Management Science

Forward-Looking Market Risk Premium

Jin-Chuan Duan, Weiqi Zhang

To cite this article:
Jin-Chuan Duan, Weiqi Zhang (2014) Forward-Looking Market Risk Premium. Management Science 60(2):521-538. http://dx.doi.org/10.1287/mnsc.2013.1758

MANAGEMENT SCIENCE
Vol. 60, No. 2, February 2014, pp. 521–538
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2013.1758

Forward-Looking Market Risk Premium

Jin-Chuan Duan
Risk Management Institute, NUS Business School, and Department of Economics, National University of Singapore, Singapore 119613, bizdjc@nus.edu.sg

Weiqi Zhang
Finance Center Münster, University of Münster, 48143 Münster, Germany, weiqi.zhang@wiwi.uni-muenster.de

A method for computing forward-looking market risk premium is developed in this paper. We first derive a theoretical expression that links forward-looking risk premium to investors' risk aversion and forward-looking volatility, skewness, and kurtosis of cumulative return. In addition, investors' risk aversion is theoretically linked to volatility spread, defined as the gap between the risk-neutral volatility deduced from option data and the physical return volatility exhibited by return data. The volatility spread formula serves as the basis for using the generalized method of moments to estimate investors' risk aversion. We adopt the generalized autoregressive conditional heteroskedasticity model for the physical return process and estimate the model using the S&P 500 daily index returns and then deduce the forward-looking variance, skewness, and kurtosis of the corresponding cumulative return. The forward-looking risk premiums are estimated monthly over the sample period of 2001–2010, and all are found to be positive. Furthermore, two asset pricing tests are conducted. First, change in forward-looking risk premiums is negatively related to the S&P 500 holding period return, reflecting that an increase in discount rate reduces current stock prices. Second, market illiquidity positively affects forward-looking risk premium, indicating that forward-looking risk premium contains an illiquidity risk premium component.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2013.1758.

Keywords: risk premium; forward looking; GARCH; options; volatility spread; skewness; kurtosis
History: Received September 21, 2012; accepted April 18, 2013, by Wei Jiang, finance. Published online in *Articles in Advance* August 2, 2013.

1. Introduction

Risk premium is the most important concept in finance, and its modelling has been at the core of the modern finance theory. Risk premium is often used in financial research and applications, such as testing asset pricing models, allocating assets between stocks and bonds, and determining the cost of capital for an investment.

Risk premium is obviously a forward-looking concept. In essence, it is compensation for holding an asset that will yield an uncertain return. In practice, however, the most commonly used method for estimating risk premium is average historical realized excess returns (Welch 2000, Damodaran 2008). Merton (1980) argued that historical risk premium fails to account for the effect of changes in the level of market risk. Simply put, when one moves into a volatile phase, investors are subject to higher uncertainty, so that forward-looking risk premium should become higher. But the historical average of excess returns cannot be expected to reflect changing market conditions when such a rise in market volatility is transient. Furthermore, the historical approach can produce a negative risk premium, especially during a crisis period when realized stock returns are negative. Empirically, stock markets can produce ex post a negative risk premium even for an estimation period longer than 10 years (e.g., from 1973 to 1984 in the U.S. market).

To gauge the magnitude of forward-looking risk premium, several studies surveyed academics, investors or business managers to get their views on risk premium (Welch 2000, Graham and Harvey 2007, Fernandez 2009a). For example, Welch (2000) surveyed 226 financial economists in 1997 and reported their forecast of long-term mean equity risk premium to be 6% to 7% per annum. Respondents claimed to revise their forecasts downward when stock markets rise. Although survey approaches may provide reasonable estimates of forward-looking risk premium, they are subject to limitations such as that (1) surveys are time consuming and thus cannot be updated frequently enough to remain timely, (2) survey approaches usually prescribe a very long prediction horizon and are not available for different horizons of interest, and (3) surveys are expressions of subjective opinions and face an unknown sample selection bias.

In this paper, we propose a practical method for estimating forward-looking market risk premiums.

We first derive the forward-looking risk premium as a function of investors' risk aversion and forward-looking physical moments (volatility, skewness, and kurtosis) of cumulative return. Then, we estimate investors' risk aversion using the volatility spread formula developed in Bakshi and Madan (2006). Instead of using the realized return time series to estimate physical return moments as in Bakshi and Madan (2006), we deduce forward-looking physical return moments from the generalized autoregressive conditional heteroskedasticity (GARCH) model estimated with daily returns. The GARCH model offers a practical way of reflecting prevailing market condition and provides us with forward-looking physical return moments for any horizon of interest.[1] An important part of our method is the forward-looking risk premium formula that links risk premium to forward-looking physical volatility, skewness, and kurtosis. With a risk aversion parameter estimate in place, we can combine it with forward-looking physical moments to produce our estimate of forward-looking risk premium for any horizon of interest.

In a recent paper, Santa-Clara and Yan (2010) used a different parametric model and derived the risk premium as a function of two latent variables (volatility and jump intensity). Their implementation avoids filtering by assuming two option prices are observed without error at any time point so as to enable them to back out the two latent variables for different points of time. In contrast to these methods, our approach relies on a generic moment expansion, which does not need any parametric option pricing model. In our empirical analysis, we utilize the result developed in the model-free risk-neutral pricing literature to extract risk-neutral volatility from option portfolios without having to deal with individual options.

Our estimates of forward-looking risk premium are based on the S&P 500 index return and option data and are repeatedly estimated on a monthly basis. The estimates range from 0.09% per month (June 2005) to 26.04% per month (October 2008) with higher premiums during extreme market periods. This result is consistent with the common belief that investors require higher compensation for taking higher risk, and thus risk premium should be high when the market is uncertain. In contrast, the estimates of risk premium using historical excess returns, the capital asset pricing model (CAPM), and the Fama and French (1996) three-factor model are unable to adequately reflect market conditions. It is worth noting that forward-looking risk premiums are positive throughout the entire sample period, whereas the risk premiums estimated from other methods (historical average excess returns, the CAPM, and the Fama–French three-factor model) are often negative.

Similarly, the Fama and French (2002) approach to estimating equity risk premium using fundamentals (dividend and earnings growth rates) can be quite volatile. Our empirical analysis shows that it can yield substantially negative risk premiums during bad times. In summary, comparing different risk premium measures suggests that forward-looking risk premium is a more reasonable way of gauging the appropriate level of compensation for bearing risk in a fast-moving equity market.

Two asset pricing implications are tested using the forward-looking risk premium. First, we confirm the theoretical relationship that an increase in discount rate (risk premium) decreases current stock price when controlling for expected future cash flows. Second, we show that forward-looking risk premium is positively related to illiquidity, i.e., the presence of an illiquidity premium. The relationship is not simply a manifestation of the liquidity–volatility relationship.

The remainder of this paper is organized as follows. Section 2 presents the theory of forward-looking risk premium. Section 3 presents the econometrics for estimating investors' risk aversion and for deducing the variance, skewness, and kurtosis of cumulative return from the GARCH model. Section 4 describes the data, the estimates of forward-looking risk premium, and the comparisons with other risk premium measures. Section 5 presents the analysis of two asset pricing implications of forward-looking risk premium, and §6 concludes.

2. The Theory of Forward-Looking Risk Premium

In this section, we first show that the risk-free interest rate can be expressed by the risk-neutral moments of the market portfolio. Then, we derive the expression of forward-looking market risk premium under the standard assumption of a stochastic discount factor that can be justified by a power utility. It is common practice in the asset pricing literature to derive an asset pricing model by combining a particular form of the stochastic discount factor with lognormal asset returns (Hansen and Singleton 1983, Grossman and Shiller 1981, Campbell and Cochrane 2000). Instead of assuming lognormal asset returns, we allow for higher moments in the derivation of market risk premium to reflect well-known empirical irregularities.

Denote the market portfolio's value by S_t and its cumulative return over the time period t to $t+\tau$ (continuously compounded) by $R_t(\tau) = \ln(S_{t+\tau}/S_t)$. At time t, $R_t(\tau)$ is a random return to be realized later at time $t+\tau$. Let $r_t(\tau)$ and $\delta_t(\tau)$ denote

[1] To understand this point, assume that the daily return time series is governed by the GARCH(1, 1) model. The historical volatility based on, say, 90 daily returns will be different from the volatility of the 90-day cumulative return implied by the GARCH model on a forward-looking basis.

the continuously compounded risk-free interest rate and dividend yield of the market portfolio over the period from t to $t+\tau$, respectively. To characterize the distributions implied by return and option data, we need to specify two probability measures. Let $\mu_{Pt}(\tau)$, $\sigma_{Pt}(\tau)$, $\theta_{Pt}(\tau)$, and $\kappa_{Pt}(\tau)$ be the mean, standard deviation, skewness, and kurtosis of the market portfolio under the physical measure P, respectively. The use of t and τ is to make it clear that these moments can be time varying and depend on the length of the period over which the cumulative return is defined. Their equivalents under the risk-neutral measure Q are denoted with the subscript Q.

We can first derive an approximate relationship for the equilibrium risk-free interest rate, since the expected asset return inclusive of cash dividends should equal the risk-free interest rate when the expectation is performed under the risk-neutral measure. By a simple expansion argument (see Appendix A for details), we have the following result:

$$r_t(\tau) \approx \delta_t(\tau) + \mu_{Qt}(\tau) + \tfrac{1}{2}\sigma^2_{Qt}(\tau) + \tfrac{1}{6}\theta_{Qt}(\tau)\sigma^3_{Qt}(\tau) + \tfrac{1}{24}\sigma^4_{Qt}(\tau)[\kappa_{Qt}(\tau) - 3]. \tag{1}$$

In the above, the risk-free interest rate is expressed as a function of risk-neutral moments, and the approximate relationship is generic in the sense that it does not depend on the form of the stochastic discount factor.

To obtain a useful expression for the physical market risk premium, we later need to express the risk-free rate in terms of physical return moments. Our derivations are based on the following assumption.

Assumption 1. *The stochastic discount factor over time t to $t+\tau$ is $e^{-\gamma R_t(\tau)}$, and the moment generating function of $R_t(\tau)$ exists under either measure P or Q.*[2]

Under the above assumption, Bakshi and Madan (2006) derived an expression for volatility spread:

$$\frac{\sigma^2_{Qt}(\tau) - \sigma^2_{Pt}(\tau)}{\sigma^2_{Pt}(\tau)} \approx -\gamma\sigma_{Pt}(\tau)\theta_{Pt}(\tau) + \frac{\gamma^2}{2}\sigma^2_{Pt}(\tau)[\kappa_{Pt}(\tau) - 3]. \tag{2}$$

Similar to their study, the above expression later serves as the basis for our empirical estimation of the risk aversion parameter γ.

One can similarly derive analytical expressions for risk-neutral expected return, variance, skewness, and kurtosis in terms of physical return moments. Substituting these expressions for risk-neutral moments into the risk-free rate equation in (1), the following new expression for market risk premium can be derived.

Proposition 1. *Under Assumption 1, the τ-period market risk premium can be expressed as a function of investors' risk aversion, physical return variance, skewness, and kurtosis*:

$$\mu_{Pt}(\tau) + \delta_t(\tau) - r_t(\tau) \approx \left(\gamma - \frac{1}{2}\right)\sigma^2_{Pt}(\tau) - \frac{3\gamma^2 - 3\gamma + 1}{6}\sigma^3_{Pt}(\tau)\theta_{Pt}(\tau) + \frac{4\gamma^3 - 6\gamma^2 + 4\gamma - 1}{24}\sigma^4_{Pt}(\tau)[\kappa_{Pt}(\tau) - 3]. \tag{3}$$

Proof. See Appendix B.

