CrossMark

RESEARCH ARTICLE

The skewness risk premium in equilibrium and stock return predictability

Hiroshi Sasaki[1]

Abstract In this study, we investigate the skewness risk premium in the financial market under a general equilibrium setting. Extending the long-run risks (LRR) model proposed by Bansal and Yaron (J Financ 59:1481–1509, 2004) by introducing a stochastic jump intensity for jumps in the LRR factor and the variance of consumption growth rate, we provide an explicit representation for the skewness risk premium, as well as the volatility risk premium, in equilibrium. On the basis of the representation for the skewness risk premium, we propose a possible reason for the empirical facts of time-varying and negative risk-neutral skewness. Moreover, we also provide an equity risk premium representation of a linear factor pricing model with the variance and skewness risk premiums. The empirical results imply that the skewness risk premium, as well as the variance risk premium, has superior predictive power for future aggregate stock market index returns, which are consistent with the theoretical implication derived by our model. Compared with the variance risk premium, the results show that the skewness risk premium plays an independent and essential role for predicting the market index returns.

✉ Hiroshi Sasaki
   hiroshi-sasaki@fintec.co.jp

1 Mizuho-DL Financial Technology Co., Ltd., 2-4-1 Koujimachi, Chiyoda-ku, Tokyo 102-0083, Japan

🍻 Springer

JEL Classification   C · D · G

1 Introduction

The concern with the information content in option-implied distributions has been growing for the last several years. In particular, there has been a growing interest in the information of the difference between option-implied and realized distributions, which is usually recognized as a risk premium required by the representative agent, from the point of view of the financial risk management and asset pricing implications. This paper investigates the risk premiums in higher order moments, especially in the skewness, of financial asset returns under a general equilibrium setting. In this paper, each of the risk premiums in higher order moments of financial asset returns is defined by the difference between two expected values of the moment under the risk-neutral and physical probability measures, respectively.

In recent years, there remains an ever-increasing interest and challenge to develop an entirely self-contained equilibrium-based explanation for the nonzero volatility (or variance) risk premium[1] and its predictability for stock index returns. To the best of our knowledge, the first attempt to demonstrate the existence of the volatility risk premium based on a general equilibrium market model is made by Eraker (2008). Eraker (2008) develops an equilibrium explanation for the volatility risk premium based on the long-run risks (LRR) model[2] which emphasizes the role of long-run risks, that is, low-frequency movements in consumption growth rates and volatility, in accounting for a wide range of asset pricing puzzles. The LRR model features an Epstein and Zin (1989) utility function with an investor preference for early resolution of uncertainty and contains (i) a persistent expected consumption growth component and (ii) long-run variation in consumption volatility. On the basis of the LRR model, Eraker (2008) studies the volatility risk premium through the framework of a general equilibrium model.

In addition to the development of an entirely self-contained equilibrium-based explanation for the risk premiums in higher order moments, several academic studies related to those risk premiums are provided in recent years. For example, motivated by fruitful implications from the LRR model pioneered by Bansal and Yaron (2004), Bollerslev et al. (2009) investigate the stock return predictability of the variance risk premium from the point of view of a general equilibrium setting based on the LRR model framework. They show that the difference between option-implied and realized variation, or the variance risk premium, is able to explain a nontrivial fraction of the time-series variation in post-1990 aggregate stock market returns, with high (low) premia predicting high (low) future returns. The magnitude of the predictability is particularly strong at the intermediate quarterly return horizon, where it dominates that afforded by other popular predictor variables, such as the P/E ratio, the default spread, and the consumption-wealth ratio.

---

[1] The volatility (variance) risk premium is defined by the difference between two expected values of the volatility (variance) under the risk-neutral and physical probability measures, respectively.

[2] The long-run risks model, which is a stylized self-contained general equilibrium model incorporating the effects of time-varying economic uncertainty, is pioneered by Bansal and Yaron (2004).

[Drechsler and Yaron](2011) also show the predictability of the variance risk premium for stock index returns based on an extended LRR model with jumps in uncertainty and the long-run component of cash-flows. They demonstrate that a risk aversion greater than one and a preference for early resolution of uncertainty correctly signs the variance risk premium and the coefficient from a predictive regression of returns on the variance risk premium.

All of the studies cited above focus only on the variance risk premium which is required by a representative investor due to the stochastic nature of asset return variance. Conversely, as far as we know, there are few reports about the risk premium which compensates for uncertainty of the third moment, that is, the skewness, of asset returns. In this paper, we demonstrate that the skewness risk premium, defined by the difference between two expected values of the skewness under the risk-neutral and physical probability measures, respectively, also captures attitudes toward economic uncertainty as well as the variance risk premium. Among recent studies on self-contained equilibrium-based models for the nonzero variance risk premium referred above, all of the studies except for [Drechsler and Yaron](2011) model the processes of both the variance of consumption growth rate and the LRR factor as conditional normal, so that the one-step-ahead conditional distribution of the market return in each time is also conditional normal and, as a result, the skewness of that distribution must be zero. Therefore, the models proposed by those studies can not explain the *negative* risk-neutral skewness, which is found by the previous studies such as those by [Aït-Sahalia and Lo](1998) and [Aït-Sahalia et al.](2001). They document several empirical features of the state price density for the S&P500 index option market over time, including the term structures of mean returns, volatility, skewness, and kurtosis, that are implied by option-implied distributions. In particular, They show that the nonparametric state price densities are negatively skewed, have fatter tails and the amount of skewness and kurtosis both increase with maturity.

We show that jump components in the LRR factor and/or the variance of consumption growth rate can explain the nonzero (and negative) skewness of the one-step-ahead asset return distribution. To the best of our knowledge, [Drechsler and Yaron](2011) is the first paper that indicates an important role for transient non-Gaussian shocks (jumps) to fundamentals such as the LRR-factor and the variance of consumption growth rate for understanding how perceptions of economic uncertainty and cash-flow risk manifest themselves in asset prices. However, in [Drechsler and Yaron](2011), they assume that the jump intensity process λ_t has an affine structure of $\lambda_t = l_0 + l_1 \sigma_t^2$, where $l_0, l_1 > 0$ and σ_t^2 is the variance of consumption growth rate, and under this assumption, it is unfortunately not possible to explain an empirical fact on a simultaneous relation between monthly stock returns and monthly changes of the option-implied skewness:

$$
\begin{aligned}
r_{m,t+1} &= 0.006 - 0.019 \times \triangle ISkew_{t+1}, \\
&\quad (2.46) \quad (-3.46) \\
r_{m,t+1} &= 0.006 - 0.007 \times \triangle VIX_{t+1} - 0.016 \times \triangle ISkew_{t+1}, \\
&\quad (3.33) \quad (-16.00) \quad (-3.94)
\end{aligned}
\tag{1}
$$

where $r_{m,t+1}$ is the monthly return of the S&P500 Total Return Index from time t to $t+1$, $\triangle VIX_{t+1}$ is the monthly change of option-implied volatility calculated with the

CBOE's VIX from time t to $t + 1$, and $\triangle ISkew_{t+1}$ is the monthly change of option-implied skewness calculated with the CBOE's Skew Index from time t to $t + 1$.[3] These results are obtained based on the monthly data from Jan-1990 to Aug-2012. Under the assumption on the jump intensity process in Drechsler and Yaron (2011), however, we can confirm that the regression parameters to $\triangle ISkew_{t+1}$ in the above regression models should be positive.

Moreover, such assumption in Drechsler and Yaron (2011) can not explain another empirical fact of low correlation (in absolute value) between $\triangle VIX_t$ and $\triangle ISkew_t$. The jump intensity model assumed in Drechsler and Yaron (2011) leads to the *perfect* correlation between $\triangle VIX_t$ and $\triangle ISkew_t$ because the jump intensity process in their paper is driven by only one factor, that is, σ_t^2. However, actual correlation value between $\triangle VIX_t$ and $\triangle ISkew_t$ in the period from Jan-1990 to Aug-2012 is nearly zero,[4] so that the assumption for jump intensity in Drechsler and Yaron (2011) can not explain such empirical fact.

In this paper, we propose an extension of the LRR models pioneered by Bansal and Yaron (2004) and Drechsler and Yaron (2011). Our model contains a rich set of transient dynamics and can quantitatively account for the time variation and asset return predictability of the skewness premium as well as the variance risk premium. In particular, we introduce a stochastic jump intensity for transient jumps to fundamentals such as the LRR factor and the variance of consumption growth rate, and show that this additional introduction of a stochastic jump intensity enables our model to capture the various empirical aspects of the stock index returns and its option-implied moments including the facts cited above. Christoffersen et al. (2012) find very strong support for time-varying jump intensities for S&P500 index returns and they show that, compared to the risk premium on dynamic volatility, the risk premium on the dynamic jump intensity has a much larger impact on option prices. We find that the existence of the negative skewness and the skewness risk premium observed in historical data have a close relationship with the existence of the jumps and the jump risk premium of the economic uncertainty, respectively.

This paper also shows that the skewness of asset return distribution and the skewness risk premium, which compensates for the stochastic nature of the skewness, are both time-varying due to the stochastic nature of the jump intensity for transient jumps in both the LRR factor and the variance of consumption growth rate. Moreover, providing an equity risk premium representation of a linear factor pricing model with time-varying variance and skewness risk premiums, we find that those risk premiums can explain a nontrivial fraction of the time series variation in the aggregate stock market returns. Simplified preliminary tests of regression models provide empirical evidence in which the skewness risk premium, as well as the variance risk premium, has superior predictive power for future aggregate stock market index returns. Compared with the variance risk premium, the results also show that the skewness risk premium plays an independent and essential role for predicting the market index returns.

---

[3] See Sects. 4.1 and 4.2 for the details of calculation methodology for option-implied skewness with the CBOE's Skew Index.

[4] We can find that the correlation value between VIX_t^2 and $ISkew_t$ in the same period is also nearly zero (see Table 2).

The remainder of this paper is organized as follows. Section 2 outlines the basic theoretical model with jumps in consumption growth rate and its volatility, shows how equilibrium is derived for our model economy, and highlights its key features. In particular, we provide an equity risk premium representation of a linear factor pricing model with time-varying variance and skewness risk premiums. Section 3 provides the implications from a calibrated version of the theoretical equity risk premium representation of a linear factor pricing model derived in Sect. 2 to help guide and interpret our subsequent empirical reduced form predictability regressions. Section 4 describes the data used for examining the equity risk premium representation empirically and discusses the results from the predictive regressions on the stock returns to the variance and the skewness risk premiums with historical data. Section 5 provides concluding remarks.

2 Model framework

2.1 Model setup and assumptions

The underlying environment is a discrete time endowment economy. The representative agent's preferences on the consumption stream are of the Epstein and Zin (1989) form, allowing for the separation of risk aversion and the intertemporal elasticity of substitution (IES). Thus, the agent maximizes his lifetime utility, which is defined recursively as

$$V_t = \left[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \left(\mathbb{E}_t[V_{t+1}^{1-\gamma}] \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}, \tag{2}$$

where C_t is consumption at time t, $0 < \delta < 1$ reflects the agent's time preference, γ is the coefficient of risk aversion, $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, and ψ is the intertemporal elasticity of substitution (IES). This preference structure collapses to a standard CRRA utility representation if $\gamma = \frac{1}{\psi}$, that is, $\theta = 1$, and in this case, only innovations to consumption are priced. In the following, based on the result provided by Bansal and Yaron (2004) we assume that both γ and ψ are larger than one. It then holds that $\gamma > \frac{1}{\psi}$, which implies $\theta < 0$. With this choice, the investor has a preference for early resolution of uncertainty (Bansal and Yaron 2004). Then, not only consumption risk is priced, but state variables carry risk premia, too. The parameter restrictions also ensure that the signs of the risk premia are in line with economic intuition, and that a worsening of economic conditions leads to a decrease in asset prices.

Utility maximization is subject to the budget constraint:

$$W_{t+1} = (W_t - C_t) R_{c,t+1},$$

where W_t is the wealth of the agent and $R_{c,t}$ is the return on all invested wealth. As shown in Epstein and Zin (1989), for any asset j, the first-order condition yields the following Euler condition:

$$\mathbb{E}_t \left[\exp(m_{t+1} + r_{j,t+1}) \right] = 1, \tag{3}$$

where $r_{j,t+1}$ is the log of the gross return on asset j and m_{t+1} is the log of the intertemporal marginal rate of substitution (IMRS), which is given by $m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} \triangle c_{t+1} + (\theta - 1)r_{c,t+1}$. Here, $r_{c,t+1}$ is $\log R_{c,t+1}$ and $\triangle c_{t+1}$ is the change in $\log C_t$, that is, $\log \left(\frac{C_{t+1}}{C_t} \right)$.

We model consumption and dividend growth rates, $g_{t+1} \equiv \log(\frac{C_{t+1}}{C_t})$ and $g_{d,t+1} \equiv \log(\frac{D_{t+1}}{D_t})$ where D_t is dividend at time t, respectively, as containing a small persistent predictable component x_t, which determines the conditional expectation of consumption growth,

$$
\begin{aligned}
x_{t+1} &= \rho_x x_t + \varphi_e \sigma_t e_{t+1} + J_{x,t+1}, \\
g_{t+1} &= \mu_g + x_t + \varphi_\eta \sigma_t \eta_{t+1}, \\
g_{d,t+1} &= \mu_d + \rho_d x_t + \varphi_\zeta \sigma_t \zeta_{t+1},
\end{aligned}
\tag{4}
$$

where $\varphi_e, \varphi_\eta, \varphi_\zeta, \rho_x, \rho_d > 0$, $\mu_g, \mu_d \in \mathbb{R}$, e_t, η_t, and ζ_t are mutually independent $i.i.d. N(0, 1)$ processes, and $J_{x,t+1}$ is a compound-Poisson process represented by $J_{x,t+1} \equiv \sum_{j=1}^{N^x_{t+1}} \epsilon^j_x$ where N^x_{t+1} is the Poisson counting process for that jump component whose the intensity process is $\lambda_{x,t+1} \equiv l_x \lambda_{t+1}, l_x > 0$, and $\epsilon^j_x \sim i.i.d.\ N(0, \delta^2_x)$, $\delta_x > 0$, is the size of the jump that occurs upon the N^x_{t+1}.

