -stat of 3.4. The skewness coefficient ψ_s is -0.027 with a significant t -stat of -3.4, implying a significant negative price associated with skewness. 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. Exposure to dynamics of kurtosis, however, does not seem to be

Table 3

Parameter estimates: alternative specifications of the equity premium.

Parameters refer to the model summarized in Table 2. Restrictions are: $\mu = 0, \psi_s \leq 0, \psi_k \geq 0$, for the 'prudence' model; $\psi_s \leq 0, \psi_k \geq 0$, for the 'intercept' model; $\psi_s = 0, \psi_k = 0$, for the 'variance' model; $\psi_k = 0$, for the 'skewness' model; and $\psi_s = 0$, for the 'kurtosis' model. *t*-Stats are in parentheses. *lgl* is the loglikelihood.

Parameter	Pricing models					
	Prudence	Intercept	Unrestricted	Variance	Skewness	Kurtosis
μ		-0.041 (-2.247)	-0.021 (-1.186)	0.020 (3.337)	-0.041 (-2.248)	-0.022 (-1.194)
ψ_v	0.026 (3.396)	0.045 (4.051)	0.053 (5.190)	0.022 (2.216)	0.045 (4.052)	0.034 (3.151)
ψ_s	-0.027 (-3.423)	-0.069 (-3.014)	-0.270 (-4.064)		-0.069 (-3.015)	
ψ_k	0.000 (0.000)	0.000 (0.000)	-0.022 (-3.228)			0.005 (2.295)
ρ_1	0.148 (21.080)	0.147 (21.012)	0.154 (21.382)	0.149 (21.240)	0.147 (21.012)	0.149 (21.225)
ρ_2	-0.046 (-6.828)	-0.047 (-6.867)	-0.051 (-7.325)	-0.045 (-6.615)	-0.047 (-6.867)	-0.046 (-6.713)
ω	0.000 (0.154)	0.000 (0.402)	0.000 (0.547)	0.000 (0.480)	0.000 (0.403)	0.000 (0.278)
α_1	-4.672 (-29.386)	-4.752 (-28.003)	-4.781 (-28.978)	-4.657 (-29.089)	-4.752 (-28.013)	-4.687 (-29.195)
β_1	0.973 (351.850)	0.976 (338.501)	0.977 (395.158)	0.972 (350.061)	0.976 (338.244)	0.974 (341.294)
α_2	-8.195 (-3.700)	-8.140 (-3.980)	-8.110 (-3.722)	-8.249 (-3.457)	-8.145 (-3.803)	-8.197 (-3.950)
β_2	0.781 (46.827)	0.788 (53.014)	0.795 (60.768)	0.775 (41.523)	0.788 (53.027)	0.781 (46.595)
$\alpha_{a,j,1}$	-2.523 (-4.826)	-2.935 (-4.490)	-3.752 (-4.231)	-2.188 (-5.450)	-2.935 (-4.502)	-2.490 (-4.746)
$\alpha_{a,1}$	1.376 (6.338)	1.250 (5.744)	1.216 (6.211)	1.483 (7.475)	1.250 (5.747)	1.379 (6.280)
$\alpha_{a,j,2}$	-0.122 (-1.123)	-0.217 (-2.128)	-0.306 (-3.452)	-0.044 (-0.422)	-0.217 (-2.129)	-0.124 (-1.120)
$\alpha_{a,2}$	6.293 (2.841)	6.355 (3.106)	6.408 (2.939)	6.234 (2.609)	6.360 (2.969)	6.298 (3.031)
γ_0	0.005 (2.789)	0.003 (2.324)	0.002 (3.025)	0.007 (3.395)	0.003 (2.324)	0.006 (2.750)
γ_1	0.959 (88.165)	0.972 (100.115)	0.979 (163.297)	0.948 (82.171)	0.972 (100.120)	0.959 (85.437)
γ_2	0.132 (3.524)	0.100 (3.612)	0.067 (4.178)	0.169 (3.848)	0.100 (3.615)	0.135 (3.499)
θ	-0.482 (-9.101)	-0.471 (-8.519)	-0.455 (-8.077)	-0.508 (-9.845)	-0.471 (-8.520)	-0.491 (-9.281)
δ	0.977 (15.927)	1.026 (15.281)	1.047 (16.112)	0.948 (16.870)	1.026 (15.283)	0.982 (15.637)
<i>lgl</i>	-26,386.791	-26,384.271	-26,378.504	-26,390.700	-26,384.270	-26,387.465

priced in the presence of the skewness premium. For the 'intercept' case (column 3), the intercept μ , which is used to capture missing pricing factors, has a *t*-stat -2.2. Returns are very volatile for several months following September 2008. For this subperiod, the diffusive variance captures a relatively larger proportion of the variability. This significantly higher variance than average implies that a negative intercept is required to fit the average premium for the entire sample. If one were to end the sample prior

to 2008, the intercept is insignificantly different from zero supporting our maintained specification given by Eq. (12) with results reported in column 2.⁶

⁶ Note that our results are not driven by the potential structural break at the end of the 1930s. Estimation using subsamples 1940 to 2011 and 1963 to 2011 give similar results.

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 0.97); while the second GARCH component has a β_2 of about 0.78 indicating lower persistence. As reported in the third row of Panel B in Table 6, a one-component GARCH specification is rejected (p -value $4.13e-52$) in favor of the maintained two-component GARCH parameterization.

The next set of parameters (fourth panel Table 3) capture volatility asymmetry with respect to the sign of the return innovation and the inferred number of jumps as parameterized in Eq. (35). These asymmetries enter both the short-run and long-run components 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 two-component GARCH specification, there remains strong evidence of jump dynamics in daily returns. The arrival of jumps is autocorrelated with a γ_1 of 0.96 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 Eqs. (36)–(38). The likelihood ratio (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 0.13 which implies about 33 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.

4.1.1. Alternative equity premium specifications

We explore additional pricing specifications in columns 5–7 of Table 3. 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 LR test presented in Table 6 Panel A rejects $\psi_s = \psi_k = 0$.⁷

It is interesting that if we estimate the power utility special case of our maintained model, that is, Eqs. (13) and (14), 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 3 since ψ_s and ψ_k associated with Eqs. (13) and (14) are estimated to be close to zero and statistically insignificant in our sample. Note that this

special case implies a coefficient of relative risk aversion of about 2.7.⁸

Column 4 of Table 3 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 (2013) also find a large negative correlation between risk-neutral skewness and kurtosis, a price of skewness 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.

4.1.2. 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:

$$m_t = \psi_v \sigma_t^2 + \psi_j \lambda_t. \quad (42)$$

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 4 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 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

⁷ The test statistic is 12.858; the p -value of 0.0016 corresponds to two degrees of freedom (two restrictions). However, as shown in Table 3, 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 even lower.

⁸ As derived in our Appendix, the coefficient on the variance for the risk premium, associated with simple as opposed to continuously compounded returns, is $100 \times \psi_v + 0.5$.

Table 4

Importance of pricing jumps nonlinearly.

Model specifications are summarized in Table 2. The ‘prudence’ model estimates are from the second column of Table 3. The ‘linear’ model assumes that the equity premium is $\psi_v\sigma_t^2 + \psi_j\lambda_t$, with other dynamics the same as in the ‘prudence’ model. *t*-Stats are in parentheses. lgl is the loglikelihood.

Parameter	Pricing models	
	prudence	linear
ψ_v	0.026 (3.40)	0.017 (1.44)
ψ_s	-0.027 (-3.42)	
ψ_k	0.000 (0.000)	
ψ_j		0.188 (2.47)
lgl	-26,386.79	-26,392.30

comparison. In the ‘linear’ model, we have $\hat{\psi}_j = 0.188$ (*t*-stat=2.47) and $\hat{\psi}_v = 0.017$ (*t*-stat=1.44), i.e., jump risk is significantly priced whereas diffusive risk is not. Furthermore, the likelihood (-26,392.3) of the ‘linear’ model is smaller than that (-26,386.8) 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).

This comparison highlights the importance of having a nonlinear pricing structure for jumps in the equity premium. Having jumps priced through higher-order moments not only works well empirically but, theoretically, is also consistent with more general preferences as derived in Appendix A.3.