Suppose that there is no physical return skewness or excess kurtosis; that is, $\theta_{Pt}(\tau) = 0$ and $\kappa_{Pt}(\tau) = 3$. The above result implies that risk premium for the market portfolio is $\mu_{Pt}(\tau) + \delta_t(\tau) - r_t(\tau) = (\gamma - \frac{1}{2})\sigma^2_{Pt}(\tau)$, a well-known result under lognormality. The equity premium expression in Equation (3) makes it easier to understand the role played by return skewness and kurtosis. It suggests that the presence of skewness and excess kurtosis will alter the risk premium. One can show that $3\gamma^2 - 3\gamma + 1$ is always positive, which implies that negative skewness will increase risk premium. The importance of negative skewness in pricing assets has been documented previously in, for example, Kraus and Litzenberger (1976) and Harvey and Siddique (2000). Similarly, one can show that $4\gamma^3 - 6\gamma^2 + 4\gamma - 1 > 0$ when $\gamma > \frac{1}{2}$. Therefore, when investors' risk aversion exceeds one-half, leptokurtosis (fat tails) will also increase the risk premium.

If one can find a practical way to estimate γ and physical return moments for different horizons of interest on a forward-looking basis, Equation (3) will provide a way to generate forward-looking market risk premiums for different horizons of interest. Indeed, that is what we will do next.

3. Econometric Formulation

Similar to Bakshi and Madan (2006), we use the volatility spread equation in (2) to estimate γ. Let I_t be some set of instruments whose values are known at time t. A generalized method of moments (GMM) estimation can be performed using the following orthogonality condition:

$$E\left\{\frac{\sigma^2_{Qt}(\tau) - \sigma^2_{Pt}(\tau)}{\sigma^2_{Pt}(\tau)} + \gamma\sigma_{Pt}(\tau)\theta_{Pt}(\tau) - \frac{\gamma^2}{2}\sigma^2_{Pt}(\tau)[\kappa_{Pt}(\tau) - 3] \,\middle|\, I_t\right\} = 0. \tag{4}$$

To utilize the above restriction, one needs a time series of risk-neutral return variance and three time series of physical return moments (variance, skewness, and kurtosis).

[2] Note that the stochastic discount factor in Assumption 1 can be deduced from the power utility function: $U(W) = W^{1-\gamma}/(1-\gamma)$ when the economic agent maximizes the expected utility of the end-of-the-period wealth.

A model-free risk-neutral return variance $\sigma^2_{Qt}(\tau)$ can be computed by forming appropriate portfolios of broad-based market index options. Such an approach was established in Britten-Jones and Neuberger (2000), Carr and Madan (2001), Bakshi et al. (2003), and Jiang and Tian (2005). The theory linking risk-neutral return variance to an option portfolio is presented in Appendix C. Its exact empirical implementation will be elaborated in the next section.

In this paper, physical return variance, skewness, and kurtosis are obtained using the GARCH model. Our approach thus differs from the ex post sample moments approach of Bakshi and Madan (2006). Using a popular GARCH model with the feature of asymmetric volatility response (i.e., leverage effect), we are able to deduce *forward-looking* higher return moments for various horizons of interest through a combination of analytical formulas and bootstrap sampling.

We adopt the nonlinear asymmetric GARCH(1,1) model of Engle and Ng (1993), hereafter NGARCH (1,1), for the market portfolio's return dynamics under the physical probability measure P:

$$\ln\frac{S_{t+1}}{S_t} = \mu + \sigma_{t+1}\varepsilon_{t+1} \quad \text{for } t = 0, 1, \ldots, \tag{5}$$

where

$$\sigma^2_{t+1} = \beta_0 + \beta_1\sigma^2_t + \beta_2\sigma^2_t(\varepsilon_t - \eta)^2, \tag{6}$$

and ε_{t+1} are independent and identically distributed random variables with $E^P_t(\varepsilon_{t+1}) = 0$ and $E^P_t(\varepsilon^2_{t+1}) = 1$. We impose the restrictions: $\beta_0 > 0$, $\beta_1 \geq 0$, and $\beta_2 \geq 0$ to ensure that conditional variances always stay positive. Parameter η reflects the leverage effect. According to Duan (1997), the NGARCH(1,1) model is strictly stationary if $\beta_1 + \beta_2(1+\eta^2) \leq 1$.

To allow for skewness and fat tails, we do not commit to a particular conditional distribution. The parameters in the NGARCH(1,1) model are estimated by the quasi–maximum likelihood method. After obtaining the parameters, we estimate the physical return moments for multiperiod horizons, specifically, 20 trading days (corresponding to 28 calendar days).

The conditional variance of τ-period cumulative return can be estimated analytically using a simple formula. The derivation of the following formula is given in Appendix D:

$$\sigma^2_{Pt}(\tau) = \frac{1-\lambda^\tau}{1-\lambda}\sigma^2_{t+1} + \frac{(\tau-1)\beta_0}{1-\lambda} - \frac{\lambda(1-\lambda^{\tau-1})\beta_0}{(1-\lambda)^2}, \tag{7}$$

where $\lambda = \beta_1 + \beta_2(1+\eta^2)$.

However, conditional skewness and kurtosis of the cumulative return under the NGARCH model do not lead to workable closed-form formulas. Thus, we resort to parametric bootstrapping to obtain these required quantities. Basically, we use the data set available at time t to obtain an estimated NGARCH(1,1) model. The model is then applied to the data set to generate a time series of standardized residuals (mean 0 and variance 1, but not necessarily normally distributed). When simulating the NGARCH(1,1) model to obtain cumulative returns, we start from σ_{t+1} and randomly sample from the set of standardized residuals to move the system forward one day at a time until it reaches time $t+\tau$ so as to compute the cumulative return of maturity τ. The smooth stratified bootstrap method in Malik and Pitt (2011) is applied to the sampling from the standardized residuals. After repeating the sampling of the cumulative return many times, conditional skewness and kurtosis of the cumulative return of interest can be approximated by their sample equivalents.

4. Empirical Analysis

4.1. Data

The S&P 500 index returns and option prices are used in the empirical study. The S&P 500 index values, their option prices, and the risk-free interest rates over the period of January 1996 to October 2010 are taken from OptionMetrics. Monthly sampling frequency is implemented, and τ in our estimation equals 28 calendar days. At each option expiration date, we move backward 28 calendar days and refer to this point as the observation date. This procedure gives us nonoverlapping call and put options with a maturity of 28 calendar days. The corresponding 28-day risk-free rate is obtained by interpolating the zero-rate curve.

Risk-neutral variance of the 28-calendar day cumulative return is calculated for each observation date using the prices of the S&P 500 index options with the remaining maturity of 28 calendar days. The formula used to calculate risk-neutral variance is shown in Appendix C. We follow CBOE (Chicago Board Options Exchange) to set K_t (determining at time t which calls and puts are considered out of the money in the algorithm) as the first available strike price below the forward index level where the forward index level is determined by the call–put pair with the smallest price differential. Numerical integration is performed over the available strike prices. As argued in Jiang and Tian (2005), the discretization error is unlikely to have material impact on the calculation of risk-neutral variance.

The NGARCH(1,1) model is used to compute the 28-calendar day conditional physical variance, skewness, and kurtosis. The daily S&P 500 closing index values over the five years immediately before each

Table 1 Summary Statistics of Return Moments

	$\sigma_Q(\tau)$	$\sigma_P(\tau)$	$\theta_P(\tau)$	$\kappa_P(\tau)$
No. of months	178	178	178	178
Mean	22.11	18.44	−1.11	6.88
Standard deviation	9.97	10.05	0.41	2.80
Minimum	9.95	7.71	−2.07	3.65
Maximum	84.29	82.66	−0.23	19.46

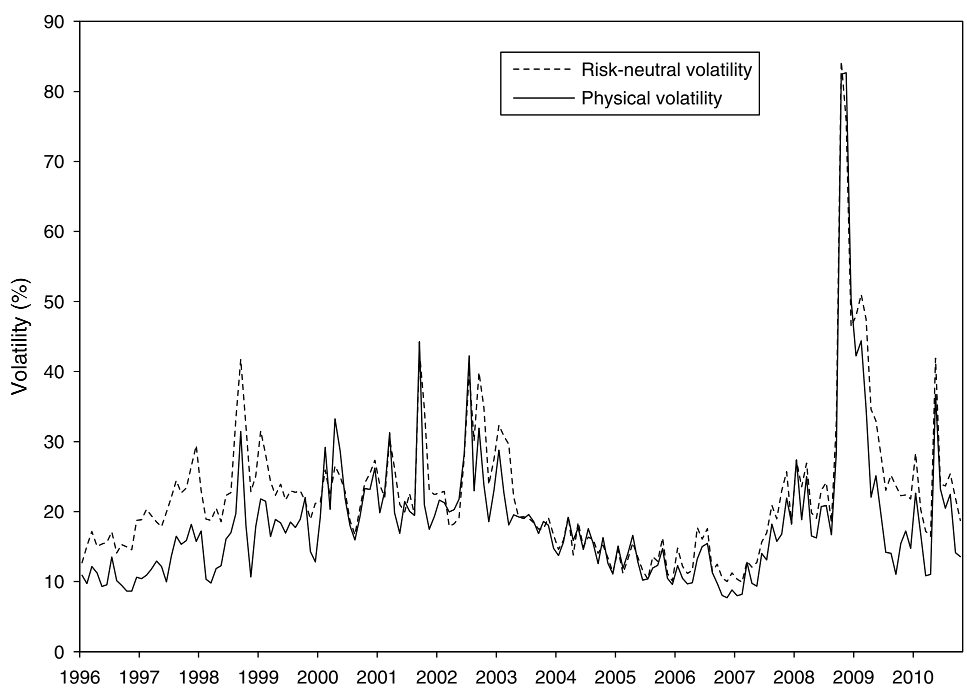
Notes. This table shows the summary statistics for the 28-calendar day risk-neutral volatility (σ_Q) and physical forward-looking volatility (σ_P), skewness (θ_P), and kurtosis (κ_P). The sample period is from January 1996 to October 2010. All volatilities are expressed in annualized percentage terms. Risk-neutral volatility (σ_Q) is calculated using S&P 500 index option prices. Physical forward-looking volatility (σ_P) is calculated analytically using the NGARCH(1, 1) model estimated to a five-year moving window of the S&P 500 index returns. Physical forward-looking skewness (θ_P) and kurtosis (κ_P) are computed by smoothed bootstrap simulations. τ equals 28 calendar days.

observation date are used in the quasi–maximum likelihood estimation of the NGARCH(1, 1) model to obtain the parameter estimates for μ, β_0, β_1, β_2, and η. In addition, the conditional variance (physical) of the next trading day return, i.e., σ_{t+1}^2, corresponding to the observation date is obtained as a by-product of the estimation. For each observation date, we then calculate the 28-calendar day conditional variance analytically using Equation (7). Because the GARCH parameters are obtained on a trading-day basis, we apply 20 trading days as corresponding to the next 28-calendar day period. The conditional 28-calendar day skewness and kurtosis are obtained by smoothed bootstrap simulations using the pool of five years worth of standardized residuals corresponding to the observation date in an attempt to preserve skewness and kurtosis in the data. A bootstrapped sample size of 100,000 is used to advance the system one trading day at a time until reaching the 28-calendar day maturity. Again, the actual number of trading days in the next 28 calendar days is used in simulation. We then compute the averages of simulated cumulative returns raised to various powers to the sample equivalent moments.

Table 1 presents the summary statistics for risk-neutral volatility and physical forward-looking volatility, skewness, and kurtosis. Qualitatively consistent with the prior findings in the literature, risk-neutral volatility has a higher mean value vis-à-vis physical forward-looking volatility. It is also revealed by the summary statistics that the 28-calendar day forward-looking returns are negatively skewed ($\theta_P(\tau) < 0$) and leptokurtic ($\kappa_P(\tau) > 3$).