Furthermore, we also model the dynamics of the volatility as follows:

$$
\begin{aligned}
\sigma^2_{t+1} &= \mu_\sigma + \rho_\sigma \sigma^2_t + \sqrt{q_t} w_{t+1} + J_{\sigma^2,t+1}, \\
q_{t+1} &= \mu_q + \rho_q q_t + \varphi_\xi \sqrt{q_t} \xi_{t+1},
\end{aligned}
\tag{5}
$$

where the parameters satisfy $\mu_\sigma > 0, \mu_q > 0, |\rho_\sigma| < 1, |\rho_q| < 1, \varphi_\xi > 0$, and w_t and ξ_t are mutually independent $i.i.d. N(0, 1)$ processes and are also independent of each of e_t, η_t, and ξ_t. $J_{\sigma^2,t+1}$ is a compound-Poisson process, which is represented by $J_{\sigma^2,t+1} \equiv \sum_{j=1}^{N^{\sigma^2}_{t+1}} \epsilon^j_{\sigma^2}$ where $N^{\sigma^2}_{t+1}$ is the Poisson counting process for that jump component whose the intensity process is $\lambda_{\sigma^2,t+1} \equiv l_{\sigma^2} \lambda_{t+1}, l_{\sigma^2} > 0$, and $\epsilon^j_{\sigma^2} \sim i.i.d.\ N(0, \delta^2_{\sigma^2}), \delta_{\sigma^2} > 0$, is the size of the jump that occurs upon the $N^{\sigma^2}_{t+1}$. We assume that N^x_{t+1} and $N^{\sigma^2}_{t+1}$ are mutually independent and ϵ^j_x and $\epsilon^j_{\sigma^2}$ are too. The stochastic variance process σ^2_t represents time-varying economic uncertainty in consumption growth with the variance-of-variance process q_t in effect inducing an additional source of temporal variation in that same process. We model the variance-of-variance process q_t in the same fashion as Bollerslev et al. (2009).

Importantly, the jump intensity dynamics for λ_t is newly introduced in our economy, which is represented by the following discrete-time stochastic process,

$$
\lambda_{t+1} = \mu_\lambda + \rho_\lambda \lambda_t + \varphi_u \sqrt{q_t}(\rho \xi_{t+1} + \sqrt{1 - \rho^2} u_{t+1}),
\tag{6}
$$

where $\mu_\lambda > 0, |\rho_\lambda| < 1, |\rho| \leq 1$, and u_t is an $i.i.d. N(0, 1)$ process which is independent of each of $e_t, \eta_t, \zeta_t, w_t$, and ξ_t.

One of the notable features of our model setup is this introduction for the jump intensity process (6). Christoffersen et al. (2012) also find very strong support for time-varying jump intensities for S&P500 index returns, and they show that, compared to the risk premium on dynamic volatility, the risk premium on the dynamic jump intensity has a much larger impact on option prices. In the previous studies, Drechsler and Yaron (2011) is the first paper that introduces transient jumps to fundamentals such as the LRR-factor x_t and the variance of consumption growth rate σ_t^2. However, it assumes that the jump intensity process λ_t is represented by an affine structure of $\lambda_t = l_0 + l_1 \sigma_t^2$ where $l_0, l_1 > 0$. As mentioned in the introduction of this paper, such assumption for the jump intensity process can not explain empirical facts of regression (1) and nearly zero correlation between option-implied volatility and option-implied skewness. We extend the LRR models of Bansal and Yaron (2004) and Drechsler and Yaron (2011) so as to introduce a stochastic jump intensity of (6) into the economy. As shown in the following, this introduction enables our model to have a consistency with the empirical facts mentioned above and plays a key role in describing the characteristics of asset return distributions.

2.2 The model solution in equilibrium

We distinguish between the unobservable return on a claim to aggregate consumption, $R_{c,t+1}$, and the observable return on the market portfolio, $R_{m,t+1}$: the latter is the return on the aggregate dividend claim. Solving our model numerically, we demonstrate the mechanisms working in our model via approximate analytical solutions in the same fashion as the previous studies such as those by Bansal and Yaron (2004), Bollerslev et al. (2009), Drechsler and Yaron (2011), etc. To derive these solutions for our model, we use the standard approximation utilized in Campbell and Shiller (1988),

$$r_{c,t+1} = \kappa_0 + \kappa_1 v_{t+1} - v_t + g_{t+1}, \tag{7}$$

where lowercase letters refer to logs, so that $r_{c,t+1} = \log(R_{c,t+1})$ is the continuous return, $v_t = \log(\frac{P_t}{C_t})$ is the log price-consumption ratio of the asset that pays the consumption endowment, $\{C_{t+i}\}_{i=1}^{\infty}$, and κ_0 and κ_1 are approximating constants that both depend only on the average level of v.[5] Analogously, $r_{m,t+l}$ and $v_{m,t+1}$ correspond to the market return and its log price-dividend ratio and the similar approximation presented below can also be derived:

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} v_{m,t+1} - v_{m,t} + g_{d,t+1}. \tag{8}$$

The standard solution method for finding the equilibrium in a model like the one defined above then consists in conjecturing solutions for v_t and $v_{m,t}$ as an affine

---

[5] Note that $\kappa_1 = \frac{\exp(\bar{v})}{1+\exp(\bar{v})}$ and this value is approximately 0.997 (cf) Bansal and Yaron (2004), which is also consistent with magnitudes used in Campbell and Shiller (1988).

function of the state variables, x_t, σ_t^2, q_t, and λ_t,

$$v_t = A_0 + A_x x_t + A_\sigma \sigma_t^2 + A_q q_t + A_\lambda \lambda_t, \tag{9}$$

$$v_{m,t} = A_{0,m} + A_{x,m} x_t + A_{\sigma,m} \sigma_t^2 + A_{q,m} q_t + A_{\lambda,m} \lambda_t, \tag{10}$$

respectively, solving for the coefficients A_0, A_x, A_σ, A_q, and A_λ in v_t and for the coefficients $A_{0,m}$, $A_{x,m}$, $A_{\sigma,m}$, $A_{q,m}$, and $A_{\lambda,m}$ in $v_{m,t}$.

Substituting (9) for (7), we have a temporal representation for $r_{c,t+1}$ with the state variables, x_t, σ_t^2, q_t, and λ_t, and furthermore, substituting this $r_{c,t+1}$ for the Euler equation (3), we can derive an identity with those state variables. Solving the identity in the same manner as Bansal and Yaron (2004), Bollerslev et al. (2009) and Drechsler and Yaron (2011), we can derive the equilibrium solutions for the four parameters as follows:

$$A_x = \frac{\gamma - 1}{\theta(\kappa_1 \rho_x - 1)},$$

$$A_\sigma = -\frac{1}{2} \frac{(1-\gamma)^2 \varphi_\eta^2 + \theta^2 \kappa_1^2 A_x^2 \varphi_e^2}{\theta(\kappa_1 \rho_\sigma - 1)},$$

$$A_\lambda = \frac{2 - \exp(\frac{1}{2}\theta^2 \kappa_1^2 A_x^2 \delta_x^2) - \exp(\frac{1}{2}\theta^2 \kappa_1^2 A_\sigma^2 \delta_\sigma^2)}{\theta(\kappa_1 \rho_\lambda - 1)}, \tag{11}$$

A_q is a solution of the quadratic equation presented below:

$$\theta A_q(\kappa_1 \rho_q - 1) + \frac{\theta^2 \kappa_1^2}{2} \left[A_\sigma^2 + A_q^2 \varphi_\xi^2 + 2 A_q A_\lambda \varphi_\xi \varphi_u \rho + A_\lambda^2 \varphi_u^2 \right] = 0.$$

Considering the expressions of (11), the following proposition can be proven easily:

Proposition 1 *If $\gamma > 1$ and $\psi > 1$, then, $A_x > 0$, $A_\sigma < 0$, $A_q < 0$, and $A_\lambda < 0$.*

The above proposition suggests that if the IES and risk aversion are higher than 1, a rise in each of the state variables of σ_t^2, q_t, and λ_t lowers the price-consumption ratio.

Having solved for $r_{c,t+1}$ with the four parameters derived above, we can substitute it (and $\triangle c_{t+1} = g_{t+1}$) into m_{t+1} to obtain an expression for the conditional innovation to the log pricing kernel at time $t + 1$:

$$
\begin{aligned}
& m_{t+1} - \mathbb{E}_t[m_{t+1}] \\
&= \theta \log \delta - \frac{\theta}{\psi} \triangle c_{t+1} + (\theta - 1)r_{c,t+1} \\
&\quad - \mathbb{E}_t \left[\theta \log \delta - \frac{\theta}{\psi} \triangle c_{t+1} + (\theta - 1)r_{c,t+1} \right] \\
&= -\Lambda^t (G_t z_{t+1} + J_{t+1} - \mathbb{E}_t[J_{t+1}]), \tag{12}
\end{aligned}
$$

where

$$\Lambda \equiv \begin{pmatrix} \gamma & (1-\theta)\kappa_1 A_x & (1-\theta)\kappa_1 A_\sigma & (1-\theta)\kappa_1 A_q & (1-\theta)\kappa_1 A_\lambda & 0 \end{pmatrix}^t,$$

$$G_t \equiv \begin{pmatrix} \varphi_\eta \sigma_t & 0 & 0 & 0 & 0 & 0 \\ 0 & \varphi_e \sigma_t & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{q_t} & 0 & 0 & 0 \\ 0 & 0 & 0 & \varphi_\xi \sqrt{q_t} & 0 & 0 \\ 0 & 0 & 0 & \rho\varphi_u\sqrt{q_t} & \varphi_u\sqrt{1-\rho^2}\sqrt{q_t} & 0 \\ 0 & 0 & 0 & 0 & 0 & \varphi_\zeta \sigma_t \end{pmatrix}, \tag{13}$$

$$z_{t+1} \equiv \begin{pmatrix} \eta_{t+1} & e_{t+1} & w_{t+1} & \xi_{t+1} & u_{t+1} & \zeta_{t+1} \end{pmatrix}^t,$$

$$J_{t+1} \equiv \begin{pmatrix} 0 & J_{x,t+1} & J_{\sigma^2,t+1} & 0 & 0 & 0 \end{pmatrix}^t,$$

$$\mathbb{E}_t[J_{t+1}] \equiv \begin{pmatrix} 0 & \mathbb{E}_t[J_{x,t+1}] & \mathbb{E}_t[J_{\sigma^2,t+1}] & 0 & 0 & 0 \end{pmatrix}^t.$$

Λ can be interpreted as the price of risk for Gaussian shocks and also the sensitivity of the IMRS to the jump shocks. From the expression of Λ, one can see that the prices of risks are determined by the A coefficients, that is, A_x, A_σ, A_q, and A_λ. The expression of Λ also shows that the signs of the risk prices depend on the signs of the A coefficients and $(1-\theta)$. In particular, when $\gamma = \frac{1}{\psi}$, $\theta = 1$, and we are in the case of constant relative risk aversion (CRRA) preferences, it is clear that only the transient shock to consumption $z_{c,t+1} \equiv \eta_{t+1}$ is priced, and prices do not separately reflect the risk of shocks to x_t (long-run risk), σ_t^2 (volatility-related risk), q_t (variance-of-variance-related risk), and λ_t (jump intensity-related risk).

In the discussion and calibrations explored below, we especially focus on the case in which the agent's risk aversion γ and the IES ψ are both greater than 1, which implies that $A_x > 0$, $A_\sigma < 0$, $A_q < 0$, and $A_\lambda < 0$ by the proposition provided above. Thus, positive shocks to long-run growth decrease the IMRS, while positive shocks to the levels of the other state variables, σ_t^2, q_t, and λ_t, increase the IMRS. Note that in this case, since $(1-\theta) > 0$, each of the A coefficients has the same sign as the corresponding price of risk.

To study the risk premiums in higher-order moments of the market returns, we first need to solve for the market return. A share in the market is modeled as a claim to a dividend with growth process given by $g_{d,t}$. To solve for the price of a market share, we proceed along the same lines as for the consumption claim and solve for $v_{m,t+1}$, the log price-dividend ratio of the market, by using the the conjecture (10), Campbell and Shiller (1988)-approximation (8), and the Euler equation (3).[6]

With the equilibrium solutions for the parameters of $A_{x,m}$, $A_{\sigma,m}$, $A_{q,m}$, and $A_{\lambda,m}$ in (10), we can obtain an expression for $r_{m,t+1}$ in terms of the state variables and its innovations by substituting the expression for $v_{m,t(+1)}$ into (8):

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} A_{0,m} + \kappa_{1,m} A_{\sigma,m}\mu_d + \kappa_{1,m}A_{q,m}\mu_g + \kappa_{1,m}A_{\lambda,m}\mu_\lambda - A_{0,m} + \mu_d$$
$$+ (\kappa_{1,m}A_{x,m}\rho_x - A_{x,m} + \rho_d)x_t$$

[6] Because the datails of the four parameters, $A_{x,m}$, A_{σ_m}, $A_{q,m}$, and $A_{\lambda,m}$, are insignificant and do not affect the discussion explored in the following at all. For simplicity, we express the parameters, $A_{x,m}$, A_{σ_m}, $A_{q,m}$, and $A_{\lambda,m}$, as they are and do not show explicit representations of those parameters in this paper.

$$+ (\kappa_{1,m} A_{\sigma,m} \rho_\sigma - A_{\sigma,m}) \sigma_t^2$$
$$+ (\kappa_{1,m} A_{q,m} \rho_q - A_{q,m}) q_t$$
$$+ (\kappa_{1,m} A_{\lambda,m} \rho_\lambda - A_{\lambda,m}) \lambda_t$$
$$+ \kappa_{1,m} A_{x,m} \varphi_e \sigma_t e_{t+1} + \kappa_{1,m} A_{\sigma,m} \sqrt{q_t} w_{t+1}$$
$$+ \kappa_{1,m} (A_{q,m} \varphi_\xi + A_{\lambda,m} \varphi_u \rho) \sqrt{q_t} \xi_{t+1}$$
$$+ \kappa_{1,m} A_{\lambda,m} \varphi_u \sqrt{1-\rho^2} \sqrt{q_t} u_{t+1} + \varphi_\zeta \sigma_t \zeta_{t+1}$$
$$+ \kappa_{1,m} A_{x,m} J_{x,t+1} + \kappa_{1,m} A_{\sigma,m} J_{\sigma^2,t+1}$$
$$= r_0 + (B_r^t F - A_m^t) Y_t + B_r^t G_t z_{t+1} + B_r^t J_{t+1}, \tag{14}$$

where

$$r_0 \equiv \kappa_{0,m} + (\kappa_{1,m} - 1) A_{0,m} + (\kappa_{1,m} A_{\sigma,m} + 1) \mu_d + \kappa_{1,m} A_{q,m} \mu_g + \kappa_{1,m} A_{\lambda,m} \mu_\lambda,$$
$$B_r \equiv \kappa_{1,m} A_m + e_d,$$

$$A_m \equiv \begin{pmatrix} 0 \\ A_{x,m} \\ A_{\sigma,m} \\ A_{q,m} \\ A_{\lambda,m} \\ 0 \end{pmatrix}, \ e_d \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \ F \equiv \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \rho_x & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_\sigma & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_q & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho_\lambda & 0 \\ 0 & \rho_d & 0 & 0 & 0 & 0 \end{pmatrix}, \ Y_t \equiv \begin{pmatrix} g_t \\ x_t \\ \sigma_t^2 \\ q_t \\ \lambda_t \\ g_{d,t} \end{pmatrix}. \tag{15}$$

2.3 Risk premiums in higher-order moments in equilibrium

Before proceeding to investigating the risk premiums in higher-order moments in equilibrium, we need to provide some further explanation on the jump dynamics and the features of the pricing kernel introduced above.