4.1.3. Alternative specifications of dynamics

Table 5 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 (GIM) specification (columns 2 and 3, respectively). As evident from the resulting likelihoods in Table 5, compared with our maintained model in columns 2 or 3 of Table 3, including jumps results in a much superior fit. As reported in Panel B of Table 6, the LR test statistics of 1,879.9 and 1,465.1 reject the no-jump one-component and two-component GIM special cases with extremely small *p*-values.

Nevertheless, even with our most general specification of jumps, the second GARCH component is still very important; the LR test reported in Panel B of Table 6 comparing column 3 of Table 3 with a special case without the second 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 6.

Column 4 of Table 5 shows the results of assuming constant jump arrival. The LR test reported in Panel B of Table 6 decisively rejects constant jump arrival in favor of our autoregressive parameterization of the jump-arrival process (*p*-value of 3.74e-23).

Finally, column 5 of Table 5 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 3: both ψ_v and ψ_s are significantly priced and ψ_k is close to zero.

4.2. 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 3, 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. Fig. 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.

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 0.1062, suggesting that one more jump in a year increases the equity premium by 0.1062%. The average jump-arrival rate ($\bar{\lambda} = 0.1361$) implies about 34 jumps a year. Therefore,

Table 5

Parameter estimates: alternative specifications of dynamics.

Parameters refer to the model summarized in Table 2. The model label 'GIM-1' refers to a GARCH-in-Mean model with one variance component and no jumps; 'GIM-2' refers to a GARCH-in-Mean specification with two variance components and no jumps; 'constant λ ' refers to the 'intercept' model from Table 3 but with a constant jump arrival rate λ ; and the 'no AR(2)' specification is the 'intercept' model with no autoregressive terms in the mean. *t*-Stats are in parenthesis. *lgl* is the loglikelihood.

Parameter	Pricing models			
	GIM-1	GIM-2	constant λ	no AR(2)
μ	0.028 (4.241)	0.019 (2.938)	0.010 (0.806)	-0.053 (-2.952)
ψ_v	0.001 (0.123)	0.023 (2.690)	0.017 (1.652)	0.043 (4.213)
ψ_s			-0.013 (-0.981)	-0.107 (-3.660)
ψ_k			0.000 (0.000)	0.000 (0.000)
ρ_1	0.158 (21.613)	0.158 (21.355)	0.150 (21.301)	
ρ_2	-0.026 (-3.661)	-0.030 (-4.207)	-0.045 (-6.640)	
ω	0.013 (12.937)	0.001 (7.163)	0.000 (1.713)	0.000 (-0.379)
α_1	-3.393 (-33.523)	-4.467 (-39.724)	-4.375 (-34.660)	-4.789 (-26.000)
β_1	0.896 (193.941)	0.982 (583.814)	0.973 (296.804)	0.978 (335.339)
α_2		-8.302 (-2.352)	-8.303 (-3.329)	-8.165 (-4.070)
β_2		0.805 (79.699)	0.799 (68.915)	0.798 (56.954)
$\alpha_{aj,1}$			-5.729 (-1.539)	-3.031 (-5.307)
$\alpha_{a,1}$	1.459 (15.119)	-1.643 (-3.545)	0.621 (0.839)	1.199 (5.588)
$\alpha_{aj,2}$			-0.578 (-6.580)	-0.241 (-2.472)
$\alpha_{a,2}$		6.686 (1.894)	6.857 (2.748)	6.340 (3.160)
γ_0			0.126 (6.317)	0.003 (2.171)
γ_1				0.980 (127.773)
γ_2				0.092 (4.198)
θ			-0.382 (-8.035)	-0.454 (-8.789)
δ			0.895 (15.433)	1.048 (16.913)
<i>lgl</i>	-27,324.196	-27,116.827	-26,426.173	-26,615.734

based on our maintained model reported in column 2 of Table 3, the contribution of jumps to the market equity premium is about 3.61% per annum. Note that Pan (2002) and Elkamhi and Ornthanalai (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.

The solid line in Fig. 3 reports the level of the equity premium for different levels of the diffusive (GARCH) volatility σ_t , holding the jump-arrival rate at its average level $\bar{\lambda}$. The dotted line shows the equity premium due to the total variance, that is, $\psi_v \times v_t$. Finally, the dashed line shows the equity premium component due to skewness, that is, $\psi_s \times s_t$. Interestingly, in contrast to the effect of an increase in λ_t on m_t , an increase in GARCH volatility has a negative effect on the premium for small values of σ_t and

Table 6

Likelihood ratio tests.

Null specification: Table 3 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 skewness or kurtosis;

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

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

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

$\lambda_t = 0$: no jumps;

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

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

*** Indicates that the alternative hypothesis is rejected at the 1% level.

Panel A: Equity premium specification

H_1	Test statistic
$\psi_v = \psi_s = \psi_k = 0$	17.746***
$\psi_s = \psi_k = 0$	12.858***

Panel B: Specification of dynamics

$\alpha_{aj,1} = \alpha_{aj,2} = 0$	27.804***
$\gamma_1 = \gamma_2 = 0$	83.804***
$g_2(\lambda, \Phi_{t-1}) = \beta_2 = 0$	288.97***
$\lambda_t = 0$	1465.1***
$\lambda_t = g_2(\lambda, \Phi_{t-1}) = \beta_2 = 0$	1879.9***
$\rho_1 = \rho_2 = 0$	462.93***

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).⁹

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 skewness measure than from the variance. Intuitively, investors value the increase in skewness (s_t has a smaller negative value) and the potential increase in upside more, 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 .

We also present a numerical example in Table 7 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 2.3% (annualized) and the equity premium from the conditional skewness is 6.5% (annualized), resulting in a total equity premium of 8.8% (annualized) for this particular day. It is clear that the

⁹ This can be verified from Eqs. (36) to (38) and the fact that the estimated jump-size mean θ is negative. That is

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

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

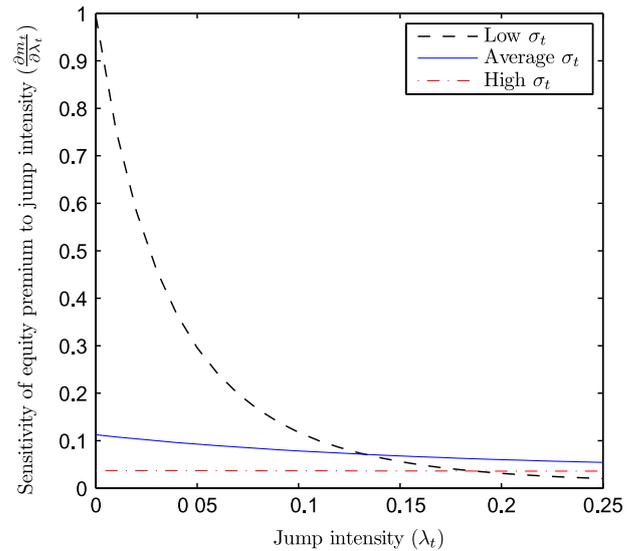


Fig. 2. Sensitivity of m_t to λ_t at different levels of σ_t . Low $\sigma_t = 0.35$ is the 5th percentile of the estimated σ_t , corresponding to an annualized 5.47%. Average $\sigma_t = 0.79$ is the mean of the estimated σ_t , corresponding to an annualized 12.44%. High $\sigma_t = 1.83$ is the 95th percentile of the estimated σ_t , corresponding to an annualized 28.89%.