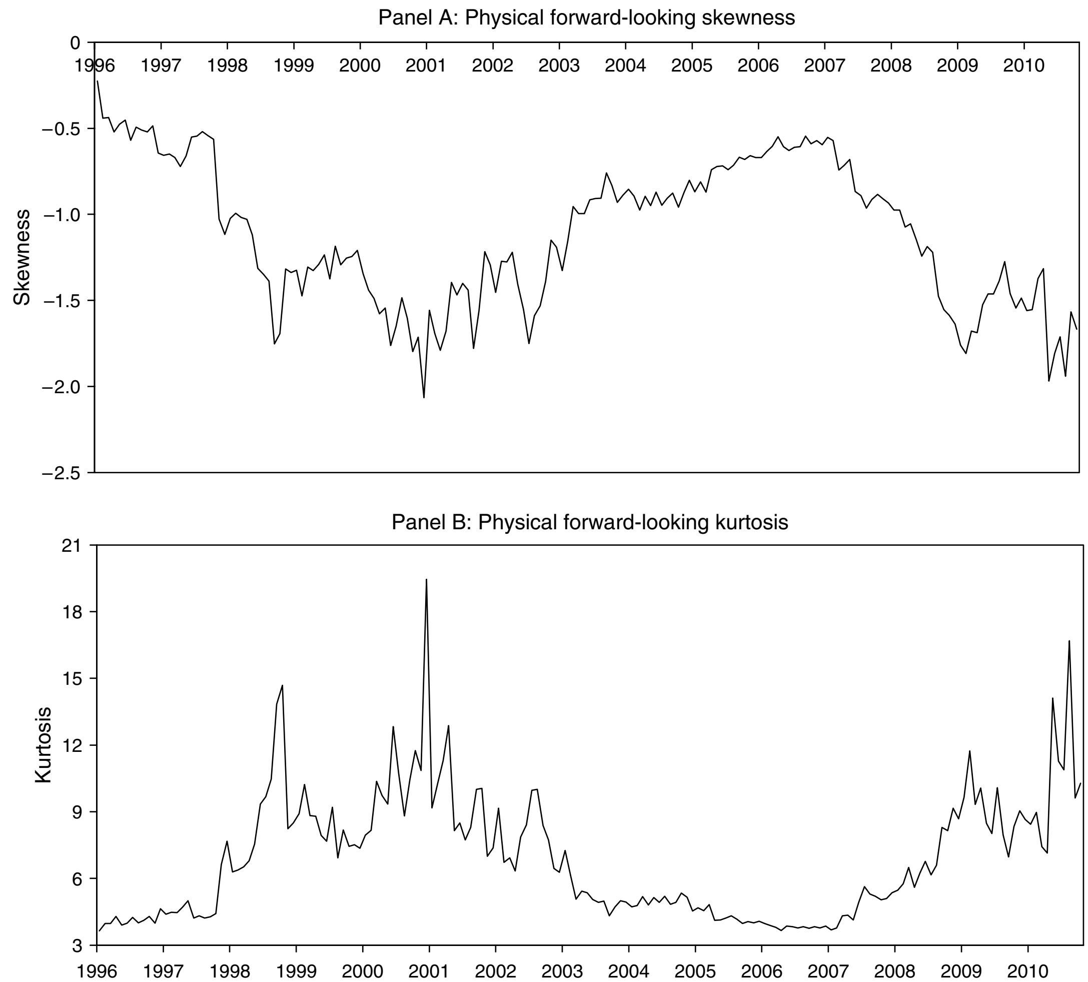
Figure 1 plots the 28-calendar day risk-neutral volatility versus physical forward-looking volatility. The curve representing risk-neutral volatility generally lies above the one for physical forward-looking volatility, especially for the period of 1996–1999. The average volatility spread between risk-neutral volatility and physical forward-looking volatility over the whole sample period is 3.7%. The average volatility spread for the period from January 1996 to December 1999 is 6.9%, and for the period from January 2000

Figure 1 Time-Series Plots of Risk-Neutral and Physical Forward-Looking Volatilities



Notes. This figure presents the time series of the 28-calender day risk-neutral volatility and physical forward-looking volatility for the S&P 500 index return. The data period is from January 1996 to October 2010. Volatilities plotted are annualized.

Figure 2 **Time-Series Plots of Physical Forward-Looking Skewness and Kurtosis**



Note. This figure presents the time-series plots of the 28-calendar day physical forward-looking skewness (panel A) and kurtosis (panel B) using the S&P 500 index data from January 1996 to October 2010.

to October 2010 is 2.5%. Although we compute physical return volatility differently, our results are qualitatively consistent with the literature that has documented volatility spread (Bakshi and Madan 2006, Christensen and Prabhala 1998). Figure 2 plots the 28-calendar day physical forward-looking skewness and kurtosis computed from the NGARCH(1, 1) model. Panel A shows that returns are generally negatively skewed, and negative skewness is more pronounced from 1998 to 2003, and again from 2007 to 2010. Panel B shows that returns are more leptokurtic from 1998 to 2003, and again from 2007 to 2010.

4.2. Investors' Risk Aversion

To obtain the risk aversion at different time points, we use a five-year moving window of data (updated monthly) to estimate γ. Specifically, γ is estimated for every observation date (one per month) using the five-year data of prior to and including the observation date to generate 60 monthly volatility spreads for the GMM estimation. The GMM method adopted here is the one with the Newey and West (1987) adjusted covariance matrix. Three sets of instruments are used and they are the same as in Bakshi and Madan (2006). Set 1 contains a constant plus $\sigma^2_{Q,\,t-1}(\tau)$. Set 2 contains a constant, $\sigma^2_{Q,\,t-1}(\tau)$, and $\sigma^2_{Q,\,t-2}(\tau)$. Finally, Set 3 contains a constant, $\sigma^2_{Q,\,t-1}(\tau)$, $\sigma^2_{Q,\,t-2}(\tau)$, and $\sigma^2_{Q,\,t-3}(\tau)$. The results from the three sets are qualitatively similar. Therefore, we only report in Table 2 those from using Set 3.

Although we have 178 monthly results for all return moments of interest, we can only conduct the GMM estimation and test on a moving-window basis for 118 times, because the first test needs return moments for five years (60 months). None of the 118 rolling tests of the model are rejected, based on testing the overidentifying restrictions at the 5% significance level. The estimated γ's are all significant with the mean being 3.71, and the smallest t-statistic equal to 2.25. The estimated γ's range from 1.3 to 6.0. It should be noted

Table 2 Investors' Risk Aversion Estimates

End month	γ	$t(\gamma)$	J-stat.	Model p	End month	γ	$t(\gamma)$	J-stat.	Model p
Jan 01	3.90	2.76	7.41	0.06	Dec 05	2.01	4.13	2.96	0.40
Feb 01	3.89	2.74	6.67	0.08	Jan 06	2.30	4.45	2.63	0.45
Mar 01	4.11	3.04	5.27	0.15	Feb 06	2.37	4.30	2.77	0.43
Apr 01	4.05	3.19	5.03	0.17	Mar 06	2.60	4.70	2.63	0.45
May 01	4.06	3.17	4.80	0.19	Apr 06	3.04	4.99	2.08	0.56
Jun 01	4.33	3.50	3.92	0.27	May 06	3.20	5.08	1.88	0.60
Jul 01	4.40	3.79	3.46	0.33	Jun 06	3.20	4.94	1.94	0.58
Aug 01	4.32	3.87	3.34	0.34	Jul 06	3.43	5.05	1.89	0.60
Sep 01	4.72	4.19	2.16	0.54	Aug 06	3.32	5.24	1.89	0.59
Oct 01	5.89	5.72	1.16	0.76	Sep 06	3.51	5.06	1.86	0.60
Nov 01	6.01	6.06	1.30	0.73	Oct 06	3.76	5.35	1.37	0.71
Dec 01	5.87	5.84	1.46	0.69	Nov 06	3.58	4.62	0.27	0.97
Jan 02	5.50	5.93	1.73	0.63	Dec 06	3.69	4.70	0.85	0.84
Feb 02	5.38	5.95	1.56	0.67	Jan 07	3.96	5.91	0.49	0.92
Mar 02	5.28	6.12	1.45	0.69	Feb 07	4.12	6.81	0.83	0.84
Apr 02	5.24	5.97	1.80	0.61	Mar 07	4.26	7.20	0.82	0.85
May 02	5.20	6.39	1.77	0.62	Apr 07	4.36	7.74	0.74	0.86
Jun 02	4.99	5.82	1.65	0.65	May 07	4.54	8.25	0.72	0.87
Jul 02	4.59	5.44	1.47	0.69	Jun 07	4.34	7.83	1.48	0.69
Aug 02	4.58	5.90	1.49	0.69	Jul 07	5.16	8.77	1.22	0.75
Sep 02	4.71	6.01	1.08	0.78	Aug 07	4.42	9.97	2.16	0.54
Oct 02	4.60	7.57	1.52	0.68	Sep 07	4.82	6.48	2.26	0.52
Nov 02	5.20	7.65	1.20	0.75	Oct 07	5.13	6.33	2.38	0.50
Dec 02	4.08	6.35	1.67	0.64	Nov 07	4.72	7.36	1.99	0.57
Jan 03	3.91	7.03	2.42	0.49	Dec 07	4.46	6.37	2.68	0.44
Feb 03	3.90	7.19	2.27	0.52	Jan 08	5.12	5.16	2.85	0.41
Mar 03	3.89	7.15	2.43	0.49	Feb 08	4.21	3.86	2.93	0.40
Apr 03	3.80	7.14	2.55	0.47	Mar 08	4.93	5.00	3.54	0.32
May 03	3.69	6.91	2.96	0.40	Apr 08	2.97	2.57	2.91	0.41
Jun 03	3.54	6.57	3.36	0.34	May 08	3.45	3.74	2.24	0.52
Jul 03	3.54	6.09	3.63	0.30	Jun 08	3.72	4.85	2.05	0.56
Aug 03	3.38	6.30	3.61	0.31	Jul 08	3.82	5.46	1.65	0.65
Sep 03	3.26	5.46	4.18	0.24	Aug 08	4.20	6.64	2.15	0.54
Oct 03	3.45	5.29	3.08	0.38	Sep 08	3.86	5.96	1.95	0.58
Nov 03	3.17	5.41	2.10	0.55	Oct 08	3.19	4.57	2.45	0.48
Dec 03	3.03	4.94	2.26	0.52	Nov 08	2.80	5.03	3.16	0.37
Jan 04	2.73	4.97	2.67	0.45	Dec 08	3.00	3.49	2.73	0.44
Feb 04	2.88	4.52	2.29	0.51	Jan 09	2.84	3.03	2.78	0.43
Mar 04	2.70	4.73	1.94	0.58	Feb 09	2.48	2.84	3.27	0.35
Apr 04	2.48	4.45	2.23	0.53	Mar 09	2.51	3.12	3.34	0.34
May 04	2.48	4.40	2.07	0.56	Apr 09	2.91	3.47	2.90	0.41
Jun 04	2.31	4.36	2.23	0.53	May 09	3.15	3.66	2.80	0.42
Jul 04	2.22	4.21	2.24	0.52	Jun 09	3.43	3.94	2.53	0.47
Aug 04	2.17	4.13	2.32	0.51	Jul 09	3.77	3.91	2.37	0.50
Sep 04	1.93	4.05	2.58	0.46	Aug 09	4.24	3.77	2.23	0.53
Oct 04	1.77	3.68	3.38	0.34	Sep 09	4.65	3.45	2.25	0.52
Nov 04	2.06	3.98	2.45	0.48	Oct 09	4.69	3.42	2.25	0.52
Dec 04	1.93	4.01	1.97	0.58	Nov 09	4.82	3.44	2.25	0.52
Jan 05	1.30	2.50	2.97	0.40	Dec 09	4.90	3.39	2.29	0.52
Feb 05	1.33	2.69	3.35	0.34	Jan 10	4.95	3.46	2.22	0.53
Mar 05	1.33	2.61	3.31	0.35	Feb 10	5.02	3.50	2.22	0.53
Apr 05	1.27	2.25	3.55	0.31	Mar 10	5.20	3.45	2.32	0.51
May 05	1.42	2.73	3.41	0.33	Apr 10	5.39	3.41	2.34	0.50
Jun 05	1.61	3.28	3.16	0.37	May 10	5.29	3.76	2.07	0.56
Jul 05	1.64	3.31	3.13	0.37	Jun 10	5.15	3.78	2.14	0.54
Aug 05	1.75	3.60	3.01	0.39	Jul 10	5.12	3.78	2.00	0.57
Sep 05	1.77	3.60	3.00	0.39	Aug 10	4.84	3.91	2.24	0.52
Oct 05	1.86	3.80	2.99	0.39	Sep 10	5.03	4.04	2.17	0.54
Nov 05	1.91	3.85	3.06	0.38	Oct 10	5.14	4.12	2.14	0.54
					Average	3.71	4.84	2.45	0.51