To handle the jumps, we introduce some notation. $\psi_k(u_k) = \mathbb{E}[\exp(u_k \epsilon_k)]$ (k is x or σ^2) denotes the moment-generating function (mgf) of the jump size ϵ_k. The mgf for the jump component of k, $\mathbb{E}[\exp(u_k J_{k,t+1})]$, then equals $\exp(\Psi_{t,k}(u_k))$, where $\Psi_{t,k}(u_k) = \lambda_{k,t}(\psi_k(u_k) - 1)$. $\Psi_{t,k}$ is called the cumulant-generating function (cgf) of $J_{k,t+1}$ and is a very helpful tool for calculating asset pricing moments. The reason is that its n-th derivative evaluated at 0 equals the n-th central moment of $J_{k,t+1}$.

Regarding the features of the pricing kernel, we can show what described below in line with Drechsler and Yaron (2011). Let us set the Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{M_{t+1}}{\mathbb{E}_t[M_{t+1}]}$, where $\mathbb{P}$ is the physical probability measure and $\mathbb{Q}$ is the risk-neutral probability measure in our economy. From (12), we have $\frac{M_{t+1}}{\mathbb{E}_t[M_{t+1}]} \propto \exp(-\Lambda^t(G_t z_{t+1} + J_{t+1}))$. Since z_{t+1} and J_{t+1} are independent, we can treat their measure transformations between $\mathbb{P}$ and $\mathbb{Q}$ separately. As a consequence, Drechsler and Yaron (2011) show that

$$z_{t+1} \overset{\mathbb{Q}}{\sim} N(-G_t'\Lambda, I), \tag{16}$$

where I is the identity matrix in $\mathbb{R}^{6\times 6}$. That is to say that, under $\mathbb{Q}$, z_{t+1} is still a vector of independent normals with unit variances, but with a shift in the mean.

For the case of J_{t+1}, we could also proceed by transforming the probability density function directly. As guided in Drechsler and Yaron (2011), Proposition (9.6) in Cont and Tankov (2004) shows that under $\mathbb{Q}$, the $J_{t+1,k}$ are still compound Poisson processes, but with cgf given by

$$\Psi^{\mathbb{Q}}_{t,k}(u_k) = \lambda_{k,t}\psi_k(-\Lambda_k)\left(\frac{\psi_k(u_k - \Lambda_k)}{\psi_k(-\Lambda_k)} - 1\right), \qquad (17)$$

where $k = x$ or $k = \sigma^2$, Λ_x denotes the price of risk for the LRR-factor x_t, that is, $(1 - \theta)\kappa_1 A_x$, and Λ_{σ^2} denotes the price of risk for the variance of consumption growth rate, that is, $(1 - \theta)\kappa_1 A_\sigma$. (see (13)) In the following discussion, we use the facts mentioned above to calculate the higher-order moments of the market return and to investigate the risk premiums in the moments.

2.3.1 The variance risk premium in equilibrium

According to Bollerslev et al. (2009) and Drechsler and Yaron (2011), the variance risk premium in equilibrium, vp_t, is defined by

$$vp_t \equiv \mathbb{E}^{\mathbb{Q}}_t[\text{Var}^{\mathbb{Q}}_{t+1}(r_{m,t+2})] - \mathbb{E}^{\mathbb{P}}_t[\text{Var}^{\mathbb{P}}_{t+1}(r_{m,t+2})], \qquad (18)$$

where $\text{Var}^{\mathbb{P}}_{t+1}$ ($\text{Var}^{\mathbb{Q}}_{t+1}$) is the variance operator under the physical (risk-neutral) probability measure at time $t + 1$. We have the following proposition:

Proposition 2 (The Conditional Variance of the Market Return) *The conditional variance of the market return $r_{m,t+2}$ at time $t + 1$ under $\mathbb{P}$ can be expressed as follows:*

$$\text{Var}^{\mathbb{P}}_{t+1}(r_{m,t+2}) = B_r^t(H_{\sigma^2}\sigma^2_{t+1} + H_q q_{t+1})B_r + B_r^{2t}\text{diag}\left(\psi^{(2)}(0)\right)\Pi_{t+1}, \qquad (19)$$

where

$$H_{\sigma^2} \equiv \begin{pmatrix} \varphi_\eta^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \varphi_e^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \varphi_\zeta^2 \end{pmatrix}, \quad H_q \equiv \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varphi_\xi^2 & \rho\varphi_\xi\varphi_u & 0 \\ 0 & 0 & 0 & \rho\varphi_\xi\varphi_u & \varphi_u^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\text{diag}\left(\psi^{(2)}(0)\right) \equiv \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \psi_x^{(2)}(0) & 0 & 0 & 0 & 0 \\ 0 & 0 & \psi_{\sigma^2}^{(2)}(0) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Pi_{t+1} \equiv \begin{pmatrix} 0 \\ \lambda_{x,t+1} \\ \lambda_{\sigma^2,t+1} \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Proof See the "Appendix". □

Under the risk-neutral probability measure $\mathbb{Q}$, the conditional variance of the market return $r_{m,t+2}$ at time $t+1$ also can be obtained in the same manner demonstrated above. As a consequence, we can show the following proposition based on the definition of the variance risk premium (18).

Proposition 3 (The Variance Risk Premium in Equilibrium) *In equilibrium, the variance risk premium at time t, vp_t, is linear to the variance-of-variance, q_t, and the jump intensity, λ_t, and the representation of it is provided as follows:*

$$vp_t = \beta_{vp,c} + \beta_{vp,q} q_t + \beta_{vp,\lambda} \lambda_t, \tag{20}$$

where

$$\beta_{vp,c} \equiv \left[l_x B_r^2(2)(\psi_x^{(2)}(-\Lambda_x) - \psi_x^{(2)}(0)) + l_{\sigma^2} B_r^2(3)(\psi_{\sigma^2}^{(2)}(-\Lambda_{\sigma^2}) - \psi_{\sigma^2}^{(2)}(0)) \right] \mu_\lambda,$$

$$\beta_{vp,q} \equiv -B_r^t \left[\Lambda_{\sigma^2} H_{\sigma^2} + \varphi_\xi (\varphi_\xi \Lambda_q + \rho \varphi_u \Lambda_\lambda) H_q \right] B_r$$
$$- \varphi_u (\rho \varphi_\xi \Lambda_q + \varphi_u \Lambda_\lambda)(l_x B_r^2(2)\psi_x^{(2)}(-\Lambda_x) + l_{\sigma^2} B_r^2(3)\psi_{\sigma^2}^{(2)}(-\Lambda_{\sigma^2})),$$

$$\beta_{vp,\lambda} \equiv B_r^t H_{\sigma^2} B_r \psi_{\sigma^2}^{(1)}(-\Lambda_{\sigma^2})$$
$$+ \left[l_x B_r^2(2)(\psi_x^{(2)}(-\Lambda_x) - \psi_x^{(2)}(0)) + l_{\sigma^2} B_r^2(3)(\psi_{\sigma^2}^{(2)}(-\Lambda_{\sigma^2}) - \psi_{\sigma^2}^{(2)}(0)) \right] \rho_\lambda.$$

Proof See the "Appendix". □

A number of interesting implications arise from the expression (20). In particular, any temporal variation in the endogenously generated variance risk premium is solely due to the variance-of-variance q_t and the jump intensity λ_t. Moreover, provided that $\theta < 0$, $\Lambda_x > 0$, and $\Lambda_{\sigma^2} < 0$, as would be implied by $\gamma > 1$ and $\psi > 1$, the factor loading to the jump intensity, that is, $\beta_{vp,\lambda}$, is guaranteed to be positive, but that to the variance-of-variance, that is, $\beta_{vp,q}$, can be both positive and negative in general. However, if the correlation between the dynamics of the variance-of-variance and that of the jump intensity, that is , ρ, is positive, then $\beta_{vp,q}$ is also guaranteed to be positive due to the facts that $\Lambda_q < 0$ and $\Lambda_\lambda < 0$.

2.3.2 The skewness risk premium in equilibrium

On the basis of the same manner used to derive the expression (19) in the previous subsection, we can also derive the representations for the skewness of the market return under $\mathbb{P}$ and $\mathbb{Q}$, respectively, as follows:

$$\begin{aligned} \text{Skew}_t^{\mathbb{P}}(r_{m,t+1}) &= B_r^{3^t} \text{diag}(\psi^{(3)}(0))\Pi_t, \\ \text{Skew}_t^{\mathbb{Q}}(r_{m,t+1}) &= B_r^{3^t} \text{diag}(\psi^{(3)}(-\Lambda))\Pi_t, \end{aligned} \tag{21}$$

where

$$
B_r^3 \equiv
\begin{pmatrix}
B_r^3\,(1) \\
B_r^3\,(2) \\
B_r^3\,(3) \\
B_r^3\,(4) \\
B_r^3\,(5) \\
B_r^3\,(6)
\end{pmatrix},
\operatorname{diag}\!\left(\psi^{(3)}(0)\right) \equiv
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & \psi_x^{(3)}(0) & 0 & 0 & 0 & 0 \\
0 & 0 & \psi_{\sigma^2}^{(3)}(0) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
$$

$$
\operatorname{diag}\!\left(\psi^{(3)}(-\Lambda)\right) \equiv
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & \psi_x^{(3)}(-\Lambda_x) & 0 & 0 & 0 & 0 \\
0 & 0 & \psi_{\sigma^2}^{(3)}(-\Lambda_{\sigma^2}) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
$$

In this paper, we define the skewness risk premium in equilibrium at time t, skp_t, as the following expression, which is the same manner with the definition of the variance risk premium (18):

$$
skp_t \equiv \mathbb{E}_t^{\mathbb{Q}}[\mathrm{Skew}_{t+1}^{\mathbb{Q}}(r_{m,t+2})] - \mathbb{E}_t^{\mathbb{P}}[\mathrm{Skew}_{t+1}^{\mathbb{P}}(r_{m,t+2})]. \tag{22}
$$

Substituting (21) into (22), the explicit representation for the skewness risk premium can be obtained as follows:

$$
skp_t \equiv B_r^{3^t}\operatorname{diag}(\psi^{(3)}(-\Lambda))\mathbb{E}_t^{\mathbb{Q}}[\Pi_{t+1}] - B_r^{3^t}\operatorname{diag}(\psi^{(3)}(0))\mathbb{E}_t^{\mathbb{P}}[\Pi_{t+1}]. \tag{23}
$$

We also find a number of interesting implications from the expressions of (21) and (23). First, in the case that there is no jump to fundamentals in the economy, that is, in the case of $\Pi_t \equiv 0$, it is clear that the conditional skewness of the market return should be zero due to (21). Thus, the existence of the nonzero skewness of the market return crucially depend on the existence of the jumps to fundamentals in the economy. Second, any temporal variation in endogenously generated skewness and skewness risk premium are solely due to the temporal variation in the jump intensity process λ_t. For example, if the jump intensity is constant, then it is clear that the skewness (under $\mathbb{P}$ and $\mathbb{Q}$) and skewness risk premium should be constant by (21) and (23). Third, since we have the fact of $A_\sigma < 0$ by the proposition 1, then in the case that the jump to the variance of consumption growth rate exists, that is, in the case that $\lambda_{\sigma^2,t} > 0$, we can show easily by (21) that the risk-neutral skewness at time t, $\mathrm{Skew}_t^{\mathbb{Q}}(r_{m,t+1})$, should be negative. Finally, we can also find via (23) that in the case that either $\lambda_{x,t} > 0$ or $\lambda_{\sigma^2,t} > 0$ is satisfied, the skewness risk premium at time t, skp_t, in equilibrium also should be negative due to the facts of $A_x > 0$ and $A_\sigma < 0$.

On the basis of the definition of (22), let us provide the proposition for the representation of the skewness risk premium in equilibrium.

Proposition 4 (The Skewness Risk Premium in Equilibrium) *In equilibrium, the skewness risk premium at time t, skp_t, is linear to the variance-of-variance, q_t, and the jump intensity, λ_t, and the representation of it is provided as follows:*

$$skp_t = \beta_{sp,c} + \beta_{sp,q}q_t + \beta_{sp,\lambda}\lambda_t,$$

where

$$
\begin{aligned}
\beta_{sp,c} &\equiv \left[l_x B_r^3(2)\psi_x{}^{(3)}(-\Lambda_x) + l_{\sigma^2} B_r^3(3)\psi_{\sigma^2}{}^{(3)}(-\Lambda_{\sigma^2})\right]\mu_\lambda, \\
\beta_{sp,q} &\equiv \left[l_x B_r^3(2)\psi_x{}^{(3)}(-\Lambda_x) + l_{\sigma^2} B_r^3(3)\psi_{\sigma^2}{}^{(3)}(-\Lambda_{\sigma^2})\right]\varphi_u(-\rho\varphi_\xi\Lambda_q - \varphi_u\Lambda_\lambda), \\
\beta_{sp,\lambda} &\equiv \left[l_x B_r^3(2)\psi_x{}^{(3)}(-\Lambda_x) + l_{\sigma^2} B_r^3(3)\psi_{\sigma^2}{}^{(3)}(-\Lambda_{\sigma^2})\right]\rho_\lambda.
\end{aligned}
\tag{24}
$$

Proof Considering (6), (16), and the definition of the moment-generating function, it is trivial to derive the above expression. □

From the above proposion, we find that any temporal variation in endogenously generated skewness risk premium is also solely due to the variance-of-variance q_t and the jump intensity λ_t as well as the volatility risk premium (Proposition 2). Moreover, provided that $\Lambda_x > 0$ and $\Lambda_{\sigma^2} < 0$, the factor loading to the jump intensity, that is, $\beta_{sp,\lambda}$, is guaranteed to be negative, but that to the variance-of-variance, that is, $\beta_{sp,q}$, can be both positive and negative in general. However, if the correlation between the dynamics of the variance-of-variance and that of the jump intensity, that is , ρ, is positive, then $\beta_{sp,q}$ is also guaranteed to be negative due to the facts that $\Lambda_q < 0$ and $\Lambda_\lambda < 0$.

Before we turn to the next discussion, it will be useful to mention about some features of the higher-order moments of the market return and the risk premiums in them.