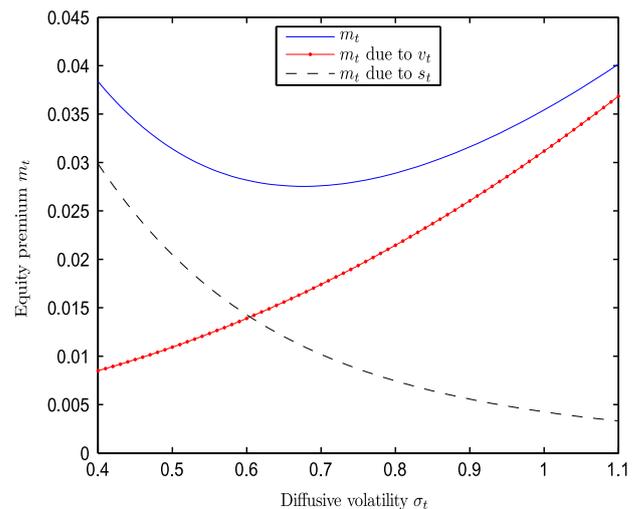


Fig. 3. m_t as a function of σ_t at average λ_t . This figure plots the level of the equity premium (the solid line) for different levels of the diffusive (GARCH) volatility σ_t , holding the jump-arrival rate at its average level. The dotted line shows the equity premium due to the total variance. The dashed line shows the equity premium component due to skewness.

skewness component dominates in this case because the magnitude of the skewness is quite high (-1.0). 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 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 7.0%, despite an increase in the equity premium due to the variance (from 2.3% to 3.7%). Therefore, if

Table 7

Risk-return tradeoff with higher-order moments.

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.

σ	s	$\psi_v \times v$	$\psi_s \times s$	m
0.07	-1.0	2.3%	6.5%	8.8%
0.10	-0.5	3.7%	3.3%	7.0%

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 8.8% to 7.0%. 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.

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.

4.3. Equity premium size and dynamics

Table 8 presents the descriptive statistics of the equity premium from our maintained model in column 2 of Table 3, as well as those for the unrestricted model in column 4 of the same table. The median (average) equity premium estimate is about 7.5% (10%) per annum for the maintained prudence model and 6.6% (9.8%) for the unrestricted parameterization.¹⁰ Fig. 4 illustrates the time-series dynamics of the equity premium for our prudence specification. The restriction implied from preference theory for the prudence model ensures that the (expected) equity premium is always positive. Notice that the average premium will be affected by a few large outliers associated particularly with 2008, 1987, and the

Table 8

Summary statistics: daily equity premium m_t .

'Prudence' and 'unrestricted' models in columns 2 and 4 of Table 3.

Equity premium	Median	Mean	StDev
m_{prudence}	0.030	0.040	0.039
$m_{\text{unrestricted}}$	0.026	0.039	0.070

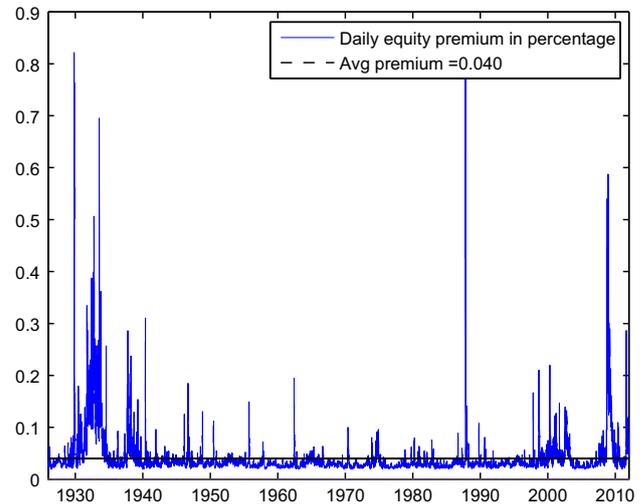


Fig. 4. Prudence model: dynamics of the conditional equity premium m_t . This figure plots the time series of the daily conditional equity premium estimated from the prudence model. The time period is from January 1926 to December 2011.

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 (1940–2011) subsample.

According to our parameter estimates, the average expected number of jumps per year is 34. Combining this with the average impact of a change in jump risk implies that the equity premium associated with jumps is about 3.61% 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 3), using the average of the estimated jump-arrival rate $\bar{\lambda}$, jumps contribute 1.06% to the equity premium through the variance dynamics¹¹ and 2.55% 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, 3.4%, which contributes about 34% of the overall equity premium for our sample.¹² 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

¹⁰ Daily returns were scaled by 100. Therefore, to annualize the median or average daily premiums reported in Table 8, we scale by (252/100).

¹¹ $\psi_v \times (\bar{\lambda} \times (\theta^2 + \delta^2)) \times 2.52 = 0.026 \times (0.136 \times ((-0.482)^2 + 0.977^2)) \times 2.52$.

¹² We have $\psi_s = -0.028$, average $s_t = -0.496$. Therefore, the skewness premium is $(-0.027) \times (-0.496) \times \frac{252}{100} = 0.034$ or 3.4% per annum.

monthly stock returns. In our parameterization of the time-series of daily market excess returns, jumps account for about 76% of that skewness premium.

As noted above, when we impose the preference restriction of a nonnegative price associated with kurtosis, our findings show that this price 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.

Table 9 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 3, respectively). The former are plotted in Fig. 5.

Table 10 reports the average jump-arrival rate $\bar{\lambda}$ and the percentage of total variance due to the jump component versus the diffusive component. As noted in Section 4.1 above, the large increase in return variability in the final quarter of 2008 implies that total variance was much higher. Although the frequency of jumps was also much higher for that subperiod relative to the entire sample, the proportion of the total variance due to the diffusive component increased dramatically to 95% relative to 74% on average for the entire sample.

5. 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 riskfree 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

Table 9
Summary statistics: higher-order moments.
'Prudence' and 'unrestricted' models in columns 2 and 4 of Table 3.

	Median	Mean	StDev
'Prudence model'			
Variance	0.541	1.042	1.658
Skewness	-0.444	-0.496	0.338
Kurtosis	4.670	5.317	2.278
'Unrestricted model'			
Variance	0.535	1.053	1.766
Skewness	-0.408	-0.458	0.324
Kurtosis	4.710	5.385	2.379

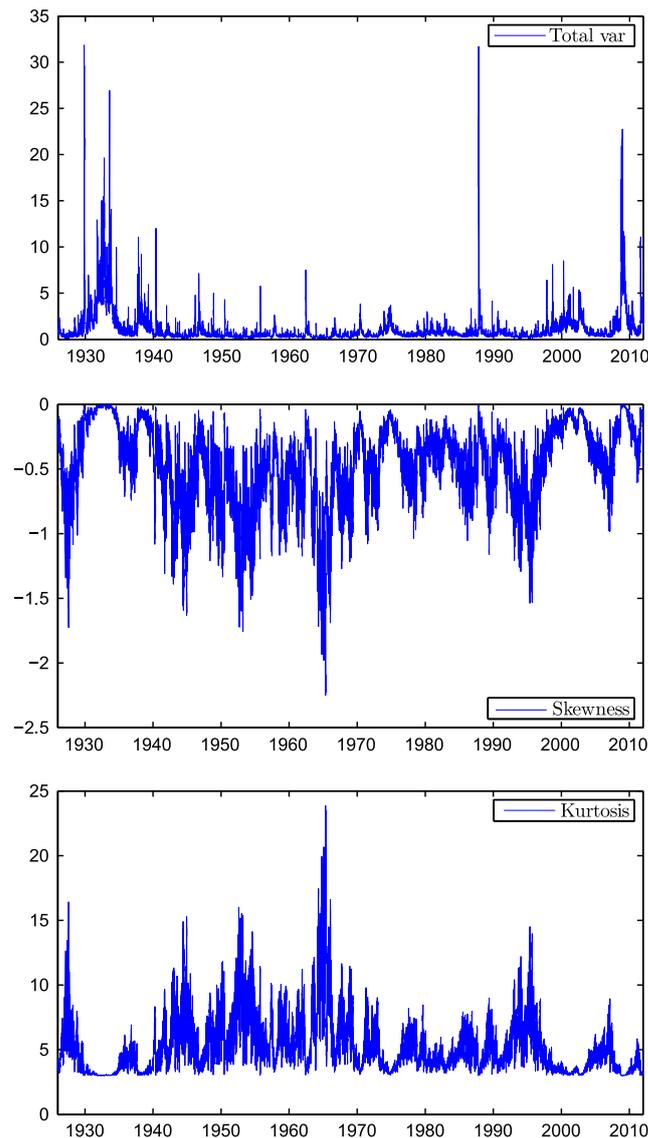


Fig. 5. Prudence model: dynamics of the conditional moments. This figure plots the time series of the daily conditional variance, conditional skewness, and conditional kurtosis estimated from the prudence model. The time period is from January 1926 to December 2011.