Notes. This table reports the GMM estimation results based on the following orthogonality condition: $E\{(\sigma_{Qt}^2(\tau) - \sigma_{Pt}^2(\tau))/\sigma_{Pt}^2(\tau) + \gamma\theta_{Pt}(\tau)\sigma_{Pt}(\tau) - (\gamma^2/2)\sigma_{Pt}^2(\tau)[\kappa_{Pt}(\tau)-3] \mid I_t\} = 0$. For every month, we use the five years of data preceding the month to estimate the risk aversion (γ). The instruments are constant and the risk-neutral variances are lagged by one, two, and three periods. The t-statistics for γ, denoted by $t(\gamma)$; the J-statistics for model overidentification (J-stat.); and the J-statistics' p-values (model p) are provided.

that Equation (4) for volatility spread is not scale free, meaning that one must apply the physical and risk-neutral volatilities in their original scale specific to the maturity; for example, one should not annualize monthly volatility. These estimates for risk aversion seem intuitively sensible and are comparable to the ones obtained in some previous studies such as Bliss and Panigirtzoglou (2004), which reports a risk aversion estimate of 4.08 (power utility) or 6.33 (exponential utility) using the risk-neutral and physical density functions of the S&P 500 return and option data.

For comparison, we also estimate γ by applying the approach of using ex post sample moments as in Bakshi and Madan (2006). The results show that the volatility spread model still passes the test on the overidentifying restrictions, but the estimated γ is approximately 93, an unreasonably large risk aversion parameter value. In contrast, the full sample γ estimated using forward-looking physical moments is 5.24. The cause for the huge difference in the parameter estimates can be attributed to the fact that the sample moments based on ex post realized returns are available only at the end of the period of interest, which are incompatible with the spirit of the theoretical relationship. Consequently, it forces the risk aversion parameter to accommodate the gap between forward-looking risk-neutral volatility and ex post physical volatility, and causes a distorted estimate of risk aversion. Another possible reason is that the use of a relatively small sample of returns (daily returns over one month) may have underestimated the magnitude of higher return moments (Jackwerth and Rubinstein 1996) and the underestimated skewness and kurtosis in turn need a much larger risk aversion to match the volatility spread.

4.3. Forward-Looking Risk Premium

Using the estimated risk aversion along with physical forward-looking variance, skewness, and kurtosis, we can compute forward-looking risk premium for each observation date. Table 3 presents the results. The estimated forward-looking risk premiums are all positive and vary from 0.09% (June 2005) to 26.04% (October 2008) per month. Furthermore, the results reveal very high risk premiums during the internet bubble bursting period (2001 to early 2003), the subprime mortgage crisis (late 2007 to 2009), and the subsequent European debt crisis period (late 2009 to the end of our sample). There are several particularly large risk premiums during these two periods. In September 2001, the 9–11 terrorist attack resulted in the closure of the New York Stock Exchange (NYSE) from September 11 to 17, and during the first reopening day, the S&P 500 index fell 4.9% and the Dow Jones Industry Average fell 7.1%, which was the single biggest one-day drop over our sample period. Therefore, it is not surprising that the forward-looking risk premium for that month reaches 11.48%. In July 2002, WorldCom, which was the second largest long distance phone company, filed for bankruptcy. It was the largest bankruptcy up to that time, and investors' confidence was severely shaken (*The Economist* 2002). Our estimated forward-looking risk premium is 9.64% for that month.

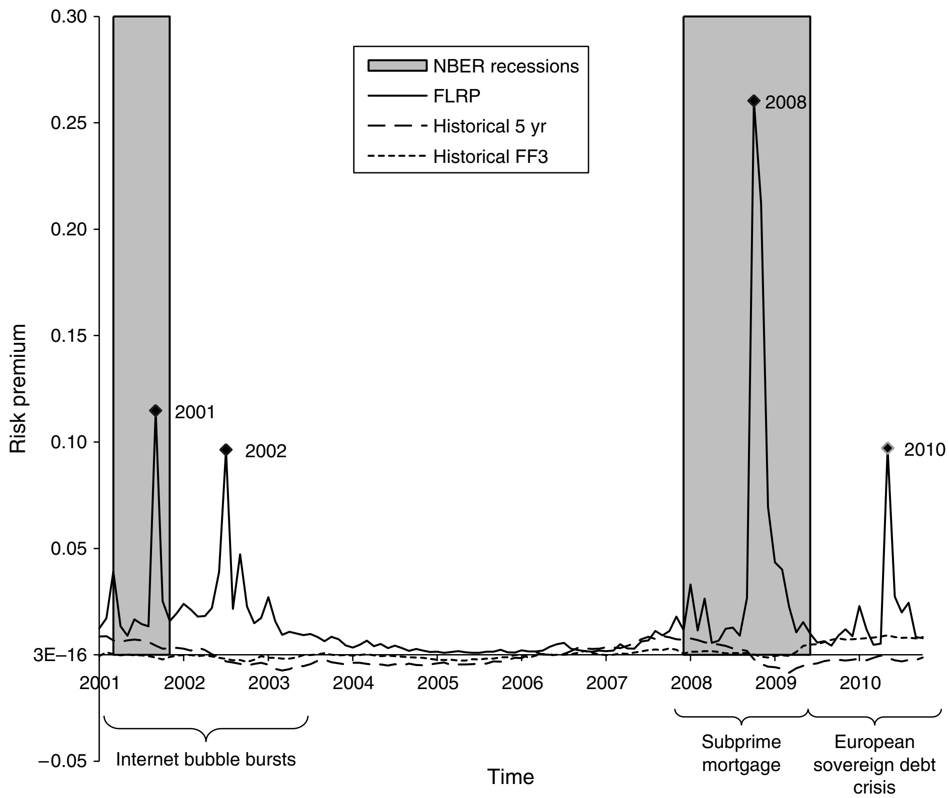
More recently, January 2008 was an especially volatile month for stock markets around the world for fears of the subprime mortgage crisis (*BBC News* 2008). Two months later, Bear Stearns collapsed and was merged with JPMorgan Chase in a distressed sale. The subprime mortgage crisis reached its peak in September and October of 2008. Several major institutions (e.g., Lehman Brothers and Merrill Lynch) either failed or were acquired with government assistance. Interestingly, our forward-looking risk premium hits its highest point of 26.04% in October 2008 and its second highest point of 21.20% in November 2008. Subsequently, there was a flash crash in the U.S. stock market in May 2010, which raised the estimated

Table 3 **S&P 500 Forward-Looking Risk Premium Estimate**

	Average (%)	Jan. (%)	Feb. (%)	Mar. (%)	Apr. (%)	May (%)	Jun. (%)	Jul. (%)	Aug. (%)	Sept. (%)	Oct. (%)	Nov. (%)	Dec. (%)
2001	2.59	1.24	1.72	3.92	1.34	0.91	1.67	1.45	1.34	**11.48**	2.52	1.59	1.95
2002	3.02	2.40	2.13	1.80	1.83	2.20	3.88	**9.64**	2.16	4.72	2.26	1.48	1.73
2003	1.06	2.71	1.57	0.94	1.08	1.01	0.93	0.98	0.83	0.64	0.85	0.72	0.45
2004	0.35	0.34	0.47	0.67	0.41	0.51	0.31	0.43	0.32	0.18	0.27	0.20	0.14
2005	0.13	0.14	0.09	0.13	0.17	0.12	0.09	0.10	0.14	0.15	0.23	0.12	0.11
2006	0.27	0.22	0.16	0.16	0.20	0.38	0.49	0.57	0.29	0.23	0.17	0.15	0.20
2007	0.74	0.18	0.19	0.50	0.30	0.29	0.63	0.68	1.13	0.92	1.13	1.80	1.16
2008	5.71	3.30	1.14	2.64	0.56	0.66	1.23	1.28	0.90	2.67	**26.04**	**21.20**	6.94
2009	1.57	4.34	4.01	2.27	1.05	1.53	1.00	0.57	0.65	0.43	0.91	1.20	0.87
2010	2.30	2.29	1.19	0.48	0.52	**9.72**	2.73	1.99	2.45	0.83	0.78		

Notes. This table reports monthly forward-looking risk premium estimates based on the S&P 500 index values and options from January 2001 to October 2010. The numbers reported are monthly percentage risk premiums. Note that one month is 28 days from the observation date in the month to the subsequent option maturity date (every third Friday of the month). For example, January 2001 is from January 19, 2001, to February 16, 2001. The numbers in bold are the particularly large risk premiums.

Figure 3 Market Risk Premiums



Notes. This figure presents the times series plot of S&P 500 historical and forward-looking monthly risk premiums with NBER recession periods. The sample period is from January 2001 to October 2010. FLRP, forward-looking risk permium; FF3, Fama–French three-factor model.

forward-looking risk premium to 9.72%. The flash crash occurred during the period when the eurozone economies experienced a sovereign debt crisis. In summary, these aforementioned events are coupled with extremely high forward-looking risk premiums that could not possibly be captured by any backward-looking risk premium measure.

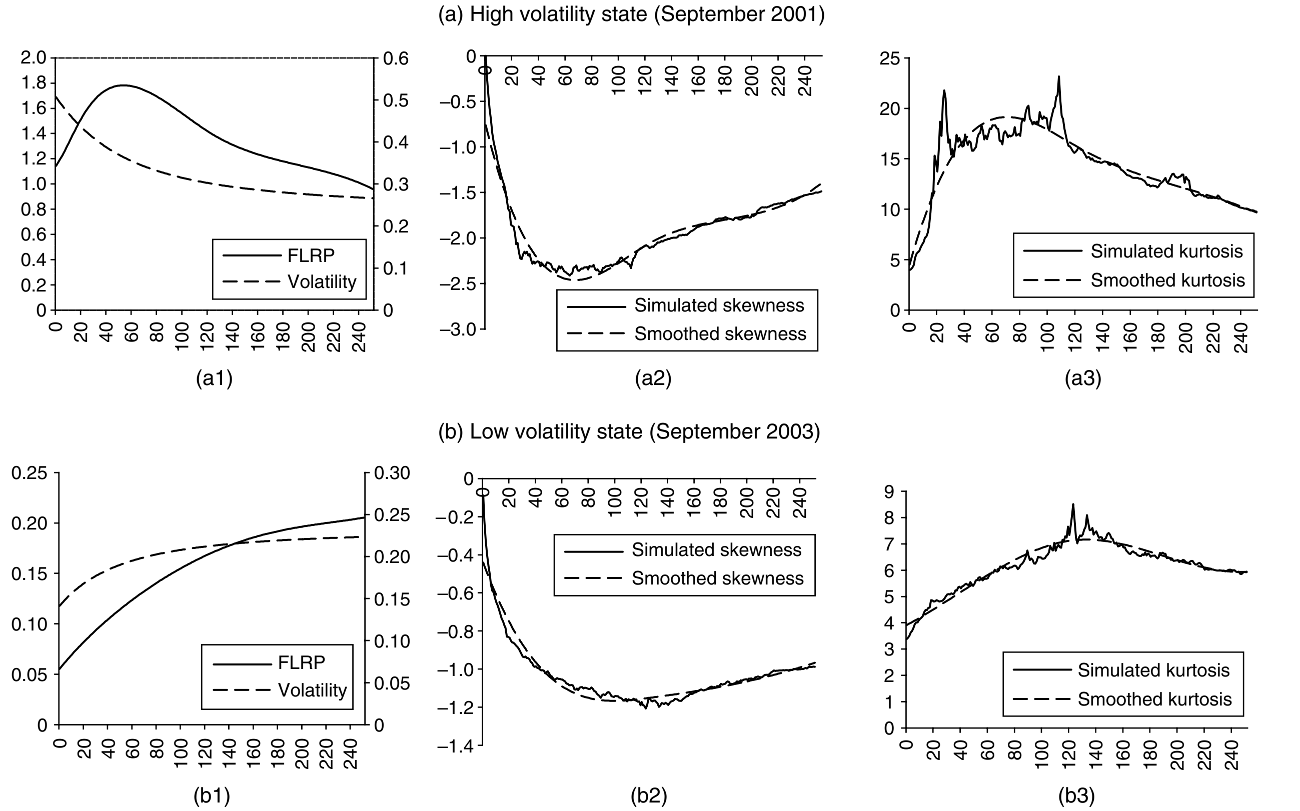
To further explore the relationship between forward-looking risk premiums and economic conditions, in Figure 3 we plot the time series of monthly forward-looking risk premiums along with the National Bureau of Economic Research (NBER) recessions (the shaded area).[3] Along with recessions, we also indicate in the plot the internet bubble bursting period, the subprime mortgage crisis period, and the European sovereign debt crisis period. This figure shows that during a recession or a crisis period, the forward-looking risk premiums are usually higher. This result is consistent with the common belief that during bad times, investors usually demand higher returns.