First, as mentioned in the introduction in this paper, the usual assumption of an affine structure on the jump intensity process λ_t, that is, $\lambda_t = l_0 + l_1\sigma_t^2$ where $l_0, l_1 > 0$ and σ_t^2 is the variance of consumption growth rate, in the previous studies such as Drechsler and Yaron (2011) can not explain an empirical fact on a simultaneous relation between monthly stock returns and monthly changes of the option-implied skewness shown by (1). It is because, under such assumption, we can show analytically that the regression parameters to $\triangle ISkew_{t+1}$ in (1) should be positive even if empirical evidence is consistent with the negative parameters.[7] On the other hand, based on our model provided above, the correlation between the one-step-ahead market return, $r_{m,t+1}$, and the one-step-ahead change of risk-neutral skewness, $\triangle Skew_{t+1}^{\mathbb{Q}} \equiv Skew_{t+1}^{\mathbb{Q}} - Skew_t^{\mathbb{Q}}$, can be derived analytically with (14) and (21) as follows:

---

[7] The empirical evidence of the negative correlation between the one-step-ahead market return, $r_{m,t+1}$, and the one-step-ahead change of risk-neutral skewness, $\triangle Skew_{t+1}^{\mathbb{Q}} \equiv Skew_{t+1}^{\mathbb{Q}} - Skew_t^{\mathbb{Q}}$, also essentially shown by Neuberger (2012). He shows that the period when the S&P500 index volatility was very low by historic standards (from the year of 2003 to 2007) was also one of relatively low skewness, whereas skewness was actually rather high in the volatility spike of the year of 2008. Under the leverage effect of Black (1976), that evidence suggests the same essentially as that of the negative correlation between the one-step-ahead market return and the one-step-ahead change of risk-neutral skewness.

$$\text{Corr}\left(r_{m,t+1}, \triangle Skew_{t+1}^{\mathbb{Q}}\right) = K\varphi_u \kappa_{1,m}(\rho A_{q,m} + \varphi_u A_{\lambda,m})q_t,$$

$$\text{where} \quad K \equiv l_x B_r^3(2)\psi_x{}^{(3)}(-\Lambda_x) + l_{\sigma^2} B_r^3(3)\psi_{\sigma^2}{}^{(3)}(-\Lambda_{\sigma^2}).$$

From the above expression, we can show that, when $\rho < -\varphi_u \frac{A_{\lambda,m}}{A_{q,m}}$, the correlation between $r_{m,t+1}$ and $\triangle Skew_{t+1}^{\mathbb{Q}} \equiv Skew_{t+1}^{\mathbb{Q}} - Skew_t^{\mathbb{Q}}$ should be negative because, in the case of $\gamma > 1$ and $\psi > 1$, it is proven that K is negative. This observation is consistent with the empirical fact of (1) shown in the introduction of this paper. If we assume the "zero" correlation between the λ_t and q_t, we can not show the negative correlation between the one-step-ahead market return, $r_{m,t+1}$, and the one-step-ahead change of risk-neutral skewness, $\triangle Skew_{t+1}^{\mathbb{Q}} \equiv Skew_{t+1}^{\mathbb{Q}} - Skew_t^{\mathbb{Q}}$, with our model. Thus, we would like to emphasize that there is considerable validity in our model setting with the stochastic jump intensity compared with the previous studies such as Drechsler and Yaron (2011), etc.

Second, although both the variance risk premium and the skewness risk premium are linear to the variance-of-variance q_t and the jump intensity λ_t, we can find that they are linearly independent because of the fact that $det \equiv \beta_{vp,q}\beta_{sp,\lambda} - \beta_{vp,\lambda}\beta_{sp,q}$ will not be zero under suitable parameter condition, which will be also proven in Sect. 3 with a model calibration result. If we assume that jump intensity process λ_t has an affine structure as mentioned above, the correlation between the variance and skewness risk premium should be one in absolute value and should not be independent each other. This is because the jump intensity process is driven by only one factor, that is, σ_t^2, under such assumption. This observation also shows considerable validity in our model setting with the stochastic jump intensity.

2.4 An equity risk premium representation

In this subsection, let us show an equity risk premium representation with the variance and skewness risk premiums in equilibrium. In the beginning, we start with an expression for the equity risk premium provided by Drechsler and Yaron (2011) as follows:

$$\log \mathbb{E}_t(R_{m,t+1}) - r_{f,t} = B_r^t G_t G_t^t \Lambda + \Pi_t^t(\psi(B_r) - 1 - \psi(B_r - \Lambda) + \psi(-\Lambda)),$$

where

$$\psi(B_r) \equiv \begin{pmatrix} 0 \\ \psi_x(B_r(2)) \\ \psi_{\sigma^2}(B_r(3)) \\ 0 \\ 0 \\ 0 \end{pmatrix}, \psi(B_r - \Lambda) \equiv \begin{pmatrix} 0 \\ \psi_x(B_r(2) - \Lambda_x) \\ \psi_{\sigma^2}(B_r(3) - \Lambda_{\sigma^2}) \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\psi(-\Lambda) \equiv \begin{pmatrix} 0 \\ \psi_x(-\Lambda_x) \\ \psi_{\sigma^2}(-\Lambda_{\sigma^2}) \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

As mentioned in Drechsler and Yaron (2011), the first term, $B_r^t G_t G_t^t \Lambda$, represents the contributions of the Gaussian shocks to the equity risk premium. In particular, according to the expression of $G_t G_t^t = H_{\sigma^2} \sigma_t^2 + H_q q_t$ (see (19)), this term aggregates both the risk-return tradeoff relationship and a true premium for variance risk. The next terms, $\Pi_t^t(\psi(B_r) - 1 - \psi(B_r - \Lambda) + \psi(-\Lambda))$, represents the contributions from the jump processes. The derivation of this expression is presented in the "Appendix" in Drechsler and Yaron (2011).

The $r_{f,t}$ is the risk-free rate at time t in the economy and the explicit expression of this $r_{f,t}$ is provided in the "Appendix".

With the expression of $G_t G_t^t = H_{\sigma^2} \sigma_t^2 + H_q q_t$ and $\Pi_t \equiv \begin{pmatrix} 0 & \lambda_{x,t} & \lambda_{\sigma^2,t} & 0 & 0 & 0 \end{pmatrix}^t$, the following representation can be obtained via the expression for the equity risk premium shown above:

$$
\begin{aligned}
\log \mathbb{E}_t(R_{m,t+1}) - r_{f,t} &= \beta_{er,\sigma} \sigma_t^2 + \beta_{er,q} q_t + \beta_{er,\lambda} \lambda_t, \; where \; \beta_{er,\sigma} \equiv B_r^t H_\sigma \Lambda, \\
\beta_{er,q} &\equiv B_r^t H_q \Lambda, \\
\beta_{er,\lambda} &\equiv l_x \left[\psi_x(B_r(2)) - 1 - \psi_x(B_r(2) - \Lambda_x) + \psi_x(-\Lambda_x) \right] \\
&\quad + l_\sigma \left[\psi_{\sigma^2}(B_r(3)) - 1 - \psi_\sigma(B_r(3) - \Lambda_{\sigma^2}) + \psi_\sigma(-\Lambda_{\sigma^2}) \right].
\end{aligned}
\tag{25}
$$

As shown in (25), the equity risk premium is driven by the state variables of σ_t^2, q_t, and λ_t and have a time-varying nature essentially because each of those variables has the stochastic nature. In particular, in the case of $\gamma > 1$ and $\psi > 1$, it is proven that $\beta_{er,\sigma} > 0$, $\beta_{er,q} > 0$, and $\beta_{er,\lambda} > 0$ because of the facts that $\Lambda_x > 0$, $\Lambda_{\sigma^2} < 0$, $\Lambda_q < 0$, and $\Lambda_\lambda < 0$, which are provided in Proposition 1, so that if each of the state variables increases, then the equity risk premium also increases, and vice versa.

The conditional variance of the equity return at time t, $\mathrm{Var}_t^{\mathbb{P}}(r_{m,t+1})$, is also expressed by

$$
\begin{aligned}
\mathrm{Var}_t^{\mathbb{P}}(r_{m,t+1}) &= B_r^t G_t G_t^t B_r + B_r^{2^t} \mathrm{diag}\left(\psi^{(2)}(0) \right) \Pi_t \\
&= B_r^t H_{\sigma^2} B_r \sigma_t^2 + B_r^t H_q B_r q_t + (l_x B_r^2(2) \psi_x^{(2)}(0) + l_{\sigma^2} B_r^2(3) \psi_{\sigma^2}^{(2)}(0)) \lambda_t \\
&\equiv \beta_{var,\sigma} \sigma_t^2 + \beta_{var,q} q_t + \beta_{var,\lambda} \lambda_t,
\end{aligned}
$$

so that with (20), (24), (25), and the above expression for the conditional variance of the equity return we can derive an explicit equity risk premium representation of a linear factor pricing model with the variance and skewness risk premiums and the conditional variance of the equity return.

Proposition 5 (An Explicit Representation for the Equity Risk Premium)

$$
\log \mathbb{E}_t(R_{m,t+1}) - r_{f,t} = \pi_c + \pi_{var} \mathrm{Var}_t^{\mathbb{P}}(r_{m,t+1}) + \pi_{vp} vp_t + \pi_{sp} skp_t,
\tag{26}
$$
$$
where
$$
$$
\pi_c \equiv \left(-\frac{\beta_{er,\sigma} \beta_{var,q}}{\beta_{var,\sigma}} + \beta_{er,q} \right) \frac{-\beta_{sp,\lambda} \beta_{vp,c} + \beta_{vp,\lambda} \beta_{sp,c}}{det}
$$

$$+ \left(-\frac{\beta_{er,\sigma}\beta_{var,\lambda}}{\beta_{var,\sigma}} + \beta_{er,\lambda} \right) \frac{\beta_{sp,q}\beta_{vp,c} - \beta_{vp,q}\beta_{sp,c}}{det},$$

$$\pi_{var} \equiv \frac{\beta_{er,\sigma}}{\beta_{var,\sigma}},$$

$$\pi_{vp} \equiv \left(-\frac{\beta_{er,\sigma}\beta_{var,q}}{\beta_{var,\sigma}} + \beta_{er,q} \right) \frac{\beta_{sp,\lambda}}{det} - \left(-\frac{\beta_{er,\sigma}\beta_{var,\lambda}}{\beta_{var,\sigma}} + \beta_{er,\lambda} \right) \frac{\beta_{sp,q}}{det},$$

$$\pi_{sp} \equiv -\left(-\frac{\beta_{er,\sigma}\beta_{var,q}}{\beta_{var,\sigma}} + \beta_{er,q} \right) \frac{\beta_{vp,\lambda}}{det} + \left(-\frac{\beta_{er,\sigma}\beta_{var,\lambda}}{\beta_{var,\sigma}} + \beta_{er,\lambda} \right) \frac{\beta_{vp,q}}{det},$$

$$det \equiv \beta_{vp,q}\beta_{sp,\lambda} - \beta_{vp,\lambda}\beta_{sp,q}.$$

This representation of (26) suggests that the skewness risk premium, as well as the variance risk premium and the conditional variance of the market return, constitutes the dominant source of the variation in the equity risk premium. In the following section, we will show that *det* in (26) is not zero under the suitable parameter condition. Thus, from the above proposition, it is found that the skewness risk premium has an essential source of the variation in the equity risk premium, which is different from that of the variance risk premium (see Fig. 1).

Some recent studies such as those by Bali and Hovakimian (2009), Yan (2009), Chang et al. (2012), Driessen et al. (2012), and Rehman and Vilkov (2012) focus on a significant relationship between skewness or jump risks and expected stock returns, and they provide empirical evidence for a significantly positive link between the expected stock returns and the jump or skewness risks. To the best of our knowledge, this result of (26), which suggests an explicit relationship between the skewness risk premium and the expected equity excess return, is the first to provide a theoretical implication in their empirical evidence in terms of the LRR model approach pioneered by Bansal and Yaron (2004).
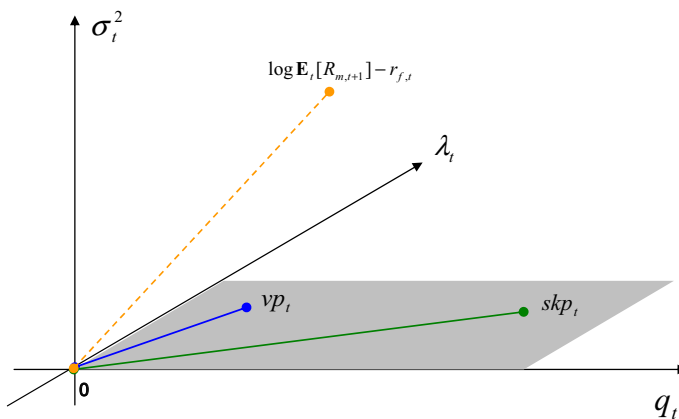

Fig. 1 The risk premiums in higher-order moments and the equity risk premium

Table 1 The set of model parameters

Parameter	Source	(Calibrated) Values
(1) Preference		
ψ	BST	2.5
(2) Consumption growth		
φ_η	BY	1.0
(3) Long run risk		
ρ_x	BY	0.979
φ_e	BY	0.044
(4) Variance		
ρ_σ	BTZ	0.978
(5) Variance-of-variance		
ρ_q	BTZ	0.8
φ_ξ	BTZ	0.001
(6) Campbell–Shiller		
κ_1	BTZ	0.9
(7) Jump intensity		
ρ_λ	CM	0.9
φ_u	–	0.01
$\delta_x, \delta_{\sigma^2}$	–	0.01
μ_λ	–	1.0
l_x, l_{σ^2}	–	1.0

This table reports the parameter values used in the calibration of the factor loadings in the theoretical model (26). CM, BY, BTZ, and BST in this table denote values taken directly from Chan and Maheu (2002), Bansal and Yaron (2004), Bollerslev et al. (2009), and Bollerslev et al. (2012), respectively

3 Model implications

Before proceeding to an empirical analysis based on the representation of (26), we show the implications from a calibrated version of the theoretical model (26) to help guide and interpret our subsequent empirical reduced form predictability regressions.

Table 1 reports the parameter values used in the calibration of the factor loadings in the theoretical model (26). CM, BY, BTZ, and BST in this table denote values taken directly from Chan and Maheu (2002), Bansal and Yaron (2004), Bollerslev et al. (2009), and Bollerslev et al. (2012), respectively. Those previous studies refer to the unit time interval in the calibrated equilibrium models as a month, and we also refer to the unit time as the same. On the basis of the parameters exhibited in Table 1, we calibrate the factor loadings for the variance risk premium, which appear in the representation of (20), and for the skewness risk premium, which appear in the representation of (24), in equilibrium.