Table 10
Summary statistics for jump intensity and total variance.

Sample period	$\bar{\lambda}$	$\bar{\nu}$	ν Due to jump part	ν Due to diffusive part
1926:01–2011:12	0.136	1.042	26%	74%
1926:01–1939:12	0.161	2.192	18%	82%
1940:01–2011:12	0.131	0.784	27%	73%
2008:01–2011:12	0.199	2.796	14%	86%
2008:09–2009:01	0.345	9.563	5%	95%

challenging. However, the equity premium associated with our maintained prudence model, summarized in Section 2.1, provides additional parameters (prices of risk) associated with the higher-order terms in the risk premium specification. As

shown in Appendix A.3, 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 3. The *GIM-1* model, which is the traditional single-component GARCH-in-Mean model, is reported in column 2 of Table 5. 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 2011 to evaluate the out-of-sample performance.

The out-of-sample portfolio performance results are reported in Table 11. 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 11, 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).¹³ There are clear benefits to pricing jumps (*prudence* versus *GIM-1*), and to pricing higher-order moments (*prudence* versus *variance*).

6. Robustness and alternative interpretations

6.1. Variance risk

In our maintained (*prudence*) model, derived and estimated above, we evaluated the potential importance of equity premium components due to the dynamics of conditional variance, skewness, and kurtosis. Following Chabi-Yo (2012) and Chabi-Yo, Ghysels, and Renault (2007), we derive (see our Appendix A.4 below) a two-period intertemporal parametrization of our model in order to add variance risk to our equity premium specification. As shown in Appendix A.5, abstracting from conditional kurtosis which is not significantly priced in our maintained (*prudence*) model, the conditional equity premium in this extended model is

$$m_t = \psi_v v_t + \psi_s s_t + \psi_\zeta \frac{\text{Cov}_t(v_{t+1}, \epsilon_{t+1})}{v_t^{3/2}}, \quad (43)$$

where the first two terms are premium components due to the dynamics of conditional variance and conditional

¹³ As in Guidolin and Timmermann (2008), the certainty equivalent return is calculated based on the average realized utility.

Table 11
Out-of-sample portfolio performance.

Model	Realized utility		CEQ	Portfolio weight	
	Mean	Std		Mean	Std
Prudence	3.32E-03	5.05E-02	4.005	0.220	0.046
Variance	3.09E-03	5.73E-02	3.719	0.266	0.078
GIM-1	2.88E-03	5.51E-02	3.469	0.276	0.101
Constant	2.80E-03	5.70E-02	3.375	0.275	0.096

skewness and the last term captures the premium due to variance risk. Note that due to conditioning on time t information, $\text{Cov}_t(v_{t+1}, \epsilon_{t+1})$ is equivalent to $\text{Cov}_t(\Delta v_{t+1}, \epsilon_{t+1})$. When an investor is facing a multi-horizon portfolio allocation problem, the uncertainty in future variance matters. As shown in our Appendices A.4 and A.5, the variance risk term in Eq. (43) captures the covariation of a future return innovation with those changes in conditional variance.

For our sample, the average variance risk is negative and the estimated price, ψ_ζ , is statistically insignificant, as in Chabi-Yo (2012) who uses generalized method of moments (GMM) estimation in a cross-section of stock returns. Our conclusion is based on a likelihood ratio test which shows the expanded model to be insignificantly different ($\psi_\zeta = 0$) than the *prudence* model.

6.2. Recursive utility

Our analyses have focused on the time-separable utility framework. In this framework, our empirical model could be linked to investors' preferences over higher-order moments and volatility risk. Many asset pricing models have used time non-separable utility, particularly recursive utility as in Epstein and Zin (1989), Bansal and Yaron (2004), and Routledge and Zin (2010). For example, Boguth and Kuehn (2013) examine consumption volatility risk using a cross-section of stock returns and conclude that the pricing of consumption volatility risk reflects investors' preference for early resolution of uncertainty; that is, that the elasticity of intertemporal substitution (EIS) is greater than the inverse of the relative risk aversion parameter ($1/\text{RRA}$). Their average beta associated with variance risk for the median portfolio sorted by this beta is negative (-0.05 in Table V of Boguth and Kuehn, 2013).¹⁴ The price of variance risk is also negative (-0.12) in their Fama-MacBeth cross-sectional regression. For the market, the premium should then be positive.

In Appendix A.6 below, following the parametrization in Boguth and Kuehn (2013), we show that our model can also be linked to a pricing equation derived from a parametric recursive utility framework. Appealing to the

¹⁴ Average beta for the whole sample is not reported but can be inferred from the sample average of their five quintile portfolios. The inferred number would be $(-0.32 - 0.15 - 0.05 + 0.05 + 0.15)/5 = -0.052$. The beta for a value-weighted portfolio should be $(-0.32) \times 12.19\% + (-0.15) \times 18.54\% + (-0.05) \times 23.13\% + 0.05 \times 24.39\% + 0.21 \times 21.74\% = -0.021$.

assumptions in Campbell (1993), we derive a time-series application for the market portfolio and discuss our results in the context of a recursive utility setting. In this case, the equity premium specification is

$$E_t[R_{m,t+1}] - R_f = \psi_v \text{Var}_t(R_{m,t+1}) + \psi_\eta \text{Cov}_t(\sigma_{t+1}, R_{m,t+1}). \quad (44)$$

This is analogous to a special case of our prudence model, that is, the variance model reported in column 5 of Table 3, extended to include variance risk.

If we estimate a model with variance only (no intercept), the loglikelihood is $-26,395.149$, compared to $-26,388.651$ for the model with variance and variance risk. That is, variance risk is significantly priced if no higher-order moments are present (p -value is 0.0003 for the LR test). For this parametrization, the average value of variance risk and the coefficient for variance risk are both negative so, as in Boguth and Kuehn (2013), the premium due to variance risk is positive.

In the previous subsection we discussed a parametrization which extends our more general prudence model to include variance risk. In that case, we found that adding variance risk did not improve the fit of our maintained prudence model for which conditional skewness is significantly priced.¹⁵ Furthermore, the loglikelihood for our prudence model ($-26,386.791$) is better than that ($-26,388.651$) for the case which includes variance risk but excludes higher-order moments.

7. 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. 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, one more jump per year would increase the equity premium by 0.1062%. The average number of jumps per year according to our parameterization and sample is 34 which implies an overall jump risk premium of 3.61% per annum. Jump risk is priced in the market equity premium both through the jump component of variance dynamics and skewness contributing 1.06% and 2.55%, respectively, to those premiums.

¹⁵ The conditional skewness and quantity of variance risk are highly correlated. To the extent that they capture similar variation, our maintained empirical model (prudence model) can also be interpreted in the recursive utility framework as supporting the assumption that EIS is greater than 1/RRA.

The total skewness premium is about 3.4% 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 close to zero in the presence of the skewness factor. We explore different specifications of the equity premium by pricing skewness and kurtosis separately and find them to have the expected signs. At the market level, any contribution of conditional kurtosis to the equity premium seems to be largely captured by dynamics of the conditional skewness.

Finally, we check the robustness of our results by extending the model to include a variance risk term as proposed by Chabi-Yo (2012). We find that the variance risk is not significantly priced in our maintained model which includes the contribution of conditional skewness to the equity premium. On the other hand, including variance risk in a special case of our model which excludes the contribution of conditional skewness and kurtosis does result in a significant price of variance risk, as in the cross-sectional results in Ang, Hodrick, Xing, and Zhang (2006) and Boguth and Kuehn (2013). To explore this further, we derive and estimate a special case of our model in a recursive utility setting applied to our time-series of market returns, and discuss our results in the context of those in the cross-sectional application in Boguth and Kuehn (2013). Nevertheless, our prudence model which prices higher-order moments still fits the data better than this recursive utility parameterization which includes variance risk but excludes higher-order moments.

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.