The forward-looking risk premium estimation can be extended to longer horizons by setting a larger τ using the physical forward-looking return variance, skewness, and kurtosis specific to the horizon. Since physical forward-looking skewness and kurtosis need to be computed with a bootstrapping method, we need to address the simulation errors arising from simulating over a longer horizon. Since the estimated GARCH model has a vary high volatility persistence, simulation noise cannot be attenuated very quickly, which causes simulated skewness and kurtosis to exhibit larger swings when the horizon is initially lengthened up to some point. Nevertheless, the pattern of forward-looking skewness (or kurtosis) as a function of horizon clearly presents itself, and a spline smoothing can be applied to obtain the smoothed values for forward-looking skewness (or kurtosis). Figure 4 presents these smoothed higher moments along with the smoothed annualized forward-looking risk premiums for different horizons up to one year (252 trading days). These plots are presented for two specific time points: a relatively volatile time (September 2001) and a relatively quiet time (September 2003). The smoothing is done by a cubic polynomial spline with one knot at 125 trading days using the least-square estimation on 252 data points.

It is evident from the plots in Figure 4 that the forward-looking risk premium term structure can have various shapes. Its pattern has a great deal to do

[3] Obtained from http://wwwdev.nber.org/cycles/cyclesmain.html.

Figure 4 Term Structure of Forward-Looking Risk Premiums



Notes. This figure plots the smoothed forward-looking risk premium (FLRP) for the next one year (252 trading days) at two points in time, along with the physical forward-looking volatility, skewness (smoothed and unsmoothed), and kurtosis (smoothed and unsmoothed). The three plots at the top are for September 2001 when the market was volatile. The bottom three plots are for September 2003 when the market was relatively quiet. The left plots ((a1) and (b1)) are the term structures of forward-looking risk premiums and volatilities where the left axis is for the risk premium, and the right axis is for the volatility. The center plots ((a2) and (b2)) are the term structures of physical forward-looking skewness (smoothed and unsmoothed), whereas and the right plots ((a3) and (b3)) are the term structures of physical forward-looking kurtosis (smoothed and unsmoothed). The plotted forward-looking risk premiums are annualized.

with the term structures of physical forward-looking skewness and kurtosis. Basically, the physical return distribution becomes more negatively skewed and with fatter tails as the horizon is lengthened. This has the effect of increasing forward-looking risk premium. Once the horizon passes a certain point, the behavior of skewness and kurtosis begin to reverse. As expected, the volatility behaves in a typical mean-reverting manner. But the behaviors of skewness and kurtosis are more complex. This interesting feature of the GARCH model is not generally understood and rarely explored in the literature. In essence, cumulative return moves further away from normality due to stochastic mixture effect of the time-varying volatility. But once the horizon is long enough, the effect of the central limit theorem will kick in and move the cumulative return back toward normality.

4.4. Comparison with Other Measures of Risk Premium

The prior literature points out that expected risk premium and realized risk premium are fundamentally different concepts, and confusions arise from not properly distinguishing the two concepts (Elton 1999; Arnott and Bernstein 2002; Fernandez 2009a, b). Perhaps because of the lack of a better alternative, many measures of expected risk premium continue to rely on some form of ex post market risk premium as an input to obtain the estimate for expected risk premium. In this section, we present five measures of risk premium and compare them to forward-looking risk premium.

4.4.1. Historical Measures. We analyze the historical average of realized excess returns which is the most commonly used estimate of expected risk premium. In addition, two measures using average historical cross-sectionally estimated risk premium (the CAPM and the Fama and French (1996) three-factor model) are constructed.

Table 4 presents the S&P 500 risk premium based on historical average of daily excess returns over three years (panel A) and over five years (panel B). For each observation date, we estimated its historical risk premium by averaging daily excess returns over three or five years immediately before the observation

Table 4 S&P 500 Historical Risk Premium Estimates

	Average	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Panel A: Three-year historical risk premium (%)													
2001	−0.15	0.40	0.31	−0.17	−0.27	−0.04	−0.17	−0.16	0.04	−0.12	−0.50	−0.54	−0.61
2002	−1.14	−0.60	−0.70	−0.73	−0.84	−0.87	−1.11	−1.28	−1.41	−1.34	−1.63	−1.56	−1.56
2003	−1.28	−1.49	−1.63	−1.84	−1.60	−1.43	−1.32	−1.26	−1.32	−1.07	−0.95	−0.79	−0.66
2004	−0.12	−0.56	−0.25	−0.06	−0.30	−0.38	−0.22	−0.17	−0.06	0.25	0.09	0.06	0.15
2005	0.61	0.25	0.19	0.21	0.26	0.33	0.61	0.94	0.87	1.12	0.84	0.80	0.93
2006	0.77	1.13	1.18	1.07	0.92	0.80	0.60	0.61	0.57	0.59	0.60	0.64	0.54
2007	0.56	0.49	0.52	0.47	0.58	0.67	0.68	0.77	0.56	0.59	0.66	0.42	0.36
2008	−0.20	0.32	0.11	0.13	0.22	0.23	0.14	−0.13	−0.09	−0.11	−0.74	−1.16	−1.26
2009	−1.07	−1.23	−1.41	−1.76	−1.43	−1.09	−0.97	−1.02	−0.84	−0.81	−0.78	−0.80	−0.74
2010	−0.75	−0.69	−0.73	−0.65	−0.62	−0.74	−0.91	−0.85	−0.75	−0.91	−0.63		
Panel B: Five-year historical risk premium (%)													
2001	0.59	0.86	0.87	0.69	0.61	0.68	0.72	0.68	0.60	0.44	0.28	0.31	0.33
2002	−0.12	0.29	0.19	0.34	0.22	0.03	−0.11	−0.31	−0.36	−0.43	−0.51	−0.42	−0.38
2003	−0.48	−0.46	−0.66	−0.74	−0.67	−0.54	−0.52	−0.44	−0.29	−0.23	−0.35	−0.40	−0.45
2004	−0.45	−0.37	−0.37	−0.43	−0.47	−0.45	−0.50	−0.45	−0.51	−0.41	−0.47	−0.49	−0.46
2005	−0.37	−0.40	−0.36	−0.47	−0.45	−0.45	−0.44	−0.41	−0.42	−0.31	−0.33	−0.24	−0.14
2006	0.12	−0.15	0.01	0.12	0.01	0.00	0.01	0.04	0.18	0.32	0.30	0.28	0.29
2007	0.62	0.36	0.35	0.31	0.42	0.53	0.73	0.85	0.74	0.88	0.82	0.73	0.70
2008	0.25	0.77	0.69	0.58	0.52	0.48	0.42	0.30	0.25	0.19	−0.21	−0.43	−0.55
2009	−0.45	−0.58	−0.66	−0.85	−0.63	−0.51	−0.45	−0.40	−0.30	−0.25	−0.23	−0.27	−0.23
2010	−0.19	−0.17	−0.26	−0.14	−0.04	−0.14	−0.25	−0.31	−0.24	−0.24	−0.13		

Notes. This table reports historical monthly risk premium estimates using the S&P 500 index returns from January 2001 to October 2010. Panel A reports the historical risk premium estimated by averaging excess returns over three years. Panel B reports historical risk premium estimated by averaging excess returns over five years. Note that one month is 28 days from the observation date in the month to the subsequent option maturity date (every third Friday of the month). For example, January 2001 is from January 19, 2001, and February 16, 2001. The numbers reported are monthly percentage risk premiums.

date. Panel A shows that the three-year historical risk premiums are mostly negative for the periods 2001–2004 and 2008–2010. Similarly, panel B shows that the five-year historical risk premiums are mostly negative for the periods 2002–2005 and 2009–2010. In contrast, the forward-looking risk premiums reported in Table 3 are all positive.

Table 5 presents the S&P 500 historical risk premiums estimated from the CAPM and the Fama–French three-factor model.[4] We use the Fama and MacBeth (1973) approach to obtain monthly factor risk premiums. We also estimate the betas for the S&P 500 index (against the Center for Research in Security Prices (CRSP) value-weighted index and the factor portfolios when appropriate). The monthly risk premium for the S&P 500 index is equal to the S&P 500 index's beta times the corresponding factor risk premium. The risk premiums reported in Table 5 are the five-year averages of the monthly risk premiums for the S&P 500 index. Table 5 shows that the estimated risk premiums from the CAPM are mostly negative in 2002 and 2005 (panel A), whereas for the Fama–French three-factor model, the risk premiums are generally negative in 2001–2005 (panel B).

The correlation coefficients reported in Table 6 show that the five-year historical risk premium is not significantly correlated with the forward-looking risk premium. However, the three-year historical risk premium is negative correlated (−0.29) with the forward-looking risk premium. A shorter time span such as three years makes the historical risk premium more reflective of recent returns. In a down market, the historical risk premium becomes negative, but the bad news pushes up the forward-looking risk premium to result in a negative correlation. Neither of the risk premiums from the CAPM or the Fama–French three factor model is significantly correlated with the forward-looking risk premium. It is also evident from Figure 3 that these two risk premiums are hardly reflective of the NBER recessions or crises. The message from the empirical analysis is clear: forward-looking premium differs from historical risk premium both in concept and reality.