The figures from Figs. 2, 3, 4 and 5 show the factor loadings, $\beta_{vp,q}$, $\beta_{vp,\lambda}$, $\beta_{sp,q}$, and $\beta_{sp,\lambda}$, corresponding to the parameters of the risk aversion parameters γ and the correlation ρ between the volatility of volatility q_t and the jump intensity λ_t. As is shown in the previous section, $\beta_{vp,\lambda}$, which is the factor loading to the jump intensity λ_t in the variance risk premium representation (20), is essentially positive, and under the parameter values exhibited in Table 1, $\beta_{vp,q}$, which is the factor loading to the
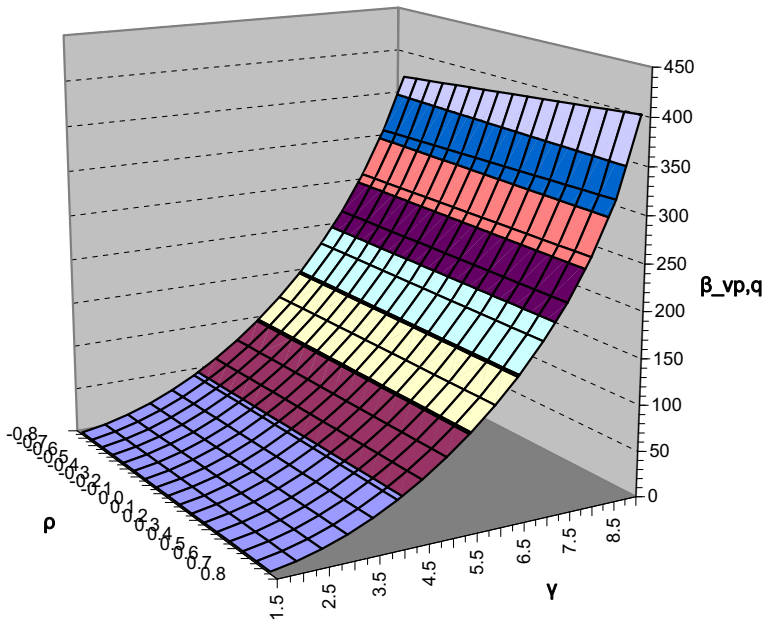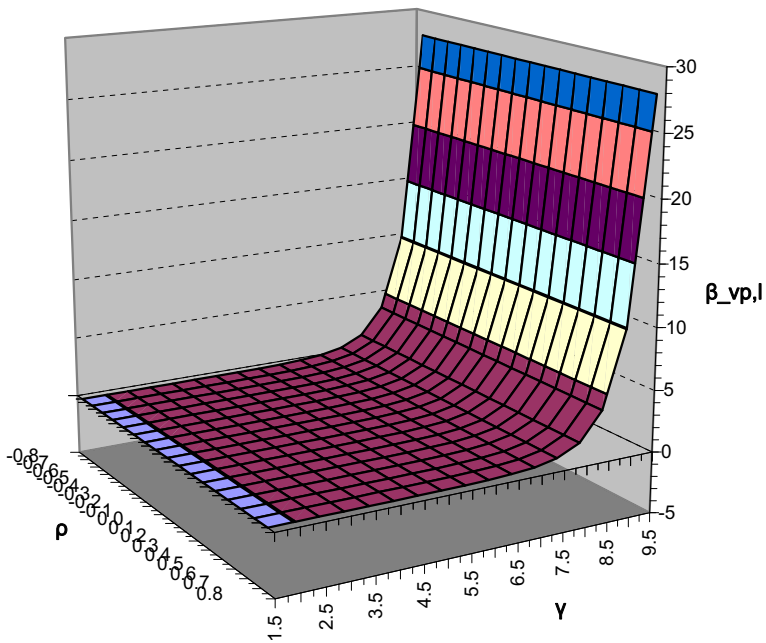
Fig. 2 The factor loading $\beta_{vp,q}$

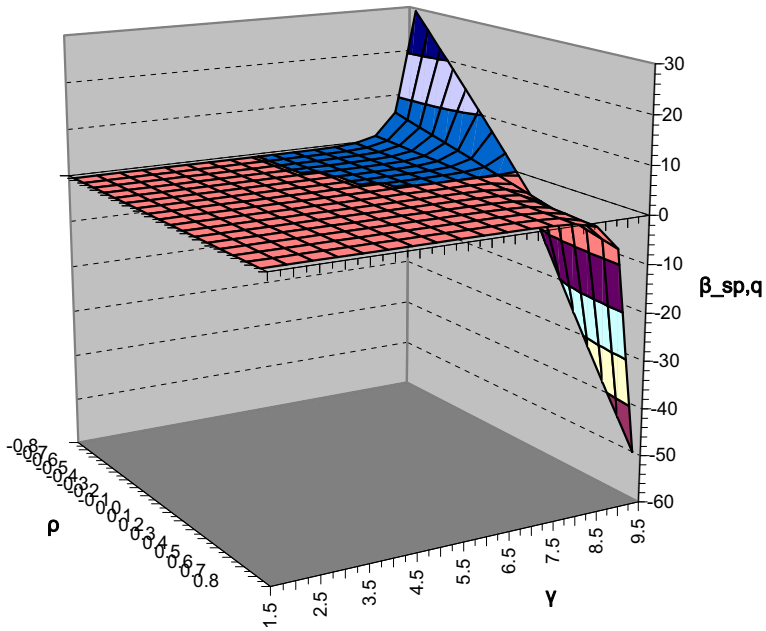**Fig. 3** The factor loading $\beta_{vp,\lambda}$

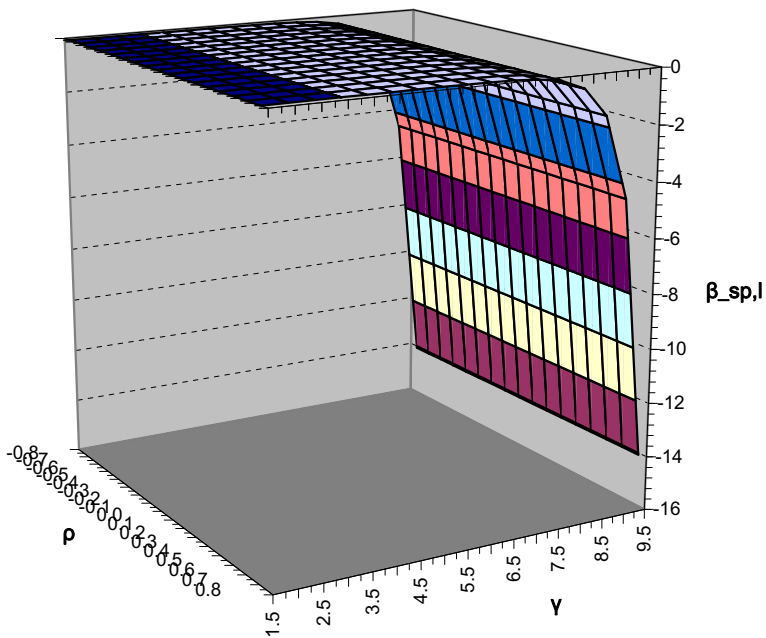**Fig. 4** The factor loading $\beta_{sp,q}$

Fig. 5 The factor loading $\beta_{sp,\lambda}$

variance-of-variance q_t in (20), also seems to be positive. These results indicate that when the variance-of-variance and (or) the jump intensity rise(s), the level of the variance risk premium also increases. In contrast, $\beta_{sp,\lambda}$, which is the factor loading to the jump intensity in the skewness risk premium representation (24), is essentially negative and this result is consistent with the discussion explored in the previous section. However, interestingly, $\beta_{sp,q}$, which is the factor loading to the variance-of-variance in (24), can be both positive and negative corresponding to the parameters of γ and ρ. These results on the $\beta_{sp,\lambda}$ and the $\beta_{sp,q}$ indicate that although an increase in the jump intensity reduces the level of the skewness risk premium essentially, but an increase in the variance-of-variance will raise or reduce the level of the skewness risk premium corresponding to the values of γ and ρ.

The figures from Figs. 6, 7 and 8 show the factor loadings to the variance of the market return, the variance risk premium, and the skewness risk premium in the equity risk premium representation (26). It is interesting that both of the π_{Var} and π_{vp} are essentially positive and these results are irrelevant to the values of γ and ρ. Moreover, these results are consistent with the previous studies such as those by Bollerslev et al. (2009) and Drechsler and Yaron (2011). An important point to emphasize is that the factor loading of π_{skp}, which is the loading to the skewness risk premium in (26), can be positive corresponding to the risk aversion parameter γ. In particular, when the γ is over 4, it is clear from Fig. 8 that the π_{skp} is strictly positive. This result indicates that a decrease in the skewness risk premium, which is the case that the risk-neutral skewness is going to be much smaller than the skewness under the physical measure, reduces
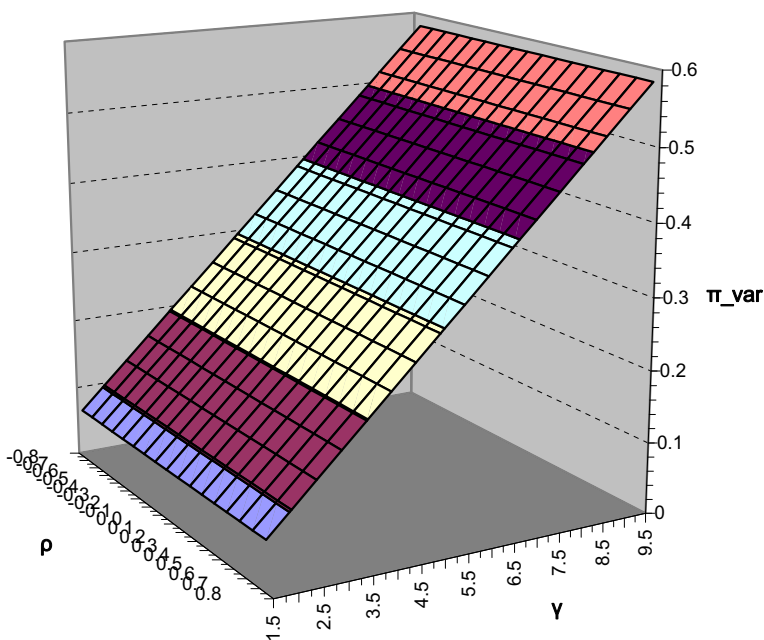

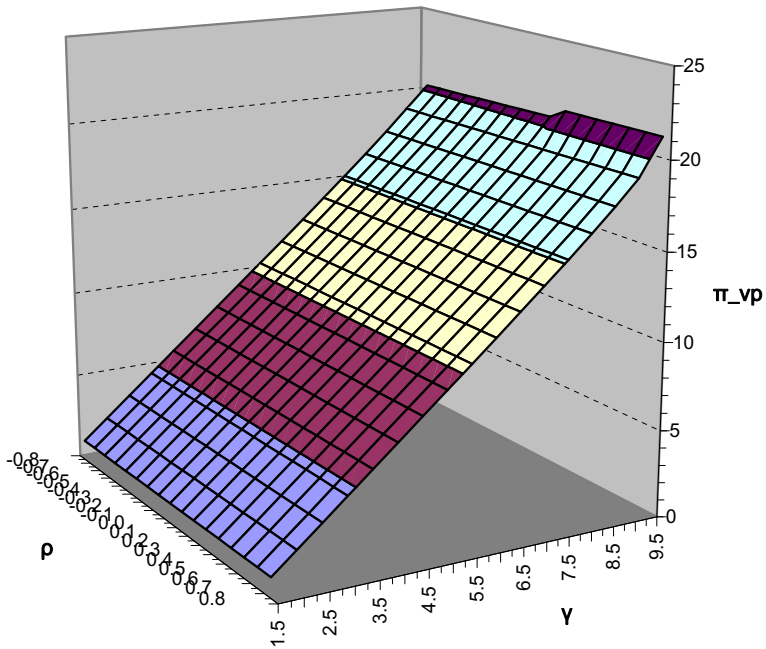**Fig. 6** The factor loading to $Var_t^{\mathbb{P}}$: π_{Var}
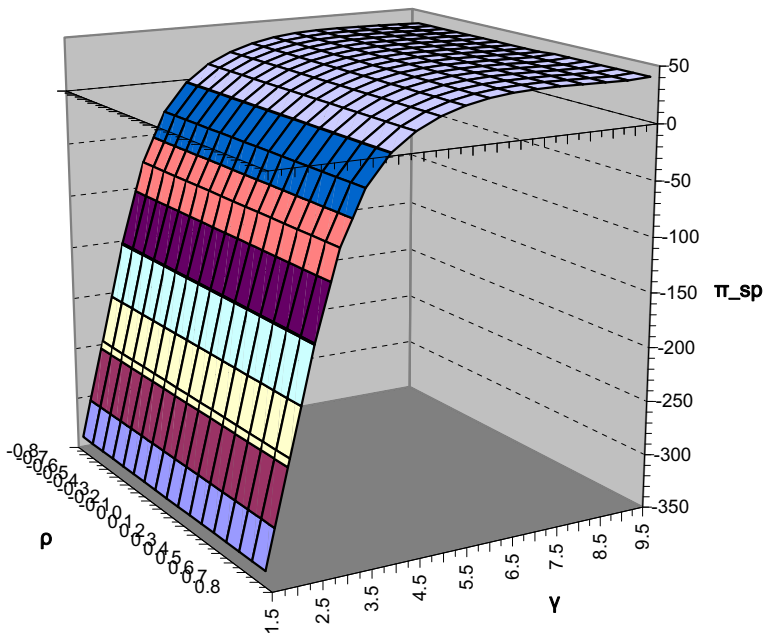
Fig. 7 The factor loading to vp_t : π_{vp}

Fig. 8 The factor loading to skp_t : π_{skp}

the equity risk premium when γ is over 4. This implication is interesting as it shows the essential contribution of the skewness risk premium to the equity risk premium explicitly implying the sign of the π_{skp} corresponding to the values of γ and ρ. As mentioned above, some recent studies such as those by Bali and Hovakimian (2009), Yan (2011), Chang et al. (2012), Driessen et al. (2012), and Rehman and Vilkov (2012) focus on a significant relationship between skewness or jump risks and expected stock returns, and they provide empirical evidence for a significantly positive link between the expected stock returns and the jump or skewness risks. In particular, Bali and Hovakimian (2009) and Yan (2011) provide evidence for a significantly positive link between expected returns and the call-put options' implied volatility spread that can be considered as a proxy for jump risk. Moreover, using data on individual stock options, Rehman and Vilkov (2012) show that the currently observed option-implied ex ante skewness is positively related to future stock returns. There has been no study that tried to provide the theoretical equilibrium model which is consistent with the empirical results cited above. To the best of our knowledge, this is the first paper that demonstrates what mentioned above with a stylized model that accounts for a close relationship between the skewness risk premium and the equity risk premium.

4 Empirical measurements

The theoretical model developed in the previous section suggests that the variance and skewness risk premiums, as well as the variance of the market return, may serve as useful predictor variables for the future market returns (see Proposition 4). To examine this suggestion empirically and compare the results with those by Bollerslev et al. (2009) and Drechsler and Yaron (2011), we plan for running some statistical tests based on simple linear regression models of the S&P500 excess return on different sets of lagged predictor variables including the variance and skewness risk premiums. We always rely on monthly and quarterly observations and focus our discussion on the estimated slope coefficients and their statistical significance as determined by the t-statistics. We also report the forecasts' accuracy of the regression models as measured by the corresponding adjusted $R^2 s$.

Before showing the results of the predictive regression models of the S&P500 excess return, let us note some key points on the measurements for the variance and skewness risk premiums and describe the data used in our analysis explored in the following subsection.