Appendix A. Technical appendix

A.1. 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 2 that

$$r = m + \epsilon_1 + \epsilon_2, \quad (45)$$

$$m = \mu + \psi_v v + \psi_s s + \psi_k k, \quad (46)$$

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} ,

$$r_{100} = m_{100} + \epsilon_{1,100} + \epsilon_{2,100}, \quad (47)$$

in which

$$m_{100} \equiv 100m = 100(\mu + \psi_v v + \psi_s s + \psi_k k), \quad (48)$$

$$\epsilon_{1,100} \equiv 100\epsilon_1, \tag{49}$$

$$\epsilon_{2,100} \equiv 100\epsilon_2. \tag{50}$$

Estimating Eq. (47) is equivalent to estimating Eq. (45) using maximum likelihood estimation.

One can verify that

$$100\epsilon_1 \sim N(0, 100^2\sigma^2). \tag{51}$$

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

$$100Y_k \sim N(100\theta, 100^2\delta^2). \tag{52}$$

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

$$\sigma_{100} = 100\sigma \tag{53}$$

$$\lambda_{100} = \lambda \tag{54}$$

$$\theta_{100} = 100\theta \tag{55}$$

$$\delta_{100} = 100\delta, \tag{56}$$

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 Eqs. (53)–(56), it is straightforward to see that

$$v_{100} = \sigma_{100}^2 + \lambda_{100}(\theta_{100}^2 + \delta_{100}^2) \tag{57}$$

$$v_{100} = 100^2\sigma^2 + \lambda(100^2\theta^2 + 100^2\delta^2) = 10,000v, \tag{58}$$

$$s_{100} = \frac{\lambda_{100}(\theta_{100}^3 + 3\theta_{100}\delta_{100}^2)}{(\sigma_{100}^2 + \lambda_{100}\theta_{100}^2 + \lambda_{100}\delta_{100}^2)^{3/2}} = s, \tag{59}$$

$$k_{100} = 3 + \frac{\lambda_{100}(\theta_{100}^4 + 6\theta_{100}^2\delta_{100}^2 + 3\delta_{100}^4)}{(\sigma_{100}^2 + \lambda_{100}\theta_{100}^2 + \lambda_{100}\delta_{100}^2)^2} = k. \tag{60}$$

Note that in the estimation, we only use the scaled return r_{100} , therefore, in Eq. (48)

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

we need to replace v with v_{100} using $v_{100} = 10,000v$, replace s with s_{100} using $s_{100} = s$, and replace k with k_{100} using $k_{100} = k$. After this operation we have

$$m_{100} = 100\mu + 0.01\psi_v v_{100} + 100\psi_s s_{100} + 100\psi_k k_{100}. \tag{61}$$

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

$$\mu_{100} = 100\mu, \tag{62}$$

$$\psi_{v,100} = 0.01\psi_v, \tag{63}$$

$$\psi_{s,100} = 100\psi_s, \tag{64}$$

$$\psi_{k,100} = 100\psi_k, \tag{65}$$

and obtain

$$m_{100} = \mu_{100} + \psi_{v,100}v_{100} + \psi_{s,100}s_{100} + \psi_{k,100}k_{100}. \tag{66}$$

Combining all of the rescaled terms:

$$r_{100} = \mu_{100} + \psi_{v,100}v_{100} + \psi_{s,100}s_{100} + \psi_{k,100}k_{100} + \epsilon_{1,100} + \epsilon_{2,100}. \tag{67}$$

A.2. Simple risk premium

Eq. (12) 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, labeled as $m_{s,t}$, is

$$m_{s,t} = E[\exp(r_{t+1} - \rho_1(r_t - m_{t-1}) - \rho_2(r_{t-1} - m_{t-2})) | \Phi_t] - 1 \tag{68}$$

$$m_{s,t} = E[\exp(m_t + \epsilon_{1,t+1} + \epsilon_{2,t+1}) | \Phi_t] - 1. \tag{69}$$

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

$$E[\exp(\epsilon_{1,t+1}) | \Phi_t] = \exp(0.5\sigma_t^2). \tag{70}$$

Recall from Section 2.3.1 that the compensated jump innovation is

$$\epsilon_{2,t+1} = J_{t+1} - \theta\lambda_t = \sum_{k=1}^{n_{t+1}} Y_{t+1,k} - \theta\lambda_t,$$

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

$$E[\exp(\epsilon_{2,t+1}) | \Phi_t] = E[\exp(J_{t+1}) | \Phi_t] \exp(-\theta\lambda_t) = \exp(\lambda_t (E[\exp(Y) | \Phi_t] - 1) - \theta\lambda_t).$$

As noted in Eq. (27), the jump-size distribution is normal, in which case

$$E[\exp(Y) | \Phi_t] = \exp\left(\theta + \frac{1}{2}\delta^2\right),$$

so that

$$E[\exp(\epsilon_{2,t+1}) | \Phi_t] = \exp\left(\lambda_t \left(\exp\left(\theta + \frac{1}{2}\delta^2\right) - 1 - \theta\right)\right) = \exp(\lambda_t \xi). \tag{71}$$

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

$$E[\exp(r_{t+1} - \rho_1(r_t - m_{t-1}) - \rho_2(r_{t-1} - m_{t-2})) | \Phi_t] - 1 \tag{72}$$

$$= (\exp(m_{s,t}) E[\exp(\epsilon_{1,t+1}) | \Phi_t] E[\exp(\epsilon_{2,t+1}) | \Phi_t]) - 1 = (\exp(\mu + \psi_v v_t + \psi_s s_t + \psi_k k_t + 0.5\sigma_t^2 + \lambda_t \xi) - 1) \tag{73}$$

in which

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

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

A.3. Calibrating utility coefficients to model estimates

To conduct our out-of-sample portfolio allocation analysis, we calibrate the utility coefficients to our model estimates. We have the following mapping between the coefficients of the Taylor-series expansion of the general utility and the estimated prices of risk:

$$\frac{U^{(2)}(1 + C_t)}{U'(1)} = -\frac{a_t}{1 + R_t^f} = -\frac{\psi_v}{1 + R_t^f}, \tag{74}$$

$$\frac{U^{(3)}(1 + C_t)}{U'(1)} = -\frac{2b_t}{1 + R_t^f} = -\frac{2\psi_s v_t^{-3/2}}{1 + R_t^f}, \tag{75}$$

$$\frac{U^{(4)}(1 + C_t)}{U'(1)} = -\frac{6c_t}{1 + R_t^f} = -\frac{6\psi_k v_t^{-2}}{1 + R_t^f}. \tag{76}$$

In our out-of-sample portfolio allocation evaluation, we fix v_t and R_t^f on the RHS at their sample averages. Note that the utility coefficients given by the above equations are underdetermined for the coefficients required for the approximation to the level of utility as expressed in the above equations. However, without loss of generality, since initial wealth is fixed and normalized to $W_t=1$, we can set $U'(1)$ to a constant and solve for $U^{(2)}(1 + C_t)$, $U^{(3)}(1 + C_t)$, and $U^{(4)}(1 + C_t)$ accordingly.