4.4.2. Fama and French (2002) Equity Premium. Fama and French (2002) estimated expected stock returns using the average dividend yield plus the average rate of capital gain estimated by either dividend or earnings growth rate. The S&P 500 risk premium based on earnings growth is estimated by

$$RXY_t = D_t/P_{t-1} + GY_t - F_t, \tag{8}$$

where D_t/P_{t-1} is the real dividend yield, GY_t is the estimate of real capital gains using realized earnings

[4] Monthly stock returns are obtained from CRSP. The Fama–French three factors are obtained from Kenneth R. French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

Table 5 S&P 500 CAPM and the Fama–French Three-Factor Model Risk Premium Estimates

	Average	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
							Panel A: CAPM (%)						
2001	0.11	0.04	0.35	0.19	0.10	0.15	0.14	0.14	0.09	0.01	−0.08	0.07	0.14
2002	0.00	0.16	0.14	0.08	0.14	0.07	−0.02	−0.09	−0.14	−0.14	−0.23	−0.12	0.08
2003	0.19	0.03	0.04	0.00	0.02	0.06	0.22	0.24	0.31	0.39	0.33	0.31	0.28
2004	0.13	0.25	0.26	0.27	0.23	0.18	0.18	0.14	0.07	0.05	0.02	0.00	−0.10
2005	−0.27	−0.19	−0.27	−0.42	−0.40	−0.35	−0.23	−0.35	−0.26	−0.31	−0.23	−0.19	−0.01
2006	0.02	0.07	−0.20	−0.05	0.07	0.00	−0.03	−0.02	0.01	0.10	0.19	0.08	0.01
2007	0.18	−0.01	0.00	0.09	0.06	0.15	0.20	0.28	0.31	0.30	0.37	0.30	0.10
2008	0.03	0.19	0.16	0.18	0.16	0.13	0.03	0.02	0.00	−0.01	−0.09	−0.20	−0.25
2009	0.26	−0.16	−0.18	−0.19	−0.03	0.26	0.31	0.33	0.46	0.56	0.61	0.57	0.57
2010	0.66	0.57	0.60	0.64	0.70	0.76	0.69	0.64	0.66	0.62	0.69		
							Panel B: Fama–French three-factor model (%)						
2001	−0.04	−0.05	0.13	0.03	−0.02	0.02	−0.02	−0.02	−0.06	−0.13	−0.22	−0.10	−0.03
2002	−0.14	0.00	−0.02	−0.07	−0.02	−0.07	−0.14	−0.20	−0.26	−0.24	−0.33	−0.24	−0.07
2003	−0.06	−0.14	−0.14	−0.18	−0.18	−0.13	−0.01	−0.01	0.02	0.08	0.00	−0.02	−0.01
2004	−0.07	−0.04	−0.01	0.02	−0.02	−0.07	−0.03	−0.06	−0.09	−0.11	−0.13	−0.15	−0.19
2005	−0.19	−0.23	−0.25	−0.23	−0.27	−0.27	−0.21	−0.22	−0.18	−0.18	−0.12	−0.12	−0.01
2006	0.02	0.00	−0.14	−0.04	0.04	−0.02	0.00	0.00	0.03	0.10	0.19	0.09	0.02
2007	0.14	−0.01	0.00	0.07	0.04	0.11	0.16	0.22	0.25	0.23	0.30	0.24	0.05
2008	0.07	0.14	0.14	0.17	0.17	0.14	0.07	0.08	0.08	0.07	0.01	−0.09	−0.13
2009	0.41	−0.05	−0.02	−0.05	0.13	0.42	0.49	0.50	0.61	0.71	0.77	0.72	0.74
2010	0.81	0.76	0.78	0.81	0.85	0.91	0.84	0.79	0.81	0.77	0.83		

Notes. This table reports the S&P 500 monthly risk premiums calculated from the CAPM (panel A) and the Fama–French three-factor model (panel B) by applying the Fama–MacBeth estimation method. Every risk premium is estimated using monthly returns in the five-year period immediately before the month in question. The numbers reported are monthly percentage risk premiums.

growth, and F_t is the real risk-free interest rate. When dividend growth is used, GY_t is replaced by dividend growth, GD_t, and RXY_t becomes RXD_t. Quarterly S&P 500 earnings and dividends are taken from Compustat.

Quarterly estimates for the risk premiums are reported in Table 7. During bad times, both realized earnings and dividends drop substantially. In these cases, the estimated capital gain yield is substantially negative, which in turns causes the risk premium to be negative. This effect is more pronounced for the estimates based on realized earnings rather than the ones based on dividends. This is true because realized earnings can be negative, but realized dividends are bounded below by zero.

Table 6 Correlation Between Risk Premium Estimates

	FLRP	Historical 5 yr	Historical 3 yr	CAPM	FF3
FLRP	1				
Historical 5 yr	−0.046	1			
Historical 3 yr	−0.293***	0.516***	1		
CAPM	−0.088	0.160*	−0.226**	1	
FF3	−0.019	0.096	−0.135	0.882***	1

Notes. This table presents the correlation coefficients between different measures of the S&P 500 index risk premium. FLRP is the forward-looking risk premium. Historical 5 yr and historical 3 yr are the historical risk premiums measured over five and three years, respectively. The risk premiums using the CAPM and the Fama–French three-factor model are denoted as CAPM and FF3, respectively.

*, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

We construct quarterly forward-looking risk premium at each quarter end to compare them with the Fama and French (2002) quarterly equity risk premium. The Spearman correlation coefficients are presented in panel C of Table 7. None of the risk

Table 7 S&P 500 Fama and French (2002) Equity Premium

	Average	Quarter 1	Quarter 2	Quarter 3	Quarter 4
		Panel A: Real risk premium using earnings growth ($RXY_t = D_t/P_{t-1} + GY_t - F_t$) (%)			
2001	−16.64	−10.30	−19.62	−23.60	−13.03
2002	3.04	−0.12	8.15	12.03	−7.89
2003	15.38	9.55	13.76	11.71	26.52
2004	4.93	7.06	8.12	3.07	1.47
2005	4.21	2.92	4.73	4.55	4.65
2006	3.19	3.21	1.85	4.63	3.05
2007	−5.42	1.21	1.38	−8.22	−16.03
2008	−25.54	−8.97	−14.25	−10.50	−68.42
2009	77.98	−50.44	7.03	60.85	294.47
		Panel B: Real risk premium using dividend growth ($RXD_t = D_t/P_{t-1} + GD_t - F_t$)			
	Average	Quarter 1	Quarter 2	Quarter 3	Quarter 4
2001	−0.61	−6.11	0.61	7.19	−4.15
2002	1.94	−5.25	9.72	−5.92	9.20
2003	5.00	−7.55	4.32	5.95	17.30
2004	1.75	−9.42	2.42	4.77	9.23
2005	3.18	0.07	0.11	0.84	11.68
2006	2.58	−3.23	1.04	0.34	12.15
2007	2.01	−5.72	1.78	2.12	9.86
2008	−1.37	−7.24	0.42	−0.93	2.28
2009	−4.68	−15.89	−7.94	−1.30	6.39

Table 7 (Continued)

Panel C: Correlations			
	$FLRP_{t-1}(t)$	RXD_t	RXY_t
$FLRP_{t-1}(t)$	1		
RXD_t	−0.207	1	
RXY_t	−0.266	0.099	1

Notes. This table reports the S&P 500 risk premiums using the Fama and French (2002) approach. The numbers reported are real equity premiums for the S&P 500 index being estimated with earnings or dividend growth. The results are quarterly from 2001 to 2009. RXY_t and RXD_t are the estimates for the S&P real risk premium based on earnings and dividend growth, respectively; D_t/P_{t-1} is real dividend yield; GY_t and GD_t are real earnings and dividend growth rates, respectively; F_t is the three-month Treasury-bill rate adjusted by inflation. The reported numbers are quarterly percentage risk premiums.

premiums are significantly correlated with the other, which implies that the risk premium based on the Fama and French (2002) method and the forward-looking risk premium are two distinctly different risk premium measures. The Fama and French (2002) risk premium is in essence an ex post fundamental measure of risk premium, whereas the forward-looking risk premium is an ex ante measure of risk compensation demanded by investors.

5. Asset Pricing Implications of Forward-Looking Risk Premium

5.1. Change in Forward-Looking Risk Premium and Excess Holding Period Return

A common approach to asset valuation is to set price equal to the present value of its expected future cash flows discounted by the cost of capital (the risk-free interest rate plus a risk premium). An increase in price is therefore related to either an increase of the expected future cash flows or a decrease in the risk premium (assuming the same risk-free rate). French et al. (1987) tested this idea indirectly by assuming that the change in risk premium is positively related to the unexpected change in stock market volatility. Without controlling for the cash flow effect, they found that an unexpected change in market volatility negatively affects the stock's holding period return.

In this section, we test the holding period return implication directly with respect to the change in the forward-looking risk premium while controlling for the change in the expected earnings. Our empirical model for this analysis is

$$R_{mt} - R_{ft} = \alpha + \beta_1 \Delta FLRP_t(\tau) + \beta_2 \Delta EPS_t^e + \epsilon_t, \quad (9)$$

where R_{mt} is the quarterly holding period return (from quarter $t-1$ to t) for the S&P 500 index, R_{ft} is the three-month Treasury bill return from quarter $t-1$ to t, and $\Delta FLRP_t(\tau)$ is defined as the change in the forward-looking risk premium from quarter $t-1$ to t (i.e., $FLRP_t(\tau) - FLRP_{t-1}(\tau)$). As in §4.4.2, quarterly forward-looking risk premiums are constructed for each quarter end. We define ΔEPS_t^e as the expected change in earnings per share for the S&P 500 index. Two proxies are used for EPS_t^e: the actual EPS data from Compustat and the analysts' forecast from the Institutional Brokers' Estimate System (I/B/E/S).

Our prediction on the regression coefficients is $\beta_1 < 0$ and $\beta_2 > 0$, reflecting the understanding that (1) an increase in expected risk premium decreases current stock price (holding period return), and (2) an increase in expected future earnings per share increases current stock price. The results reported in Table 8 are consistent with the predictions. Model (1) in Table 8 uses realized EPS as a proxy for expected EPS, whereas model (2) uses the mean of analysts' EPS forecasts for next quarter as expected EPS. Both tests give us consistent results on the implication for change in discount rate. The coefficients for $\Delta FLRP_t$ are significantly negative (−0.021 and −0.022). The coefficient for ΔEPS_t^e is significantly positive using realized earnings (0.015), but insignificant using analyst forecasts. The insignificant result using analyst forecasts may be caused by the high uncertainty during the recent crisis period. The constant term is insignificant in either case. The regression results confirm the theoretical prediction that an increase in discount rate (forward-looking risk premium) negatively

Table 8 S&P 500 Excess Return and Change in Forward-Looking Risk Premium

	$R_{mt} - R_{ft}$	
	(1)	(2)
Constant	−0.005	0.004
	(−0.39)	(0.30)
$\Delta FLRP_t$	−0.021***	−0.022**
	(−3.07)	(−2.73)
ΔEPS_t^e	0.015***	0.010
	(2.88)	(0.89)
Adj. R^2 (%)	34	21
No. of observations	38	34

Notes. This table reports the regression coefficients from the time-series regression of quarterly S&P 500 excess holding period return on change in forward-looking risk premium ($\Delta FLRP_t(\tau)$) and change in expected future earnings (ΔEPS_t^e). The sample is from 2001 to 2010. R_{mt} is the quarterly return for the S&P 500 index, and R_{ft} is the three-month Treasury-bill rate; $\Delta FLRP_t(\tau)$ is defined as the quarterly change in $FLRP_t(\tau)$, where $FLRP_t(\tau)$ is the forward-looking risk premium at the t for quarter $t+1$; and ΔEPS_t^e is the change in expected quarterly earnings. Model (1) uses realized earnings in the current quarter as a proxy for the next quarter's expected earnings. Model (2) uses analysts' forecasted next quarter earnings from I/B/E/S. The t-values are reported in parentheses.

** and *** denote statistical significance at the 5%, and 1% levels, respectively.

affects current stock price (holding period return) while controlling for expected change in future cash flows (change in EPS). A similar conclusion holds when the median of analysts' EPS forecasts instead of the mean is used.

5.2. Liquidity and the Forward-Looking Risk Premium

Amihud (2002) analyzed liquidity in stock returns using annual data from 1964 to 1996. He reported a positive relationship between expected market illiquidity and excess returns, and a negative relationship between unexpected illiquidity and contemporaneous excess returns. His finding serves as direct evidence that an illiquidity premium is reflected in excess returns.

In this section, we first show that the positive relationship between expected market illiquidity and excess return disappears from the monthly data between 2001 and 2010. Then, we demonstrate that illiquidity risk premium is still reflected in the forward-looking risk premium.