4.1 Measurements for the higher-order moments

Our method for measuring the risk premiums in higher-order moments is similar to that in Bollerslev et al. (2009) and Drechsler and Yaron (2011). As mentioned above, we formally define the variance risk premium as the difference between the risk-neutral and physical expectations of the variance of the market return and also define the skewness risk premium in the same manner. We focus on the one-month- and three-month-forward predictability of those risk premiums and use the squared VIX and the SKEW index from the Chicago Board of Options Exchange (CBOE) as

our measures for the risk-neutral expected variance and skewness, respectively. The VIX is calculated by the CBOE using the model-free approach to measure 30-day expected volatility of the S&P500 return. The components of the VIX are near- and next-term put and call options, usually in the first and second SPX (S&P500 index) contract months. The model-free approach used to calculate the VIX is provided by, for example, Demeterfi et al. (1999). The SKEW index from the CBOE is also calculated from the S&P 500 option prices based on the method similar to that used to calculate the VIX, which is obtained by a portfolio of S&P 500 index options that mimics an exposure to the skewness payoff of one-step-ahead cumulative return distribution of the index. The Skew index is also calculated by the model-free approach provided, for example, Bakshi et al. (2003).[8]

For the measures of the expected variance and skewness under the physical measure, we basically use the current variance and skewness of the S&P500 index return, which are respectively defined as the historical 22 days actual variance estimated based on daily return data of the index and the historical 12 months actual skewness estimated based on monthly return data of the index. To match the definition of those historical moments of the index return distribution with the risk-neutral expected moments mentioned above, we use the annualized current variance, while the current skewness, which is estimated based on historical 12 months monthly return data, is used as it is. Bollerslev et al. (2009) suggest that, for highly persistent variance dynamics, or $\rho_\sigma \approx 1$, the objective expected future variance will obviously be close to the value of the current variance so that the same qualitative implications hold true for the variance difference obtained by replacing $\mathbb{E}_t^{\mathbb{P}}[Var_{t+1}^{\mathbb{P}}(r_{m,t+2})]$ in the definition of (18) with the current variance. From a similar point of view, the same would be considered for the objective expected future skewness.[9] However, in our model setting, the time-t skewness of the one-step-ahead market return, $r_{m,t+1}$, can be expressed by the jump intensity λ_t (see (21)), and empirically some people may find non-persistent jump intensities and it may be true that the jumps are rare events and their roles should disappear very quickly. For this reason, in addition to the current skewness as a proxy of the measure of expected skewness under the physical measure, it may be very nice to introduce another approach for measuring the $\mathbb{E}_t^{\mathbb{P}}[Skew_{t+1}^{\mathbb{P}}(r_{m,t+2})]$

---

[8] According to the description of the CBOE's SKEW index, we have the proxy for the risk-neutral expected skewness, $\mathbb{E}_t^{\mathbb{Q}}[Skew_{t+1}^{\mathbb{Q}}(r_{m,t+2})]$, as $\frac{1}{10}(100 - Skew\ \ index)$.

[9] In our model setting, the time-t skewness of the one-step-ahead market return, $r_{m,t+1}$, can be expressed by (21). Thus, we can find that the first-order autocorrelation of the skewness under the physical measure $\mathbb{P}$ (or the risk-neutral measure $\mathbb{Q}$) is determined by that of the jump intensity λ_t. From (6), we can also find that the first-order autocorrelation of the jump intensity λ_t is ρ_λ. Previous studies such as those by Chan and Maheu (2002) and Christoffersen et al. (2012) also model the dynamics of the jump intensity of financial asset returns as a first-order autoregressive model in a similar fashion, and they find that the ρ_λ parameters in the jump intensity models, which are estimated with historical data of stock indices, are more than 0.9 and show the strong persistence in the conditional jump intensity. So, on the basis of these previous studies, we could assume that the objective expected future skewness, $\mathbb{E}_t^{\mathbb{P}}[Skew_{t+1}^{\mathbb{P}}(r_{m,t+2})]$, will be close to the value of the current empirical skewness so that the same qualitative implications hold true for the skewness difference obtained by replacing $\mathbb{E}_t^{\mathbb{P}}[Skew_{t+1}^{\mathbb{P}}(r_{m,t+2})]$ in the definition (22) with the current empirical skewness.

based on a forecasting regression.[10] In the empirical analysis in the following, we also investigate empirical results explored by using the expected physical skewness estimated by forecasting regressions described in the below. The models we employ in this case are the following:

$$
\begin{aligned}
Reg.(1) &: Skew_{t+1}^M = \alpha + \beta_1 Skew_t^M + \beta_2 CVol_t + \epsilon_{t+1}, \\
Reg.(2) &: Skew_{t+1}^D = \alpha + \beta_1 Skew_t^D + \beta_2 CVol_t + \epsilon_{t+1}, \\
Reg.(3) &: Skew_{t+1}^M = \alpha + \beta_1 Skew_t^M + \beta_2 \triangle CVol_t + \epsilon_{t+1}, \\
Reg.(4) &: Skew_{t+1}^D = \alpha + \beta_1 Skew_t^D + \beta_2 \triangle CVol_t + \epsilon_{t+1}, \\
Reg.(5) &: Skew_{t+1}^M = \alpha + \beta_1 Skew_t^M + \epsilon_{t+1}, \\
Reg.(6) &: Skew_{t+1}^D = \alpha + \beta_1 Skew_t^D + \epsilon_{t+1}.
\end{aligned}
\tag{27}
$$

All the models presented above are monthly-based model and one time step of the above models is one month. $Skew_t^M$ is time-t actual skewness, which is culculated based on historical 12 months monthly return data of the S&P500 index. $Skew_t^D$ is time-t actual skewness culculated based on historical 22 days daily return data of the S&P500 index,[11] and $CVol_t$ is time-t annualized volatility culculated based on historical 22 days daily return data of the S&P500 index. $\triangle CVol_t$ is one-month differential of the current volatility, which is defined by $\triangle CVol_t \equiv CVol_t - CVol_{t-1}$. The regression models introduced above are estimated based on historical 36 months rolling window procedure. Based on the one-step-ahead forecasts of the physical skewness of the S&P500 index monthly return distribution, which are estimated by each of the models introduced above, we calculate the skewness risk premium based on the definition of (22).

4.2 Data description

Our data series for the VIX, SKEW index, and expected variance and skewness under $\mathbb{P}$ covers the period from January 1990 to August 2012. The main limitation on the length of our sample comes from the VIX and SKEW index, since the time series published by the CBOE begins in January 1990. As mentioned in the previous subsection, we rely on the monthly and quarterly data for the squared VIX and SKEW index for quantifying $\mathbb{E}_t^{\mathbb{Q}}[Var_{t+1}^{\mathbb{Q}}(r_{m,t+2})]$ in (18) and $\mathbb{E}_t^{\mathbb{Q}}[Skew_{t+1}^{\mathbb{Q}}(r_{m,t+2})]$ in (22), respectively, and purposely rely on the readily available squared VIX as our measure for the risk-neutral expected variance and the value of $\frac{1}{10}(100 - Skew\ index)$ as our measure for the risk-neutral expected skewness. The expected variance $\mathbb{E}_t^{\mathbb{P}}[Var_{t+1}^{\mathbb{P}}(r_{m,t+2})]$ and the expected skewness $\mathbb{E}_t^{\mathbb{P}}[Skew_{t+1}^{\mathbb{P}}(r_{m,t+2})]$ at time t are respectively calculated based on the historical index returns as described in the previous subsection.

---

[10] We thank a referee so much for suggesting the idea on this point.

[11] In this case, under the assumption of i.i.d distribution for daily returns, $Skew_t^D$ is the historical 22 days actual skewness of daily return data, which is devided by $\sqrt{22}$.
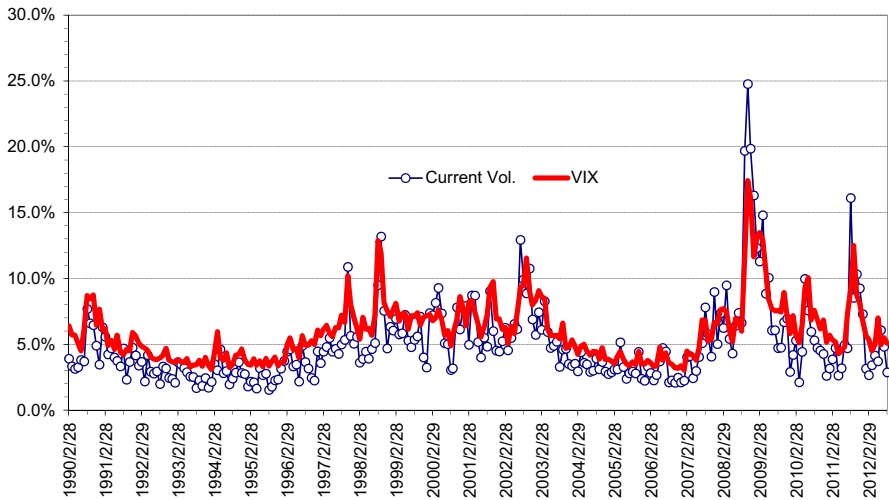
Fig. 9 The VIX and the current volatility. This figure shows the time-series data of the VIX and the current volatility (the square root of the current variance defined in the main paper). The current volatility is the historical 22 days actual volatility estimated based on daily return data of the S&P500 index
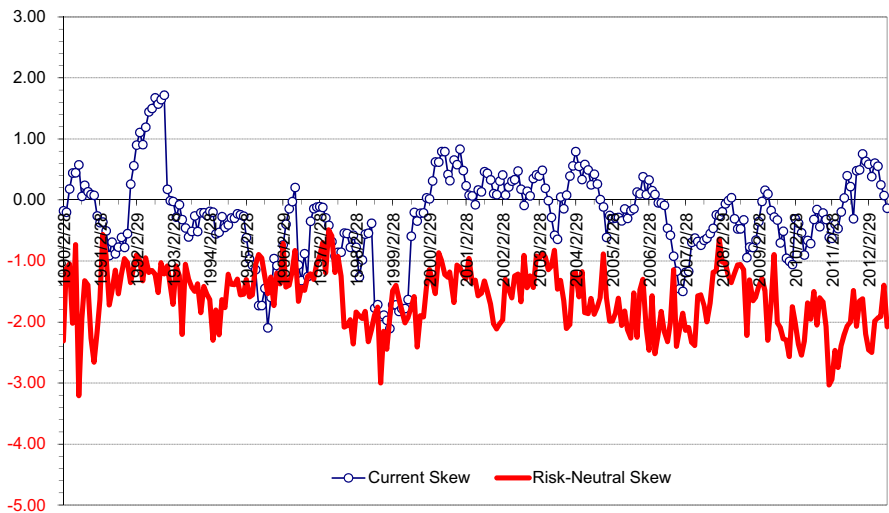
Fig. 10 The risk-neutral skewness and the current skewness. This figure shows the time-series data of the risk-neutral expected skewness extracted from the SKEW index and the current skewness. The current skewness is the historical 12 months actual skewness estimated based on monthly return data of the S&P500 index

To illustrate the data, Figs. 9 and 10 plot the monthly time-series of the risk-neutral expected volatility (VIX), the current volatility (historical 22 days annualized actual volatility), the risk-neutral expected skewness, and the current skewness (historical 12 months actual skewness). Consistent with the theoretical model developed in the previous section and the earlier empirical evidence, the spread between the risk-neutral

expected variance (the squared VIX) and the current variance is almost always positive and the spread between the risk-neutral expected skewness and the current skewness is almost always negative. Moreover, it is clear that these spreads have a time-varying nature. It is interesting that, although the value of the VIX reaches an outstanding peak at the period of the subprime crisis in 2008, the risk-neutral skewness seems to be more negative at the period of the European financial crisis in 2011 than at the period of the subprime crisis.

In addition to the variance and skewness risk premiums, we also consider a set of other more traditional predictor variables for the predictive regressions examined in the following subsection. Specifically, we obtain monthly P/E ratios and dividend yields of the S&P 500 index directly from Standard&Poor's. Data on the three-month T-bill, the high-yield spread (hys) (between Moody's BAA and AAA corporate bond yields), and the term spread (ts) (between the 1-year T-bond and the three-month T-bill yields) are taken from the Thomson Reuters Data Stream. The CAY, which represents the aggregate consumption-wealth ratio, defined in Lettau and Ludvigson (2001) is downloaded from Lettau and Ludvigson's Web site.

Basic summary statistics for the monthly excess returns of the S&P500 index and predictor variables are exhibited in Table 2. The sample period extends from January 1990 to August 2012. All variables are reported in monthly-based percentage form whenever appropriate. The $r_{m,t} - r_{f,t}$ denotes the logarithmic return on the S&P 500 index in excess of the three-month T-bill rate. VIX^2 denotes the squared VIX index. ISkew refers to the risk-neutral expected skewness extracted from the CBOE SKEW index by the formula of $ISkew = \frac{1}{10}(100 - Skew\ index)$. CVar and CSkew refer to the current variance, which is the annualized actual variance based on historical 22 days daily return data, and the current skewness, which is the actual skewness based on historical 12 months monthly return data, respectively. vp and skp respectively refer to the variance and skewness risk premiums, that is, $vp \equiv VIX^2 - CVar$ and $skp \equiv ISkew - CSkew$. The predictor variables include the log price-earning ratio (ln(pe)), the log dividend yield (ln(dy)), the high yield spread (hys) defined as the difference between Moody's BAA and AAA bond yield indices, and the term spread (ts) defined as the difference between the 10-year and 3-month Treasury yields.

The mean excess return on the S&P 500 index over the sample equals 0.3 % monthly. The sample means for the VIX^2 and the current (historical 22 days) annualized variance are 6.0 and 5.0 %, respectively, and the sample means for the risk-neutral expected skewness and the current (historical 12 months) skewness are -1.6 and -0.2, respectively. The numbers for the traditional forecasting variables are all directly in line with those reported in previous studies. In particular, all of the variables are highly persistent with first-order autocorrelations ranging from 0.95 to 0.99.