A.4. Taylor expansion at arbitrary returns for a two-period model

Following Chabi-Yo (2012), we assume that a representative agent maximizes the expected utility in a two-period model as

$$\max_{\{\omega_t\}} E_t \left(\max_{\{\omega_{t+1}\}} E_{t+1}[U(W_{t+2})] \right), \tag{77}$$

where the investor wealth is

$$W_{t+2} = W_t(1 + R_t^f + \omega_t R_{t+1}^e)(1 + R_{t+1}^f + \omega_{t+1} R_{t+2}^e) \tag{78}$$

The weights in N individual risky assets are represented by the $1 \times N$ vectors ω_t, ω_{t+1} . The R_{t+1}^e, R_{t+2}^e are the $N \times 1$ excess return vectors for individual assets. By construction, we have the market excess return as

$$R_{t+1}^W - R_t^f = \omega_t R_{t+1}^e, \tag{79}$$

$$R_{t+2}^W - R_{t+1}^f = \omega_{t+1} R_{t+2}^e. \tag{80}$$

Without loss of generality, we assume $W_t=1$. Note that the wealth at $t + 1, t + 2$ can also be expressed as

$$W_{t+1} = (1 + R_{t+1}^W) \tag{81}$$

$$W_{t+2} = W_{t+1}(1 + R_{t+2}^W) = (1 + R_{t+1}^W)(1 + R_{t+2}^W). \tag{82}$$

In this two-period intertemporal model, the marginal utility can be approximated using a recursive univariate Taylor-series expansion, or equivalently, the following

bivariate Taylor-series expansion for $f(x, y)$:

$$\begin{aligned} f(x, y) \approx & f(x_0, y_0) + \frac{\partial f}{\partial x} \Big|_{(x_0, y_0)} (x - x_0) + \frac{\partial f}{\partial y} \Big|_{(x_0, y_0)} (y - y_0) \\ & + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \Big|_{(x_0, y_0)} (x - x_0)^2 \\ & + \frac{1}{2} \frac{\partial^2 f}{\partial y^2} \Big|_{(x_0, y_0)} (y - y_0)^2 + \frac{\partial^2 f}{\partial x \partial y} \Big|_{(x_0, y_0)} (x - x_0)(y - y_0). \end{aligned} \tag{83}$$

Applying this bivariate expansion to our case, we expand $U'(W_{t+2})$ around $(1 + R_{t+1}^W, 1 + R_{t+2}^W)$ at arbitrary returns $(1 + C_t, 1 + C_{t+1})$, that is

$$\begin{aligned} U'(W_{t+2}) \approx & U'[(1 + C_t)(1 + C_{t+1})] \\ & + U''[(1 + C_t)(1 + C_{t+1})](1 + C_{t+1})(R_{t+1}^W - C_t) \\ & + U''[(1 + C_t)(1 + C_{t+1})](1 + C_t)(R_{t+2}^W - C_{t+1}) \\ & + \frac{1}{2} U'''[(1 + C_t)(1 + C_{t+1})](1 + C_{t+1})^2 (R_{t+1}^W - C_t)^2 \\ & + \frac{1}{2} U'''[(1 + C_t)(1 + C_{t+1})](1 + C_t)^2 (R_{t+2}^W - C_{t+1})^2 \\ & + (U'''[(1 + C_t)(1 + C_{t+1})](1 + C_t)(1 + C_{t+1}) \\ & + U''[(1 + C_t)(1 + C_{t+1})]) \times (R_{t+1}^W - C_t)(R_{t+2}^W - C_{t+1}). \end{aligned}$$

The first-order condition with respect to ω_t is

$$E_t[U'[W_{t+2}](R_{kt+1} - R_t^f)] = 0, \tag{84}$$

and we can get the expected return expression for asset k as

$$E_t[R_{kt+1} - R_t^f] = -\text{Cov}_t \left(\frac{U'[W_{t+2}]}{E_t[U'[W_{t+2}]}}, R_{kt+1} \right). \tag{85}$$

Substituting the bivariate Taylor-series expansion of $U'[W_{t+2}]$ into the expected return equation. We have

$$\begin{aligned} E_t[R_{kt+1} - R_t^f] \approx & -\frac{U''[(1 + C_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]}(1 + C_{t+1}) \\ & \times \text{Cov}_t \left((R_{t+1}^W - C_t), R_{kt+1} \right) - \frac{U''[(1 + r_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]}(1 + C_t) \\ & \times \text{Cov}_t \left((R_{t+2}^W - C_{t+1}), R_{kt+1} \right) \\ & - \frac{1}{2} \frac{U'''[(1 + C_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]}(1 + C_{t+1})^2 \\ & \text{Cov}_t \left((R_{t+1}^W - C_t)^2, R_{kt+1} \right) - \frac{1}{2} \frac{U'''[(1 + C_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]}(1 + C_t)^2 \\ & \text{Cov}_t \left((R_{t+2}^W - C_{t+1})^2, R_{kt+1} \right) \\ & - \left(\frac{U'''[(1 + C_t)(1 + C_{t+1})](1 + C_t)(1 + C_{t+1}) + U''[(1 + C_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]} \right) \\ & \times \text{Cov}_t \left((R_{t+1}^W - C_t)(R_{t+2}^W - C_{t+1}), R_{kt+1} \right), \end{aligned}$$

where

$$a_t = -\frac{U''[(1 + C_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]}(1 + C_{t+1}) \tag{86}$$

$$a_{1t} = -\frac{U''[(1 + C_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]}(1 + C_t) \tag{87}$$

$$b_t = -\frac{1}{2} \frac{U'''[(1 + C_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]}(1 + C_{t+1})^2 \tag{88}$$

$$b_{1t} = -\frac{1}{2} \frac{U'''[(1 + C_t)(1 + C_{t+1})]}{E_t[U'[W_{t+2}]}(1 + C_t)^2 \tag{89}$$

$$h_t = - \frac{(U'''[(1+C_t)(1+C_{t+1}))(1+C_t)(1+C_{t+1}) + U''[(1+C_t)(1+C_{t+1})]}{E_t[U''[W_{t+2}]]} \quad (90)$$

We assume that the coefficients depend on time t information. Note that this is equivalent to assuming the time $t+1$ variation is canceling out in these coefficients. In some power utility cases, this cancelation can be exact. Another special case is where $C_t = C_{t+1} = 0$, in which case we have constant coefficients:

$$a_t = a_{1t} = - \frac{U''[1]}{U'[1]} \quad (91)$$

$$b_t = b_{1t} = - \frac{1}{2} \frac{U'''[1]}{U''[1]} \quad (92)$$

$$h_t = - \frac{(U'''[1] + U''[1])}{U''[1]} \quad (93)$$

However, the corresponding expected return equation under $C_t = C_{t+1} = 0$ becomes more complicated for a time-series application for which it is useful to use centralized moments. If we have $\text{Cov}_t((R_{t+1}^W)^2, R_{t+1}^W)$ on the RHS of the expected aggregate return equation, we need to know the expected aggregate return first in order to compute $\text{Cov}_t((R_{t+1}^W)^2, R_{t+1}^W)$. Therefore, in our case, we do the Taylor-series expansion around $C_t = E_t[R_{t+1}^W], C_{t+1} = E_{t+1}[R_{t+2}^W]$.

Recall that the expected return equation for asset k is

$$\begin{aligned} E_t[R_{kt+1} - R_t^f] &= a_t \text{Cov}_t((R_{t+1}^W - C_t), R_{kt+1}) \\ &+ a_{1t} \text{Cov}_t((R_{t+2}^W - C_{t+1}), R_{kt+1}) \\ &+ b_t \text{Cov}_t((R_{t+1}^W - C_t)^2, R_{kt+1}) + b_{1t} \text{Cov}_t((R_{t+2}^W - C_{t+1})^2, R_{kt+1}) \\ &+ h_t \text{Cov}_t((R_{t+1}^W - C_t)(R_{t+2}^W - C_{t+1}), R_{kt+1}). \end{aligned} \quad (94)$$

When $C_t = E_t[R_{t+1}^W], C_{t+1} = E_{t+1}[R_{t+2}^W]$, we have

$$\text{Cov}_t((R_{t+2}^W - C_{t+1}), R_{kt+1}) = \text{Cov}_t(E_{t+1}(R_{t+2}^W - C_{t+1}), R_{kt+1}) = 0, \quad (95)$$

$$\begin{aligned} \text{Cov}_t((R_{t+1}^W - C_t)(R_{t+2}^W - C_{t+1}), R_{kt+1}) \\ = \text{Cov}_t(E_{t+1}[(R_{t+1}^W - C_t)(R_{t+2}^W - C_{t+1})], R_{kt+1}) = 0, \end{aligned} \quad (96)$$

and the expected return equation becomes

$$\begin{aligned} E_t[R_{kt+1} - R_t^f] &= a_t \text{Cov}_t((R_{t+1}^W - C_t), R_{kt+1}) \\ &+ b_t \text{Cov}_t((R_{t+1}^W - C_t)^2, R_{kt+1}) \\ &+ b_{1t} \text{Cov}_t((R_{t+2}^W - C_{t+1})^2, R_{kt+1}). \end{aligned} \quad (97)$$