Similar to Amihud (2002), monthly market illiquidity, denoted by $MILLIQ_m$, is estimated as the average of $|R_{idm}|/VOLD_{idm}$ across all NYSE stocks and over all days in the month, where $|R_{idm}|$ is the absolute return of stock i on day d of month m, and $VOLD_{idm}$ is the corresponding daily trading volume. The monthly unexpected liquidity denoted by $MILLIQ_m^U$ is the residual from applying the AR(1) model to $MILLIQ_m$. Using monthly data from January 2001 to October 2010, we replicate the Amihud (2002) regression:

$$R_{Mt} - R_{ft} = g_0 + g_1 \ln(MILLIQ_{t-1}) + g_2 \ln(MILLIQ_t^U) + g_3 JANDUM_t + w_t, \tag{10}$$

where $R_{Mt} - R_{ft}$ is the excess stock return for month t, and $JANDUM_t$ is the dummy variable for January. The regression results are reported in Table 9. Amihud and Hurvich (2004) pointed out that the predictive regression such as Equation (10) produces biased estimates, but the bias can be corrected by adjusting the coefficient of the AR(1) model for illiquidity and also adjusting the standard error for g_1. We apply their adjustment and report the t-values computed from the adjusted standard errors.[5] Consistent with Amihud (2002), the monthly unexpected illiquidity negatively affects excess stock returns. Nevertheless, the lagged illiquidity does not positively affect excess returns. As a robustness check, we rerun the regression using the full monthly sample from 1964 to 2010, and the coefficient of $\ln(MILLIQ_{t-1})$ is 0.001 with a t-value equal to 0.90. The results indicate that there does not exist a positive relationship between excess return and lagged illiquidity (as a measure of expected illiquidity) over the past decade.

The lack of a significant relationship between expected illiquidity and excess return may be indicative of the poor quality of excess return as a proxy for risk premium. Therefore, we use forward-looking risk premium to test whether risk premium is related to expected illiquidity. To be compatible with the monthly horizon used in the liquidity measure, we recalculate the monthly forward-looking risk premium at each month end using the most recent γ available at the time.[6] We run the regression

$$FLRP_{t-1}(\tau) = \beta_0 + \beta_1 \ln(MILLIQ_{t-1}) + \beta_2 JANDUM_t + \epsilon_t, \tag{11}$$

where $FLRP_{t-1}(\tau)$ is the forward-looking risk premium for the next month (month t) at the end of month $t-1$. The Newey–West adjusted standard errors are used to calculate the t-values. The regression results reported in Table 9 show that from 2001 to 2010, the lagged illiquidity (as a measure of expected illiquidity) does not positively affect realized excess return, but it does positively influence forward-looking risk premium.

Prior literature suggests that illiquidity and volatility are positively related (Grossman and Miller 1988, Deuskar 2006, Kang and Yeo 2008). The positive correlation between illiquidity and forward-looking volatility is confirmed in Table 9, column (3). To check whether the positive relationship between forward-looking risk premium and illiquidity is merely a manifestation of the illiquidity–volatility relationship, we take the component, $(\gamma - 1)\sigma_{Pt}^2(\tau)$, out of the forward-looking risk premium estimate and rerun the regression in Equation (11). The results in Table 9, column (4), show that illiquidity still positively affects forward-looking risk premium after taking out the variance component. We further examine the relationship between illiquidity and forward-looking skewness (or kurtosis). The results in columns (5) and (6) of Table 9 show that illiquidity negatively (positively) affects skewness (kurtosis). The relationship between illiquidity and skewness (or kurtosis) are consistent with the intuition that when liquidity dries up, investors face higher uncertainty and may particularly worry about a large drop in stock price (negative skewness) or extreme moves in price (fat tails). In summary, illiquidity risk premium is reflected in forward-looking risk premium, and the relationship is not merely a manifestation of the illiquidity–volatility relationship.

[5] Similar results are obtained if we apply the adjustment procedure as in Amihud (2002).

[6] Using the original forward-looking risk premium for 28 calendar days (in §4.3) does not change our conclusion.

Table 9 Market Illiquidity and Risk Premium

	$R_{mt}-R_{ft}$ (1)	$FLRP_{t-1}(\tau)$ (2)	Variance$_{t-1}(\tau)$ (3)	$FLRP_{t-1}(\tau)$ (ex. var) (4)	Skewness$_{t-1}(\tau)$ (5)	Kurtosis$_{t-1}(\tau)$ (6)
Constant	−0.041	3.209**	0.051**	1.279**	−4.045***	22.113***
	(−0.63)	(2.58)	(2.27)	(2.08)	(−8.17)	(6.94)
Ln $MILLIQ_{t-1}$	−0.003	0.186**	0.003**	0.076**	−0.182***	0.971***
	(−0.73)	(2.49)	(2.17)	(2.05)	(−5.63)	(4.76)
Ln $MILLIQ_t^U$	−0.030***					
	(−4.23)					
$JANDUM_t$	−0.006	−0.0003	0.0004	−0.006	0.036	−0.258
	(−0.32)	(−0.005)	(0.28)	(−0.23)	(0.40)	(−0.48)
Adj. R^2 (%)	11	30	24	25	35	26
No. of observations	118	118	118	118	118	118

Notes. This table reports the coefficients for monthly regression from January 2001 to October 2010. $MILLIQ_{t-1}$ is the Amihud (2002) aggregate market illiquidity measure; $MILLIQ_t^U$ is the illiquidity shock in month t; $JANDUM_t$ is the dummy variable for January; R_{mt} is the monthly return for the S&P 500 index; R_{ft} is the 30-day Treasury-bill rate; $FLRP_{t-1}(\tau)$ is the forward-looking risk premium for next month at month $t-1$; Variance$_{t-1}(\tau)$ is the forward-looking variance for next month at time $t-1$, which is computed from the GARCH model; $FLRP_{t-1}$(ex. var) is the forward-looking risk premium minus $(\gamma-1)\sigma^2 P_t(\tau)$; and Skewness$_{t-1}(\tau)$ and Kurtosis$_{t-1}(\tau)$ are the forward-looking skewness and kurtosis for next month available at month $t-1$, respectively. The numbers in parentheses are t-values. For model (1), they are based on the adjusted standard errors as in Amihud and Hurvich (2004); for models (2)–(6), they are based on the Newey–West adjusted standard errors.

** and *** denote statistical significance at the 5% and 1% levels, respectively.

6. Conclusion

We propose a practical model for estimating forward-looking risk premium. First, a forward-looking risk premium formula is developed. Then, the components of this formula—physical forward-looking volatility, skewness, and kurtosis, and investors' risk aversion—are estimated. The GARCH model is used to deduce forward-looking physical volatility, skewness, and kurtosis needed for the implementation. Investors' risk aversion is estimated by a volatility spread formula that links the gap between option implied risk-neutral volatility and forward-looking physical volatility to forward-looking skewness and kurtosis. Option implied risk-neutral volatilities are naturally forward looking, and have been shown in a different context to deliver superior performance in forward-looking asset allocations (see Kostakis et al. 2011).

Our empirical analysis uses the S&P 500 index return and option data. The estimates for investors' risk aversion are sensible with values in the range from 1.3 to 6.0. The estimated forward-looking risk premium are consistently positive. In sharp contrast, other commonly used risk premium measures, such as historical average excess return, estimates using the CAPM and the Fama and French (1996) three-factor model, and the Fama and French (2002) fundamental estimate, are often empirically negative. Obviously, negative risk premiums are theoretically questionable, intuitively unappealing, and practically unusable. Furthermore, forward-looking risk premiums are higher during crisis periods and lower during boom times, exhibiting a desirable feature that is in keeping with economic intuition.

Two asset pricing implications related to forward-looking risk premium are also examined in this paper. The change in forward-looking risk premium negatively affects current stock price, and expected illiquidity positively affects forward-looking risk premium. Both are consistent with financial theory and economic intuition. Given the prominent role played by risk premium in finance, our proposed estimation method for forward-looking risk premium can have wide-ranging implications in financial research and practice.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2013.1758.

Acknowledgments

The authors thank Yakov Amihud, Bruno Solnik, Liuren Wu, Raymond Kan, and seminar participants at ESSEC (École Supérieure des Sciences Économiques et Commerciales), University of Toronto, New York University, and University of Piraeus for their helpful comments. The inputs received at the Workshop on Econometric and Financial Studies after Crisis held at Academia Sinica in June 2010 are also gratefully acknowledged.

Appendix A. Derivation of Equation (1)

First expand

$$e^{R_t(\tau)-\mu_{Qt}(\tau)} = 1 + R_t(\tau) - \mu_{Qt}(\tau) + \frac{(R_t(\tau)-\mu_{Qt}(\tau))^2}{2} + \frac{(R_t(\tau)-\mu_{Qt}(\tau))^3}{6} + \frac{(R_t(\tau)-\mu_{Qt}(\tau))^4}{24} + O[(R_t(\tau)-\mu_{Qt}(\tau))^5].$$

Taking expectation with respect to measure Q and recognizing $E_t^Q(e^{R_t(\tau)}) = e^{r_t(\tau)-\delta_t(\tau)}$ in turn gives rise to

$$e^{r_t(\tau)-\delta_t(\tau)-\mu_{Qt}(\tau)} = 1 + \frac{\sigma_{Qt}^2(\tau)}{2} + \frac{\theta_{Qt}(\tau)\sigma_{Qt}^3(\tau)}{6} + \frac{\kappa_{Qt}(\tau)\sigma_{Qt}^4(\tau)}{24} + o[\sigma_{Qt}^4(\tau)].$$

Therefore, the equilibrium risk-free interest rate can be written as

$$\begin{aligned} r_t(\tau) &= \delta_t(\tau)+\mu_{Qt}(\tau)+\ln\big(1+\tfrac{1}{2}\sigma_{Qt}^2(\tau)+\tfrac{1}{6}\theta_{Qt}(\tau)\sigma_{Qt}^3(\tau) \\ &\qquad +\tfrac{1}{24}\kappa_{Qt}(\tau)\sigma_{Qt}^4(\tau)+o[\sigma_{Qt}^4(\tau)]\big) \\ &= \delta_t(\tau)+\mu_{Qt}(\tau)+\tfrac{1}{2}\sigma_{Qt}^2(\tau)+\tfrac{1}{6}\sigma_{Qt}^3(\tau)\theta_{Qt}(\tau) \\ &\qquad +\tfrac{1}{24}\sigma_{Qt}^4(\tau)[\kappa_{Qt}(\tau)-3]+o[\sigma_{Qt}^4(\tau)]. \end{aligned}$$

The second equality comes from a second-order Taylor expansion of the logarithmic function around 1. The term $o[\sigma_{Qt}^4(\tau)]$ can be ignored because the typical estimate suggests that $\sigma_{Qt}(\tau)$ on an annualized basis is less than 1. Applying to the monthly or quarterly cumulative return, it would be even smaller.

Appendix B. Derivation of Equation (3)

Our derivations for the following equations are essentially the same as those of Bakshi and Madan (2006) except for two subtle points. First, our approximation ignores terms with an order higher than $\sigma_{Qt}^4(\tau)$, whereas their approach drops terms with an order of γ^3 or higher. Since the estimated risk aversion coefficient is typically large (Bakshi and Madan's (2006) own estimate for γ is around 17), it is questionable to ignore terms in the order of γ^3 or higher. However, the volatility for an equity index such as S&P 500 is typically below 20% per annum, which makes its fifth or higher powers indeed negligible. Second, Bakshi and Madan (2006) assumed that the physical first moment equals zero, which is actually not needed.