As stated in the previous subsection, we additionally construct expected physical skewness for measuring the $\mathbb{E}_t^{\mathbb{P}}[Skew_{t+1}^{\mathbb{P}}(r_{m,t+2})]$ based on the forecasting regressions (see (27)). Table 3 shows basic summary statistics for the expected physical skewness ("ES"), which are estimated by the forecasting regressions described in (27), and correspondent skewness risk premium (skp). The sample period extends from January 1990 to August 2012. For example, the label "ES_1" means the expected physical skewness estimated by the model of "Reg. (1)" in (27) and the label "skp_1" means correspondent skewness risk premium. From this table, we can find that the expected physical skew-

Table 2 Summary statistics for the monthly returns and predictor variables

	$r_{m,t} - r_{f,t}$	VIX^2	ISkew	CVar	CSkew	vp	skp	ln(pe)	ln(dy)	hys	ts
(A) Summary statistics											
(1) Mean	0.3%	6.0%	−1.6	5.0%	−0.2	1.0%	−1.4	3.1	0.7	1.0%	1.9%
(2) Std. dev.	4.4%	2.3%	0.5	3.1%	0.7	1.5%	0.8	0.4	0.3	0.4%	1.2%
(3) Skewness	−0.6	1.6	−0.4	2.7	−0.2	−2.2	0.2	2.3	0.1	3.2	−0.2
(4) Kurtosis	1.1	3.9	−0.2	10.7	0.7	9.9	0.0	7.1	−0.6	12.9	−1.1
(5) AR(1)	0.07	0.85	0.56	0.75	0.91	0.22	0.74	0.95	0.99	0.96	0.97
(B) Correlation matrix											
$r_{m,t} - r_{f,t}$	1										
VIX^2	0.02	1									
ISkew	0.06	−0.00	1								
CVar	−0.11	0.87	0.03	1							
CSkew	−0.10	−0.06	0.19	−0.03	1						
vp	0.25	−0.25	−0.07	−0.69	−0.03	1					
skp	0.13	0.05	0.48	0.04	−0.77	−0.01	1				
ln(pe)	0.01	0.25	0.00	0.15	0.00	0.08	0.00	1			
ln(dy)	0.08	−0.14	0.11	−0.10	0.09	−0.02	−0.01	−0.35	1		
hys	−0.05	0.64	−0.04	0.67	0.07	−0.38	−0.09	0.16	0.21	1	
ts	−0.02	0.06	0.08	0.05	0.33	−0.02	−0.24	0.28	0.38	0.26	1

The sample period extends from January 1990 to August 2012. All variables are reported in monthly-based percentage form whenever appropriate. The $r_{m,t} - r_{f,t}$ denotes the logarithmic return on the S&P 500 in excess of the three-month T-bill rate. VIX^2 denotes the squared VIX index. ISkew refers to the risk-neutral expected skewness extracted from the CBOE SKEW index by the formula of $ISkew = \frac{1}{10}(100 - Skew \text{ index})$. CVar and CSkew refer to the current variance, which is the annualized actual variance based on historical 22 days daily return data, and the current skewness, which is the actual skewness based on historical 12 months monthly return data, respectively. vp and skp respectively refer to the variance and skewness risk premiums, that is, $vp \equiv VIX^2 - CVar$ and $skp \equiv ISkew - CSkew$. The predictor variables include the log price-earning ratio (ln(pe)), the log dividend yield (ln(dy)), the high yield spread (hys) defined as the difference between Moody's BAA and AAA bond yield indices, and the term spread (ts) defined as the difference between the ten-year and three-month Treasury yields

Table 3 Summary statistics for the expected physical skewness and skewness risk premium

	ES_1	ES_2	ES_3	ES_4	ES_5	ES_6	skp_1	skp_2	skp_3	skp_4	skp_5	skp_6
(A) Summary statistics												
(1) Mean	−0.22	−0.01	−0.24	−0.01	−0.24	−0.01	−1.39	−1.60	−1.38	−1.60	−1.37	−1.60
(2) SD	0.64	0.05	0.64	0.04	0.64	0.03	0.73	0.50	0.74	0.50	0.74	0.50
(3) Skewness	−0.04	1.22	−0.00	−0.17	−0.01	−0.03	0.13	−0.37	0.14	−0.41	0.14	−0.39
(4) Kurtosis	0.90	9.00	0.91	1.27	0.92	1.21	0.06	−0.15	0.05	−0.16	0.04	−0.17
(5) AR(1)	0.91	0.59	0.91	0.22	0.92	0.58	0.72	0.55	0.73	0.54	0.73	0.55
(B) Correlation matrix												
ES_1	1											
ES_2	0.06	1										
ES_3	0.99	0.08	1									
ES_4	0.22	0.58	0.21	1								
ES_5	0.99	0.08	1.00	0.22	1							
ES_6	0.24	0.59	0.25	0.75	0.25	1						
skp_1	−0.74	0.02	−0.74	−0.11	−0.74	−0.08	1					
skp_2	0.19	0.01	0.18	0.07	0.18	0.14	0.52	1				
skp_3	−0.72	0.00	−0.74	−0.09	−0.74	−0.09	0.99	0.53	1			
skp_4	0.18	0.06	0.17	0.05	0.17	0.13	0.53	1.00	0.54	1		
skp_5	−0.73	−0.01	−0.74	−0.11	−0.74	−0.09	0.99	0.53	1.00	0.54	1	
skp_6	0.18	0.07	0.17	0.08	0.17	0.14	0.53	1.00	0.54	1.00	0.53	1

The sample period extends from January 1990 to August 2012. For example, the label "ES_1" means the expected physical skewness estimated by the model of "Reg. (1)" in (27) and the label "skp_1" means correspondent skewness risk premium

ness and correspondent skewness risk premium have similar characteristics among monthly- or daily-based skewness regression models.

4.3 The results of simple regression tests

Table 4 provides the results of return predictability regressions with the variance and skewness risk premiums. In the first place, we use the skewness risk premium constructed with the current skewness in this table. (The regression results based on the expected physical skewness estimated by the regression models of (27) and correspondent skewness risk premium will be shown in the next table.) All of our forecasts are based on simple linear regressions of the S&P500 excess returns on different sets of lagged predictor variables. There are two sets of columns with regression estimates. The first set of columns shows OLS estimates by monthly return regressions, that is, one-month-ahead forecasts, and the second set shows OLS estimates by non-overlapped quarterly return regressions, that is, one-quarter-ahead forecasts. These regressions are examined in the period from January 1990 to August 2012 and, in particular, each of the monthly return regressions is examined by 270-month samples and each of the quarterly return regressions is examined by 88-quarter samples. Each of the sets of columns consists of five regression results. The first two regressions are one-factor regression models using the variance risk premium (vp model) or the skewness risk premium (skp model) as a univariate regressor, while the third regression is two factors regression model with both the variance and skewness risk premiums (vp+skp model). The fourth regression model, which is denoted by 3 factors model, is based on the three independent factors of the current variance (CVar), the variance premium (vp), and the skewness risk premium (skp), and it represents the theoretical linear model of (26) derived in the previous section. Finally, we also provide the stepwise selection model (Stepwise model), which the universe of independent variables consists of the risk premiums in higher-order moments, changes of those risk premiums, and one of the traditional predictor variables, that is, the log price-earning ratio (ln(pe)). The variables such as $\triangle VIX^2$ and $\triangle ISkew$ exhibited in this table are monthly or quarterly changes of the VIX^2 and $ISkew$, respectively.

From the monthly return regression results in this table, we can find that the slope coefficients of the vp and skp model are both significant at 5 % level and, in particular, the slope coefficient of the vp model is significant at 1 % level. Moreover, the slope coefficients of the vp+skp model are also significant at the same level as mentioned above and this model can account for about 7.2 % of the monthly return variation. The 3 factors model represents the theoretical implication of (26) and this model has a superior predictive power in the adjusted R^2 than the vp+skp model due to the additional variable of the current variance (CVar). Although the stepwise model is not equivalent to the theoretical implication of (26), that is, the 3 factors model, all of the independent variables of CVar, vp, and skp in (26) are significant at 5 or 1 % level. These results indicate that the theoretical model of (26) and, in particular, the variance and skewness risk premiums have superior predictive power for future aggregate stock market index returns and this indication is consistent with the theory provided in the previous section in this paper.

Table 4 The monthly and quarterly return regressions

(A) Monthly return regression models

	vp	skp	vp+skp	3 factors	Stepwise
Constant	−0.004	0.013	0.006	−0.004	−0.015
(t-stat)	−1.18	2.43**	1.18	−0.42	−1.57
VIX^2					
ISkew					
CVar				0.158	0.312
				1.37	2.54**
CSkew					
ΔVIX^2					−0.883
					−3.35***
Δ ISkew					
Δ CVar					0.369
					2.11**
Δ CSkew					
vp	0.699		0.704	0.921	1.416
	4.19***		4.25***	4.02***	4.85***
skp		0.007	0.007	0.007	0.008
		2.11**	2.22**	2.16**	2.42**
Δ vp					

(B) Quarterly return regression models

	vp	skp	vp+skp	3 facors	Stepwise
Constant	−0.007	0.027	0.016	−0.042	−0.042
(t-stat)	−0.68	1.48	0.91	−1.44	−1.44
VIX^2					
ISkew					
CVar				0.909	0.909
				2.43**	2.43**
CSkew					
ΔVIX^2					
Δ ISkew					
Δ CVar					
Δ CSkew					
vp	1.529		1.646	2.845	2.845
	2.95***		3.17***	4.03***	4.03***
skp		0.013	0.019	0.018	0.018
		1.11	1.60	1.56	1.56
Δ vp					

Table 4 continued

	(A) Monthly return regression models					(B) Quarterly return regression models				
	vp	skp	vp+skp	3 factors	Stepwise	vp	skp	vp+skp	3 facors	Stepwise
△ skp										
ln(pe)										
$Adj.R^2$	5.8%	1.3%	7.2%	7.5%	10.6%	8.1%	0.3%	9.8%	14.7%	14.7%

The sample period extends from January 1990 to August 2012. VIX^2 denotes the squared VIX index. ISkew refers to the risk-neutral expected skewness extracted from the CBOE SKEW index by the formula of $ISkew = \frac{1}{10}(100 - Skew\ index)$. CVar and CSkew refer to the current variance, which is the annualized actual variance based on historical 22 days daily return data, and the current skewness, which is the actual skewness based on historical 12 months monthly return data, respectively. vp and skp respectively refer to the variance and skewness risk premiums, that is, $vp \equiv VIX^2 - CVar$ and $skp \equiv ISkew - CSkew$. The variables such as $\triangle VIX^2$ and $\triangle ISkew$ exhibited in this table are monthly or quarterly changes of the VIX^2 and the ISkew, respectively. The predictor variables include the log price-earning ratio (ln(pe))

The quarterly regressions reported in this table further underscore the significance of the monthly return regressions and, in contrast to the monthly return regressions, all of the t-statistics for the skewness risk premium are insignificant at conventional levels. However, interestingly, we can find that the stepwise model is perfectly equivalent to the theoretical implication of (26), that is, the 3 factors model, and this model can account for about 14.7 % of the quarterly return variation. Although the slope coefficient to the skewness risk premium is not significant as mentioned above, the coefficients to the variance risk premium and the current variance are both significant at 5 % level and, in particular, at 1 % level for the variance risk premium.

Table 5 provides the results of predictive regressions with the 3 factors, that is, the physical variance $Var_t^{\mathbb{P}}(r_{m,t+1})$, the variance risk premium vp_t, and the skewness risk premium skp_t, of the S&P500 index. Particularly, the skewness risk premiums used in these regressions are constructed by the expected physical skewness estimated by the regression models described in (27). These regressions are examined in the period from Jan-1990 to Aug-2012 and this in-sample period is the same as the in-sample period of the regressions examined in Table 4. For example, the "Original 3 fac. model" means the 3 factors model which is shown in the Table Table 5 and the "Predictive Reg. (1)" means the predictive regression model based on the 3 factors in which the skewness risk premium is constructed with the expected physical skewness estimated by the Reg. (1) model in (27). We can find that the results in Table 5 show that none of the alternative 3 factors models using the one-step-ahead forecast of the physical skewness have a superior predictive power in comparison with the original 3 factors model in terms of the adjusted R-square. Moreover, the statistical significance of the factor loadings on the skewness risk premium skp_t are decreasing when compared with the original model and, in particular, the factor loadings on the skp_t of the Predictive Reg. (2), Predictive Reg. (4), and Predictive Reg. (6) are statistically insignificant at 5 % level. However, decreasing R-squares and smaller t-ratios should be the case when we use the expected quantities for the physical skewness and this is not surprising. In particular, the results of Predictive Reg. (1), (3), and (5) nevertheless indicate that the variance and skewness risk premiums have superior predictive power for future aggregate stock market index returns.

Let us show the other results to emphasize the superiority of the skewness risk premium, as well as the variance risk premium, as a predictor variable for the equity excess return. Table 6 reports monthly- and quarterly-based predictive regression results for the S&P500 index excess return with each of the traditional predictor variables exhibited in this table, that is, the price-earning ratio (pe), dividend yield (dy), high-yield spread (hys), and term spread (ts) defined in the previous subsection and the changes of those variables. As shown in this table, we can find that, in the case of the monthly return regressions, none of the predictor variables are superior in the adjusted R^2 to the variance and skewness risk premiums (see Table 4). In the case of the quarterly return regressions in this table, it seems that only $\triangle \ln(\text{pe})$ and $\triangle \text{hys}$ have superior adjusted R^2 in comparison with the skewness risk premium, but, none of the variables are superior in the adjusted R^2 to the variance risk premium (see Table 4)

Table 8 reports monthly- and quarterly-based predictive regression results for the S&P500 index excess return with the CAY, the aggregate-consumption wealth ratio

Table 5 The results of monthly-return-based predictive regressions

Models	Original 3 fac. model	Predictive Reg. (1)	Predictive Reg. (2)	Predictive Reg. (3)	Predictive Reg. (4)	Predictive Reg. (5)	Predictive Reg. (6)
Constant	−0.004	−0.003	−0.003	−0.004	−0.004	−0.004	−0.004
(t-stat)	(−0.42)	(−0.34)	(−0.26)	(−0.42)	(−0.31)	(−0.40)	(−0.33)
CVar	0.158	0.158	0.175	0.161	0.171	0.158	0.172
	(1.37)	(1.36)	(1.50)	(1.39)	(1.47)	(1.36)	(1.47)
vp	0.921	0.923	0.944	0.929	0.944	0.922	0.948
	(4.02***)	(4.03***)	(4.10***)	(4.05***)	(4.09***)	(4.02***)	(4.11***)
skp	0.007	0.008	0.007	0.007	0.007	0.007	0.007
	(2.16**)	(2.14**)	(1.43)	(2.06**)	(1.32)	(2.06**)	(1.30)
$Adj. R^2$	7.5%	7.4%	6.6%	7.3%	6.5%	7.3%	6.4%

The sample period extends from January 1990 to August 2012. CVar refer to the current variance, which is the annualized actual variance based on historical 22 days daily return data. vp and skp respectively refer to the variance and skewness risk premiums defined by the same as those in Table 4

Table 6 The univariate regressions with traditional predictor variables

	ln(pe)	△ ln(pe)	ln(dy)	△ ln(dy)	hys	△ hys	ts	△ ts
(A) Monthly return regressions								
Slope Coeff.	0.001	−0.002	0.011	−0.090	−0.516	-1.822	−0.084	−0.513
p-Value (%)	86.7	93.1	22.4	15.1	40.0	38.4	71.2	56.6
$Adj.R^2$ (%)	−0.4	−0.4	0.2	0.4	−0.1	-0.1	−0.3	−0.3
(B) Quarterly return regressions								
Slope Coeff.	0.008	0.059	0.038	−0.093	−1.118	-7.402	−0.132	−1.661
p-Value (%)	73.8	10.6	20.6	41.3	58.8	2.5**	86.4	31.6
$Adj.R^2$	−1.0	1.9	0.7	−0.4	−0.8	4.6	−1.1	0.0

The sample period extends from January 1990 to August 2012. We obtain monthly P/E ratios (pe) and dividend yields (dy) for the S&P 500 directly from Standard&Poor's. Data on the three-month T-bill, the high-yield spread (hys) (between Moody's BAA and AAA corporate bond yields), and the term spread (ts) (between the ten-year T-bond and the three-month T-bill yields) are taken from the Thomson Reuters Data Stream

Table 7 Summary statistics for the CAY

	$r_{m,t} - r_{f,t}$	cay	vp	skp
(A) Summary statistics				
(1) Mean	0.29 %	0.21 %	0.94 %	-132.51 %
(2) SD	4.38 %	2.38 %	1.55 %	75.83 %
(3) Skewness	-0.59	-0.11	-2.27	0.18
(4) Kurtosis	1.12	-1.40	10.02	0.08
(5) AR(1)	0.08	0.98	0.23	0.75
(B) Correlation matrix				
$r_{m,t} - r_{f,t}$	1			
cay	0.09	1		
vp	0.24	0.17	1	
skp	0.14	0.26	0.01	1

The sample period extends from January 1990 to January 2012. The CAY is the aggregate-consumption wealth ratio defined in Lettau and Ludvigson (2001), which is quarterly-based data and downloaded from Lettau and Ludvigson's web site

defined in Lettau and Ludvigson (2001). The CAY is quarterly-based data and downloaded from Lettau and Ludvigson's web site. The downloaded data covers January 1990 to January 2012. Table 7 shows summary statistics for the CAY as well as the variance and skewness risk premiums under the period from January 1990 to January 2012. For the monthly return regressions, we define a monthly CAY series from the most recent quarterly observation.