For the aggregate return, we have

$$\begin{aligned} E_t[R_{t+1}^W - R_t^f] &= a_t \text{Var}_t(\epsilon_{t+1}) + b_t E_t[\epsilon_{t+1}^3] \\ &+ b_{1t} \text{Cov}_t(\text{Var}_{t+1}(\epsilon_{t+2}), \epsilon_{t+1}) \end{aligned} \quad (98)$$

$$E_t[R_{t+1}^W - R_t^f] = a_t v_t + b_t v_t^{3/2} s_t + b_{1t} v_t^{3/2} \frac{\text{Cov}_t(v_{t+1}, \epsilon_{t+1})}{v_t^{3/2}} \quad (99)$$

Let $\zeta_t \equiv \text{Cov}_t(v_{t+1}, \epsilon_{t+1})/v_t^{3/2}$ and if we make the following assumptions:

$$\begin{aligned} a_t &= \psi_v, \\ b_t v_t^{3/2} &= \psi_s, \\ b_{1t} v_t^{3/2} &= \psi_\zeta, \end{aligned}$$

then we have

$$m_t = \psi_v v_t + \psi_s s_t + \psi_\zeta \zeta_t. \quad (100)$$

A.5. Implementation of volatility risk

To take into account the variance risk in a two-period model, we need to evaluate

$$\text{Cov}_t(v_{t+1}, \epsilon_{t+1}), \quad (101)$$

where all the expectations are evaluated with respect to information at time t . In our model, the closed-form formula for this quantity is not easy to obtain. We know that

$$v_{t+1} = \sigma_{t+1}^2 + \lambda_{t+1}(\theta^2 + \delta^2) \quad (102)$$

$$\sigma_{t+1}^2 = \omega + [g_1(A_1, \Phi_{t+1}) + g_2(A_2, \Phi_{t+1})] \epsilon_{t+1}^2 + \beta_1 \sigma_{1,t}^2 + \beta_2 \sigma_{2,t}^2 \quad (103)$$

$$g_i(A_i, \Phi_{t+1}) = \exp(\alpha_i + I(\epsilon_{t+1})[\alpha_{a,i} E(n_{t+1} | \Phi_{t+1}) + \alpha_{a,i}]) \quad (104)$$

$$\lambda_{t+1} = \gamma_0 + \gamma_1 \lambda_t + \gamma_2 [E(n_{t+1} | \Phi_{t+1}) - \lambda_t]. \quad (105)$$

Therefore, the volatility risk can be expanded as

$$\begin{aligned} \text{Cov}_t(v_{t+1}, \epsilon_{t+1}) &= \text{Cov}_t([g_1(A_1, \Phi_{t+1}) + g_2(A_2, \Phi_{t+1})] \epsilon_{t+1}^2, \epsilon_{t+1}) \\ &+ \gamma_2(\theta^2 + \delta^2) \text{Cov}_t(E(n_{t+1} | \Phi_{t+1}), \epsilon_{t+1}), \end{aligned} \quad (106)$$

where we have used the fact that $\sigma_{1,t}^2, \sigma_{2,t}^2, \lambda_t$ are known at time t . We need to evaluate the following two terms:

$$\begin{aligned} \text{Cov}_t([g_1(A_1, \Phi_{t+1}) \\ + g_2(A_2, \Phi_{t+1})] \epsilon_{t+1}^2, \epsilon_{t+1}) \quad \text{and} \quad \text{Cov}_t(E(n_{t+1} | \Phi_{t+1}), \epsilon_{t+1}). \end{aligned} \quad (107)$$

We know that

$$E(n_{t+1} | \Phi_{t+1}) = \sum_{j=0}^{\infty} j P(n_{t+1} = j | \Phi_{t+1}) \quad (108)$$

$$P(n_{t+1} = j | \Phi_{t+1}) = \frac{f(r_{t+1} | n_{t+1} = j, \Phi_t) P(n_{t+1} = j | \Phi_t)}{f(r_{t+1} | \Phi_t)} \quad (109)$$

$$f(r_{t+1} | n_{t+1} = j, \Phi_t) = \frac{1}{\sqrt{2\pi(\sigma_t^2 + j\delta^2)}} \exp\left[-\frac{[\epsilon_{t+1} - (j - \lambda_t)\theta]^2}{2(\sigma_t^2 + j\delta^2)}\right] \quad (110)$$

$$f(r_{t+1} | \Phi_t) = \sum_{j=0}^{\infty} f(r_{t+1} | n_{t+1} = j, \Phi_t) P(n_{t+1} = j | \Phi_t), \quad (111)$$

and then

$$\begin{aligned} \text{Cov}_t(E(n_{t+1} | \Phi_{t+1}), \epsilon_{t+1}) &= \sum_{j=0}^{\infty} j \text{Cov}_t(P(n_{t+1} = j | \Phi_{t+1}), \epsilon_{t+1}) \\ &= \sum_{j=0}^{\infty} j P(n_{t+1} = j | \Phi_t) \text{Cov}_t\left(\frac{f(r_{t+1} | n_{t+1} = j, \Phi_t)}{f(r_{t+1} | \Phi_t)}, \epsilon_{t+1}\right). \end{aligned} \quad (112)$$

It appears that

$$\text{Cov}_t\left(\frac{f(r_{t+1} | n_{t+1} = j, \Phi_t)}{f(r_{t+1} | \Phi_t)}, \epsilon_{t+1}\right)$$

would be difficult to evaluate directly as both $f(r_{t+1}|n_{t+1} = j, \Phi_t)$ and $f(r_{t+1}|\Phi_t)$ are nonlinear functions and involve exponential function of ϵ_{t+1} . Taylor-series expansion of

$$\frac{f(r_{t+1}|n_{t+1} = j, \Phi_t)}{f(r_{t+1}|\Phi_t)}$$

in terms of ϵ_{t+1} is also not straightforward to obtain as the denominator $f(r_{t+1}|\Phi_t)$ is the summation of an infinite series. The other key term

$$\text{Cov}_t([g_1(A_1, \Phi_{t+1}) + g_2(A_2, \Phi_{t+1})]\epsilon_{t+1}^2, \epsilon_{t+1})$$

is perhaps more difficult to obtain in closed-form because $g_i(A_i, \Phi_{t+1})$ involves an asymmetric random indicator $I(\epsilon_{t+1})$ and an exponential function of $E(n_{t+1}|\Phi_{t+1})$. We need to rely on simulations to evaluate these terms.

Finally, to incorporate this extra variance risk term in our empirical model, the return dynamics would be

$$r_{t+1} = m_t + \rho_1(r_t - m_{t-1}) + \rho_2(r_{t-1} - m_{t-2}) + \epsilon_{t+1}, \quad (113)$$

and

$$m_t = \psi_v v_t + \psi_s s_t + \psi_\zeta \frac{\text{Cov}_t(v_{t+1}, \epsilon_{t+1})}{v_t^{3/2}}, \quad (114)$$

where the first two terms are conditional variance and conditional skewness and the last term captures the variance risk. The information set is at time t .