As shown in Table 8, we can find that the CAY does not seem to be superior predictor variable in comparison with the variance and skewness risk premiums. This result is similar to the results in Table 6 and also suggests that the skewness risk premium, as well as the variance risk premium, has superior predictive power for future aggregate stock market index returns.

Table 8 The univariate regressions with the CAY

	cay	vp	skp
(A) Monthly return regressions			
Slope coeff.	0.157	0.683	0.008
p value (%)	17.2	0.0***	2.3**
$Adj.R^2$ (%)	0.3	5.5	1.6
(B) Quarterly return regressions			
Slope coeff.	0.418	1.618	0.016
p value (%)	27.4	0.2***	21.5
$Adj.R^2$ (%)	0.2	9.3	0.7

The sample period extends from January 1990 to January 2012. The CAY is the aggregate-consumption wealth ratio defined in Lettau and Ludvigson (2001), which is quarterly-based data and downloaded from Lettau and Ludvigson's web site. For the monthly return regressions, we define a monthly CAY series from the most recent quarterly observation

5 Concluding remarks

In this study, we investigate the skewness risk premium in the financial market under a general equilibrium setting. Extending the long-run risks (LRR) model proposed by Bansal and Yaron (2004) by introducing a stochastic jump intensity for jumps in the LRR factor and the variance of consumption growth rate, we provide an explicit representation for the skewness risk premium, as well as the volatility risk premium, in equilibrium.

On the basis of the representation for the skewness risk premium, we propose a possible reason for the empirical facts of time-varying and negative risk-neutral skewness. Moreover, we also provide an equity risk premium representation of a linear factor pricing model with the variance and skewness risk premiums. The empirical results imply that the skewness risk premium, as well as the variance risk premium, has superior predictive power for future aggregate stock market index returns, which are consistent with the theoretical implication derived by our model. Compared with the variance risk premium, the results show that the skewness risk premium plays an independent and essential role for predicting the market index returns.

Some recent studies such as those by Bali and Hovakimian (2009), Yan (2011), Chang et al. (2013), Driessen et al. (2012), and Rehman and Vilkov (2012) focus on a significant relationship between skewness or jump risks and expected stock returns and they provide empirical evidence for a significantly positive link between the expected stock returns and the jump or skewness risks. To the best of our knowledge, this study is the first to provide a theoretical implication in their empirical evidence in terms of the LRR model approach pioneered by Bansal and Yaron (2004). It remains some challenges for future research on providing more explicit theoretical explanation for the results presented by the recent studies cited above with the theoretical implication shown in this paper. And moreover, it also needs a detailed analysis on the reasons why the skewness and variance risks are priced differently and, in particular, independently of each other. Further insight into this aspect is left to further work.

Appendix 1: Proof of Proposition 2

From (14),

$$
\begin{aligned}
\mathrm{Var}^{\mathbb{P}}_{t+1}(r_{m,t+2}) &= B_r^t G_{t+1} G_{t+1}^t B_r + \sum_i B_r^2(i)\mathrm{Var}^{\mathbb{P}}_{t+1}(J_{i,t+2}) \\
&= B_r^t G_{t+1} G_{t+1}^t B_r + B_r^{2^t}\Psi^{(2)}_{t+1}(0),
\end{aligned} \tag{28}
$$

where

$$
\begin{aligned}
B_r &= \kappa_{1,m} A_m + e_d(\because (15)) \\
&\equiv \left(B_r(1)\ B_r(2)\ B_r(3)\ B_r(4)\ B_r(5)\ B_r(6)\right)^t \in \mathbb{R}^6, \\
B_r^2 &\equiv \left(B_r^2(1)\ B_r^2(2)\ B_r^2(3)\ B_r^2(4)\ B_r^2(5)\ B_r^2(6)\right)^t \in \mathbb{R}^6, \\
\Psi^{(2)}_{t+1}(0) &\equiv \left(0\ \ \Psi^{(2)}_{t+1,x}(0)\ \ \Psi^{(2)}_{t+1,\sigma^2}(0)\ 0\ 0\ 0\right)^t \in \mathbb{R}^6,
\end{aligned}
$$

and $\Psi^{(2)}_{t+1,x}(0)$ and $\Psi^{(2)}_{t+1,\sigma^2}(0)$ are respectively the second derivative of the cgf (cumulant-generating function) for $J_{x,t+1}$ and $J_{\sigma^2,t+1}$ evaluated at 0, that is,

$$
\begin{aligned}
\Psi^{(2)}_{t+1,x}(0) &\equiv \frac{\partial^2}{\partial u^2}\Psi_{t+1,x}(u)\mid_{u=0}= \frac{\partial^2}{\partial u^2}\lambda_{x,t+1}(\psi_x(u)-1)\mid_{u=0}, \\
\Psi^{(2)}_{t+1,\sigma^2}(0) &\equiv \frac{\partial^2}{\partial u^2}\Psi_{t+1,\sigma^2}(u)\mid_{u=0}= \frac{\partial^2}{\partial u^2}\lambda_{\sigma^2,t+1}(\psi_{\sigma^2}(u)-1)\mid_{u=0}.
\end{aligned}
$$

Thus the expression of (28) is rearranged to the following representation,

$$
\begin{aligned}
\mathrm{Var}^{\mathbb{P}}_{t+1}(r_{m,t+2}) &= B_r^t G_{t+1} G_{t+1}^t B_r + B_r^{2^t}\Psi^{(2)}_{t+1}(0) \\
&= B_r^t(H_{\sigma^2}\sigma^2_{t+1} + H_q q_{t+1})B_r + B_r^{2^t}\mathrm{diag}\left(\psi^{(2)}(0)\right)\Pi_{t+1},
\end{aligned}
$$

where

$$
H_{\sigma^2} \equiv \begin{pmatrix} \varphi_\eta^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \varphi_e^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \varphi_\zeta^2 \end{pmatrix}, \quad H_q \equiv \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varphi_\xi^2 & \rho\varphi_\xi\varphi_u & 0 \\ 0 & 0 & 0 & \rho\varphi_\xi\varphi_u & \varphi_u^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},
$$

$$
\mathrm{diag}\left(\psi^{(2)}(0)\right) \equiv \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \psi_x^{(2)}(0) & 0 & 0 & 0 & 0 \\ 0 & 0 & \psi_{\sigma^2}^{(2)}(0) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Pi_{t+1} \equiv \begin{pmatrix} 0 \\ \lambda_{x,t+1} \\ \lambda_{\sigma^2,t+1} \\ 0 \\ 0 \\ 0 \end{pmatrix}.
$$

Appendix 2: Proof of Proposition 3

From the definition of the variance risk premium [18] and the expressions of the conditional variance of the market return $r_{m,t+2}$ at time $t+1$ under each of the probability measures, we can derive the following expression,

$$
\begin{aligned}
vp_t = &-B_r^t \Big[\Lambda_{\sigma^2} H_{\sigma^2} + \varphi_\xi (\varphi_\xi \Lambda_q + \rho \varphi_u \Lambda_\lambda) H_q \Big] B_r q_t \\
&+ B_r^t H_{\sigma^2} B_r \Big[\mathbb{E}_t^{\mathbb{Q}} [J_{\sigma^2,t+1}^{\mathbb{Q}}] - \mathbb{E}_t^{\mathbb{P}} [J_{\sigma^2,t+1}^{\mathbb{P}}] \Big] \\
&+ B_r^{2^t} \Big[\mathrm{diag}(\psi^{(2)}(-\Lambda)) \mathbb{E}_t^{\mathbb{Q}} [\Pi_{t+1}] - \mathrm{diag}(\psi^{(2)}(0)) \mathbb{E}_t^{\mathbb{P}} [\Pi_{t+1}] \Big], \quad (29)
\end{aligned}
$$

where $\Lambda_q \equiv (1-\theta)\kappa_1 A_q$, $\Lambda_\lambda \equiv (1-\theta)\kappa_1 A_\lambda$ (see [13]), and

$$
\mathrm{diag}\Big(\psi^{(2)}(-\Lambda)\Big) \equiv
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & \psi_x^{(2)}(-\Lambda_x) & 0 & 0 & 0 & 0 \\
0 & 0 & \psi_{\sigma^2}^{(2)}(-\Lambda_{\sigma^2}) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
$$

$$
\mathbb{E}_t^{\mathbb{Q}} [\Pi_{t+1}] \equiv \Big(0 \; \mathbb{E}_t^{\mathbb{Q}} [\lambda_{x,t+1}] \; \mathbb{E}_t^{\mathbb{Q}} [\lambda_{\sigma^2,t+1}] \; 0 \; 0 \; 0 \Big)^t,
$$

$$
\mathbb{E}_t^{\mathbb{P}} [\Pi_{t+1}] \equiv \Big(0 \; \mathbb{E}_t^{\mathbb{P}} [\lambda_{x,t+1}] \; \mathbb{E}_t^{\mathbb{P}} [\lambda_{\sigma^2,t+1}] \; 0 \; 0 \; 0 \Big)^t,
$$

Substituting the following facts,

$$
\mathbb{E}_t^{\mathbb{Q}} [J_{\sigma^2,t+1}^{\mathbb{Q}}] = \lambda_{\sigma^2,t} \psi_{\sigma^2}^{(1)}(-\Lambda_{\sigma^2}),
$$

$$
\mathbb{E}_t^{\mathbb{P}} [J_{\sigma^2,t+1}^{\mathbb{P}}] = \lambda_{\sigma^2,t} \psi_{\sigma^2}^{(1)}(0),
$$

into (29) and considering [6] and [16], we can obtain the representation [20]. $\quad\square$

Appendix 3: The risk-free rate

The explicit expression of the risk-free rate can be obtained by substituting $r_{f,t}$ into $r_{j,t+1}$ in [3]. We finally provide the following proposition on the risk-free rate $r_{f,t}$.

Proposition 6 (The Risk-Free Rate) *The risk free rate is expressed as follows with the state variables of σ_t^2, q_t, and λ_t.*

$$
r_{f,t} = \beta_{rf,c} + \beta_{rf,x} x_t + \beta_{rf,\sigma} \sigma_t^2 + \beta_{rf,q} q_t + \beta_{rf,\lambda} \lambda_t,
$$

where

$$\beta_{rf,c} \equiv -\theta \log \delta + \gamma \mu_g - (\theta - 1)(\kappa_0 - A_0)$$
$$-(\theta - 1)\kappa_1(A_0 + A_\sigma \mu_\sigma + A_q \mu_q + A_\lambda \mu_\lambda),$$

$$\beta_{rf,x} \equiv \gamma - (\theta - 1)A_x(\kappa_1 \rho_x - 1),$$

$$\beta_{rf,\sigma} \equiv (1 - \theta)A_\sigma(\kappa_1 \rho_\sigma - 1) - \frac{1}{2}\Big[\gamma^2 \varphi_\eta^2 + (\theta - 1)^2 \kappa_1^2 A_x^2 \varphi_e^2\Big],$$

$$\beta_{rf,q} \equiv (1 - \theta)A_q(\kappa_1 \rho_q - 1) - \frac{1}{2}(\theta - 1)^2 \kappa_1^2$$
$$\Big[A_\sigma^2 + A_q^2 \varphi_\xi^2 + 2A_q A_\lambda \varphi_\xi \varphi_u \rho + A_\lambda^2 \varphi_u^2\Big],$$

$$\beta_{rf,\lambda} \equiv (1 - \theta)A_\lambda(\kappa_1 \rho_\lambda - 1)$$
$$-l_x\Big[\exp\Big(\frac{1}{2}(\theta - 1)^2 \kappa_1^2 A_x^2\Big) - 1\Big] - l_\sigma\Big[\exp\Big(\frac{1}{2}(\theta - 1)^2 \kappa_1^2 A_\sigma^2\Big) - 1\Big].$$

References

Aït-Sahalia, Y., Lo, A.: Nonparametric estimation of state-price-densities implicit in financial asset prices. J Financ **53**, 499–547 (1998)

Aït-Sahalia, Y., Wang, W., Yared, F.: Do option markets correctly price the probabilities of movement of the underlying asset? J Econ **102**, 67–110 (2001)

Bakshi, G., Kapadia, N., Madan, D.: Stock return characteristics, skew laws, and the differential pricing of individual equity options. Rev Financ Stud **16**, 101–143 (2003)

Bali, T., Hovakimian, A.: Volatility spread and expected stock returns. Manage Sci **55**, 1797–1812 (2009)

Bansal, R., Yaron, A.: Risks for the long run: a potential resolution of asset-pricing puzzles. J Financ **59**, 1481–1509 (2004)

Black, F.: Studies of stock market volatility changes. In: Proceedings of the American Statistical Association, Business and Economic Statistics Section, pp. 177–181, (1976)

Bollerslev, T., Sizova, N., Tauchen, G.: Volatility in equilibrium: asymmetries and dynamic dependencies. Rev Financ **16**, 31–80 (2012)

Bollerslev, T., Tauchen, G., Zhou, H.: Expected stock returns and variance risk premia. Rev Financ Stud **22**, 4463–4492 (2009)

Campbell, J., Shiller, R.J.: The dividend-price ratio and expectations of future dividends and discount factors. Rev Financ Stud **1**, 195–227 (1988)

Chan, W., Maheu, J.M.: Conditional jump dynamics in stock market returns. J Bus Econ Stat **20**, 377–389 (2002)

Chang, B., Christoffersen, Y.P., Jacobs, K.: Market skewness risk and the cross section of stock returns. J Financ Econ **107**, 46–68 (2013)

Christoffersen, P., Jacobs, K., Ornthanalai, C.: Dynamic jump intensities and risk premiums: evidence from S&P500 returns and options. J Financ Econ **106**, 447–472 (2012)

Cont, R., Tankov, P.: Financial Modeling with Jump Processes. London: Chapman & Hall (2004)

Demeterfi, K., Derman, E., Kamal, M., Zou, J.: A guide to volatility and variance swaps. J Deriv **6**, 9–32 (1999)

Drechsler, I., Yaron, A.: What's vol got to do with it? Rev Financ Stud **24**, 1–45 (2011)

Driessen, J., Lin, T., Lu, X.: Why do option prices predict stock returns? Working paper (2012)

Epstein, L., Zin, S.: Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica **57**, 937–969 (1989)

Eraker, B.: The volatility premium. Working paper, Department of Finance, University of Wisconsin (2008)

Lettau, M., Ludvigson, S.: Consumption, and expected stock returns. J Financ **56**, 815–849 (2001)

Neuberger, A.: Realized skewness. Rev Financ Stud **26**, 2174–2203 (2012)

Rehman, Z., Vilkov, G.: Risk-neutral skewness: return predictability and its sources. Working paper (2012)

Yan, S.: Jump risk, stock returns, and slope of implied volatility smile. J Financ Econ **99**, 216–233 (2011)