A.6. Pricing kernel for a recursive utility example

Following [Boguth and Kuehn \(2013\)](#), the pricing kernel can be written as

$$m_{t+1} \approx k - \gamma \Delta c_{t+1} - (1 - \theta) \Delta z_{t+1}, \quad (115)$$

where γ is the relative risk aversion (RRA), Δc_{t+1} is the consumption growth rate, $\theta = (1 - \gamma)/(1 - 1/\psi)$, ψ is the EIS, and Δz_{t+1} is the changes in the logarithm of the wealth consumption ratio. The corresponding asset pricing equation is

$$E_t[R_{i,t+1}] - R_f = \frac{\gamma}{R_f} \text{Cov}_t(\Delta c_{t+1}, R_{i,t+1}) + \frac{(1 - \theta)}{R_f} \text{Cov}_t(\Delta z_{t+1}, R_{i,t+1}). \quad (116)$$

Since the consumption wealth ratio is not available in closed-form, [Boguth and Kuehn \(2013\)](#) use an affine approximation as

$$\Delta z_{t+1} = A \Delta \mu_{c,t+1} + B \Delta \sigma_{c,t+1}, \quad (117)$$

where $\Delta \mu_{c,t+1}$ is the changes in the mean consumption growth rate, $\Delta \sigma_{c,t+1}$ is the changes in the consumption growth volatility. With this approximation, the asset pricing equation becomes

$$E_t[R_{i,t+1}] - R_f = \frac{\gamma}{R_f} \text{Cov}_t(\Delta c_{t+1}, R_{i,t+1}) + \frac{A(1 - \theta)}{R_f} \text{Cov}_t(\Delta \mu_{c,t+1}, R_{i,t+1}) + \frac{B(1 - \theta)}{R_f} \text{Cov}_t(\Delta \sigma_{c,t+1}, R_{i,t+1}). \quad (118)$$

Empirically, they test the asset pricing equation using a cross-section of portfolio and stock returns and find that the price of risk for changes in the mean consumption

growth rate is not priced whereas the changes in consumption growth volatility is significantly priced. Therefore, we can adapt the following empirical asset pricing model from [Boguth and Kuehn \(2013\)](#):

$$E_t[R_{i,t+1}] - R_f = \frac{\gamma}{R_f} \text{Cov}_t(\Delta c_{t+1}, R_{i,t+1}) + \frac{B(1 - \theta)}{R_f} \text{Cov}_t(\Delta \sigma_{c,t+1}, R_{i,t+1}). \quad (119)$$

If the market return is a scaled process of the consumption process ([Campbell, 1993](#)), then the above equation becomes

$$E_t[R_{i,t+1}] - R_f = \tilde{\psi}_v \text{Cov}_t(\Delta R_{m,t+1}, R_{i,t+1}) + \tilde{\psi}_\eta \text{Cov}_t(\Delta \sigma_{t+1}, R_{i,t+1}), \quad (120)$$

in which we have substituted the coefficients $\tilde{\psi}_v$ and $\tilde{\psi}_\eta$ for monotone transformations of γ/R_f and $B(1 - \theta)/R_f$, respectively. Applied to the market itself, we have

$$E_t[R_{m,t+1}] - R_f = \tilde{\psi}_v \text{Cov}_t(\Delta R_{m,t+1}, R_{m,t+1}) + \tilde{\psi}_\eta \text{Cov}_t(\Delta \sigma_{t+1}, R_{m,t+1}) = \tilde{\psi}_v \text{Var}_t(R_{m,t+1}) + \tilde{\psi}_\eta \text{Cov}_t(\sigma_{t+1}, R_{m,t+1}). \quad (121)$$

The market price of risk for volatility risk is captured by $\tilde{\psi}_\eta \text{Var}(\sigma_{t+1})$ and is estimated to be negative by [Boguth and Kuehn \(2013\)](#). This supports the common assumption that has been used in the long-run risk literature, that the EIS is greater than 1/RRA.

References

Andersen, T.G., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. *Journal of Finance* 62, 1239–1284.

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61, 259–299.

Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies* 16, 101–143.

Bansal, R., Viswanathan, S., 1993. No-arbitrage and arbitrage pricing: a new approach. *Journal of Finance* 48, 1231–1262.

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance* 59, 1481–1509.

Bates, D.S., 2000. Post-'87 Crash fears in the S&P 500 futures option market. *Journal of Econometrics* 94, 181–238.

Boguth, O., Kuehn, A., 2013. Consumption volatility risk. *Journal of Finance* 10.1111/jofi.12058.

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

Campbell, J., 1993. Intertemporal asset pricing without consumption data. *American Economic Review* 83, 487–512.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.

Chabi-Yo, F., 2012. Pricing kernels with stochastic skewness and volatility risk. *Management Science* 58, 24–640.

Chabi-Yo, F., Ghysels, E., Renault, E., 2007. Disentangling the effects of heterogeneous beliefs and preferences on asset prices. Unpublished Working Paper. Ohio State University and University of North Carolina at Chapel Hill.

Chan, W.H., Feng, L., 2012. Time-varying jump risk premia in stock index futures returns. *Journal of Futures Markets* 32, 639–659.

Chan, W.H., Maheu, J.M., 2002. Conditional jump dynamics in stock market returns. *Journal of Business & Economic Statistics* 20, 377–389.

Chang, B.Y., Christoffersen, P., Jacobs, K., 2013. Market skewness risk and the cross-section of stock returns. *Journal of Financial Economics* 107, 46–68.

Chapman, D., 1997. Approximating the asset pricing kernel. *Journal of Finance* 52, 1383–1410.

- Chernov, M., Gallant, R.A., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. *Journal of Econometrics* 116, 225–257.
- Christoffersen, P., Jacobs, K., Ornathanalai, C., 2012. Exploring time-varying jump intensities: evidence from S&P500 returns and options. *Journal of Financial Economics* 106, 447–472.
- Conrad, J., Dittmar, R.F., Ghysels, E., 2013. Ex ante skewness and expected stock returns. *Journal of Finance* 68, 85–124.
- Das, S.R., Sundaram, R.K., 1997. Taming the skew: higher-order moments in modeling asset price processes in finance. NBER Working Paper 5976.
- Dittmar, R.F., 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross-section of equity returns. *Journal of Finance* 57, 369–403.
- Duan, J.-C., Zhang, W., 2010. Forward-looking market risk premium. Unpublished working paper. National University of Singapore and University of Munster.
- Elkamhi, R., Ornathanalai, C., 2009. The market jump risk and the price structure of individual equity options. Unpublished Working Paper. Rotman School of Management, University of Toronto.
- Engle, R.F., Lee, G.G.J., 1999. A long-run and short-run component model of stock return volatility. In: Engle, R., White, H. (Eds.), *Cointegration, Causality, and Forecasting*, Oxford University Press.
- Epstein, L., Zin, S., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica* 57, 937–969.
- Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and options prices. *Journal of Finance* 59, 1367–1403.
- Friend, I., Westerfield, R., 1980. Co-skewness and capital asset pricing. *Journal of Finance* 35, 897–914.
- Guidolin, M., Timmermann, A., 2008. International asset allocation under regime switching, skew, and kurtosis preferences. *Review of Financial Studies* 21, 889–935.
- Harvey, C.R., Siddique, A., 1999. Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis* 34, 465–487.
- Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. *Journal of Finance* 55, 1263–1295.
- Hwang, S., Satchell, S.E., 1999. Modeling emerging market risk premia using higher moments. *International Journal of Finance and Economics* 4, 271–296.
- Johannes, M., Kumar, R., Polson, N.G., 1999. State dependent jump models: How do U.S. equity indices jump? Working Paper. Graduate School of Business, University of Chicago.
- Kimball, M.S., 1990. Precautionary saving in the small and in the large. *Econometrica* 58, 53–73.
- Kimball, M.S., 1993. Standard risk aversion. *Econometrica* 61, 589–611.
- Kraus, A., Litzenberger, R., 1976. Skewness preference and the valuation of risk assets. *Journal of Finance* 31, 1085–1100.
- Lim, K.-G., 1989. A new test of the three-moment capital asset pricing model. *Journal of Financial and Quantitative Analysis* 24, 205–216.
- Maheu, J.M., McCurdy, T.H., 2004. News arrival, jump dynamics, and volatility components for individual stock returns. *Journal of Finance* 59, 755–794.
- Maheu, J.M., McCurdy, T.H., 2007. Components of market risk and return. *Journal of Financial Econometrics* 5, 560–590.
- Merton, R.C., 1980. On Estimating the expected return on the market: an exploratory investigation. *Journal of Financial Economics* 8, 323–361.
- Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50.
- Rossi, A., Timmermann, A., 2010. What is the shape of the risk-return relation? Unpublished working paper. University of California, San Diego.
- Routledge, B.R., Zin, S.E., 2010. Generalized disappointment aversion and asset prices. *Journal of Finance* 65, 1303–1332.
- Sears, S.R., Wei, K.-C.J., 1985. Asset pricing, higher moments, and the market risk premium. *Journal of Finance* 40, 1251–1253.
- Smith, D.R., 2007. Conditional coskewness and asset pricing. *Journal of Empirical Finance* 14, 91–119.