Instead of dealing with the moment generating function of $R_t(\tau)$ directly, it is analytically more convenient to compute that of $R_t^*(\tau)\equiv R_t(\tau)-\mu_{Pt}(\tau)$. We have

$$\begin{aligned} \mathscr{C}_t(\lambda)\equiv E_t^P(e^{\lambda R_t^*(\tau)}) &= 1+\frac{\lambda^2}{2}\sigma_{Pt}^2(\tau)+\frac{\lambda^3}{6}\theta_{Pt}(\tau)\sigma_{Pt}^3(\tau) \\ &\qquad +\frac{\lambda^4}{24}\kappa_{Pt}(\tau)\sigma_{Pt}^4(\tau)+o[\lambda^4\sigma_{Pt}^4(\tau)]. \end{aligned}$$

This in turn allows one to express the moment generating function of $R_t^*(\tau)$ under measure Q using $\mathscr{C}_t(\cdot)$:

$$\begin{aligned} E_t^Q(e^{\lambda R_t^*(\tau)}) &= \frac{E_t^P(e^{\lambda R_t^*(\tau)}e^{-\gamma R_t(\tau)})}{E_t^P(e^{-\gamma R_t(\tau)})}=\frac{E_t^P(e^{(\lambda-\gamma)R_t^*(\tau)}e^{-\gamma\mu_{Pt}})}{E_t^P(e^{-\gamma R_t^*(\tau)}e^{-\gamma\mu_{Pt}})} \\ &= \frac{\mathscr{C}_t(\lambda-\gamma)}{\mathscr{C}_t(-\gamma)}. \end{aligned}$$

Thus,

$$\begin{aligned} &E_t^Q(R_t^*(\tau)) \\ &\quad=\frac{\mathscr{C}_t'(\lambda-\gamma)|_{\lambda=0}}{\mathscr{C}_t(-\gamma)} \\ &\quad=\left(1+\frac{\gamma^2}{2}\sigma_{Pt}^2(\tau)+O[\sigma_{Pt}^3(\tau)]\right)^{-1} \\ &\qquad\cdot\left(-\gamma\sigma_{Pt}^2(\tau)+\frac{\gamma^2}{2}\theta_{Pt}(\tau)\sigma_{Pt}^3(\tau)-\frac{\gamma^3}{6}\kappa_{Pt}(\tau)\sigma_{Pt}^4(\tau)+o[\sigma_{Pt}^4(\tau)]\right) \\ &\quad=-\gamma\sigma_{Pt}^2(\tau)+\frac{\gamma^2}{2}\sigma_{Pt}^3(\tau)\theta_{Pt}(\tau)-\frac{\gamma^3}{6}\sigma_{Pt}^4(\tau)[\kappa_{Pt}(\tau)-3] \\ &\qquad+o[\sigma_{Pt}^4(\tau)]; \end{aligned}$$

$$\begin{aligned} &E_t^Q[(R_t^*(\tau))^2] \\ &\quad=\frac{\mathscr{C}_t''(\lambda-\gamma)|_{\lambda=0}}{\mathscr{C}_t(-\gamma)} \\ &\quad=\left(1+\frac{\gamma^2}{2}\sigma_{Pt}^2(\tau)+O[\sigma_{Pt}^3(\tau)]\right)^{-1} \\ &\qquad\cdot\left(\sigma_{Pt}^2(\tau)-\gamma\theta_{Pt}(\tau)\sigma_{Pt}^3(\tau)+\frac{\gamma^2}{2}\kappa_{Pt}(\tau)\sigma_{Pt}^4(\tau)+o[\sigma_{Pt}^4(\tau)]\right) \\ &\quad=\sigma_{Pt}^2(\tau)-\gamma\sigma_{Pt}^3(\tau)\theta_{Pt}(\tau)+\frac{\gamma^2}{2}\sigma_{Pt}^4(\tau)[\kappa_{Pt}(\tau)-1]+o[\sigma_{Pt}^4(\tau)]; \end{aligned}$$

$$\begin{aligned} &E_t^Q[(R_t^*(\tau))^3] \\ &\quad=\frac{\mathscr{C}_t'''(\lambda-\gamma)|_{\lambda=0}}{\mathscr{C}_t(-\gamma)} \\ &\quad=\left(1+\frac{\gamma^2}{2}\sigma_{Pt}^2(\tau)+O[\sigma_{Pt}^3(\tau)]\right)^{-1} \\ &\qquad\cdot\left(\theta_{Pt}(\tau)\sigma_{Pt}^3(\tau)-\gamma\kappa_{Pt}(\tau)\sigma_{Pt}^4(\tau)+o[\sigma_{Pt}^4(\tau)]\right) \\ &\quad=\theta_{Pt}(\tau)\sigma_{Pt}^3(\tau)-\gamma\kappa_{Pt}(\tau)\sigma_{Pt}^4(\tau)+o[\sigma_{Pt}^4(\tau)]; \end{aligned}$$

$$\begin{aligned} &E_t^Q[(R_t^*(\tau))^4] \\ &\quad=\frac{\mathscr{C}_t''''(\lambda-\gamma)|_{\lambda=0}}{\mathscr{C}_t(-\gamma)} \\ &\quad=\left(1+\frac{\gamma^2}{2}\sigma_{Pt}^2(\tau)+O[\sigma_{Pt}^3(\tau)]\right)^{-1}\left(\kappa_{Pt}(\tau)\sigma_{Pt}^4(\tau)+o[\sigma_{Pt}^4(\tau)]\right) \\ &\quad=\kappa_{Pt}(\tau)\sigma_{Pt}^4(\tau)+o[\sigma_{Pt}^4(\tau)]. \end{aligned}$$

The above results immediately give rise to the risk-neutral expected return, variance, skewness, and kurtosis as follows:

$$\begin{aligned} \mu_{Qt}(\tau) &= E_t^Q(R_t^*(\tau))+\mu_{Pt}(\tau) \\ &= \mu_{Pt}(\tau)-\gamma\sigma_{Pt}^2(\tau)+\frac{\gamma^2}{2}\sigma_{Pt}^3(\tau)\theta_{Pt}(\tau) \\ &\qquad -\frac{\gamma^3}{6}\sigma_{Pt}^4(\tau)[\kappa_{Pt}(\tau)-3]+o[\sigma_{Pt}^4(\tau)]; \end{aligned} \tag{B1}$$

$$\begin{aligned} \sigma_{Qt}^2(\tau) &= E_t^Q[(R_t^*(\tau))^2]-[\mu_{Qt}(\tau)-\mu_{Pt}(\tau)]^2 \\ &= \sigma_{Pt}^2(\tau)-\gamma\sigma_{Pt}^3(\tau)\theta_{Pt}(\tau) \\ &\qquad +\frac{\gamma^2}{2}\sigma_{Pt}^4[\kappa_{Pt}(\tau)-3]+o[\sigma_{Pt}^4(\tau)]; \end{aligned} \tag{B2}$$

$$\begin{aligned} \theta_{Qt}(\tau)\sigma_{Qt}^3(\tau) &= E_t^Q[(R_t^*(\tau))^3]-3\sigma_{Qt}^2(\tau)[\mu_{Qt}(\tau)-\mu_{Pt}(\tau)] \\ &\qquad -[\mu_{Qt}(\tau)-\mu_{Pt}(\tau)]^3 \\ &= \sigma_{Pt}^3(\tau)\theta_{Pt}(\tau)-\gamma\sigma_{Pt}^4[\kappa_{Pt}(\tau)-3] \\ &\qquad +o[\sigma_{Pt}^4(\tau)]; \end{aligned} \tag{B3}$$

$$\begin{aligned} \kappa_{Qt}(\tau)\sigma_{Qt}^4(\tau) &= E_t^Q[(R_t^*(\tau))^4]-4\theta_{Qt}\sigma_{Qt}^3(\tau)[\mu_{Qt}(\tau)-\mu_{Pt}(\tau)] \\ &\qquad -6\sigma_{Qt}^2(\tau)[\mu_{Qt}(\tau)-\mu_{Pt}(\tau)]^2 \\ &\qquad -[\mu_{Qt}(\tau)-\mu_{Pt}(\tau)]^4 \\ &= \kappa_{Pt}(\tau)\sigma_{Pt}^4(\tau)+o[\sigma_{Pt}^4(\tau)]. \end{aligned} \tag{B4}$$

Note that Equation (B2) can be rewritten as the volatility spread in Equation (2).

Substituting Equations (B1)–(B4) into the risk-free interest rate Equation (1), we can express the equilibrium risk-free interest rate in terms of physical moments:

$$r_t(\tau) \approx \delta_t(\tau) + \mu_{Pt}(\tau) - \left(\gamma - \frac{1}{2}\right)\sigma^2_{Pt}(\tau) + \frac{3\gamma^2 - 3\gamma + 1}{6}\sigma^3_{Pt}(\tau) \cdot \theta_{Pt}(\tau) - \frac{4\gamma^3 - 6\gamma^2 + 4\gamma - 1}{24}\sigma^4_{Pt}(\tau)[\kappa_{Pt}(\tau) - 3].$$

Consequently, the equity risk premium can be expressed as

$$\begin{aligned}\mu_{Pt}(\tau) + \delta_t(\tau) - r_t(\tau) &\approx \left(\gamma - \frac{1}{2}\right)\sigma^2_{Pt}(\tau) - \frac{3\gamma^2 - 3\gamma + 1}{6}\sigma^3_{Pt}(\tau)\theta_{Pt}(\tau) \\ &\quad + \frac{4\gamma^3 - 6\gamma^2 + 4\gamma - 1}{24}\sigma^4_{Pt}(\tau)[\kappa_{Pt}(\tau) - 3].\end{aligned}$$

Remark B.1. Ignoring terms with an order of γ^3 or higher as in Bakshi and Madan (2006) does not affect the volatility spread equation, but it alters the equation for the risk-neutral first moment. Were γ small, the term $(\gamma^3/6)\sigma^4_{Pt}(\tau)[\kappa_{Pt}(\tau) - 3]$ in (B1) could be grouped into $O(\gamma^3)$ and ignored. Interestingly, all $O(\gamma^3)$ except for $(\gamma^3/6)\sigma^4_{Pt}(\tau)[\kappa_{Pt}(\tau) - 3]$ are of an order higher than $\sigma^4_{Pt}(\tau)$, and thus can also be ignored under our approach.

Appendix C. Risk-Neutral Variance as an Option Portfolio

A continuous twice-differentiable payoff function $f(S)$ can be represented in a integral form as follows (see Carr and Madan 2001, Bakshi and Madan 2006):

$$f(S) = f(\bar{S}) + (S - \bar{S})f_S(\bar{S}) + \int_{\bar{S}}^{\infty} f_{SS}(k)(S-k)^+\,dk + \int_0^{\bar{S}} f_{SS}(k)(k-S)^+\,dk,$$

where $f_S(\cdot)$ and $f_{SS}(\cdot)$ are the first and second derivatives, respectively. Expand $f(S_{t+\tau})$ around K_t, a point close to the forward price $F_t(\tau) = S_t e^{(r_t(\tau)-\delta_t(\tau))\tau}$ and apply the risk-neutral measure at time t to yield

$$\begin{aligned}E_t^Q[f(S_{t+\tau})] &= f(K_t) + (E_t^Q[S_{t+\tau}] - K_t)f_S(K_t) \\ &\quad + \int_{K_t}^{\infty} f_{SS}(k)E_t^Q[(S_{t+\tau}-k)^+]\,dk \\ &\quad + \int_0^{K_t} f_{SS}(k)E_t^Q[(k-S_{t+\tau})^+]\,dk \\ &= f(K_t) + [S_t(\tau)e^{(r_t(\tau)-\delta_t(\tau))\tau} - K_t]f_S(K_t) \\ &\quad + e^{r_t(\tau)\tau}\int_{K_t}^{\infty} f_{SS}(k)C(k; S_t, \tau)\,dk \\ &\quad + e^{r_t(\tau)\tau}\int_0^{K_t} f_{SS}(k)P(k; S_t, \tau)\,dk,\end{aligned}$$

where $C(k; S_t, \tau)$ and $P(k; S_t, \tau)$ are the time-t European call and put option prices, respectively, with strike price k and maturity τ.

Define $w_1(k) \equiv 1/k^2$, $w_2(k; S_t) \equiv (2[1 - \ln(k/S_t)])/k^2$. The following expressions can easily be derived (Bakshi et al. 2003):

$$\begin{aligned}E_t^Q[R_t(\tau)] &= \ln\left(\frac{K_t}{S_t}\right) + \frac{F_t(\tau) - K_t}{K_t} \\ &\quad - e^{r_t(\tau)\tau}\int_{K_t}^{\infty} w_1(k)C(k; S_t, \tau)\,dk \\ &\quad - e^{r_t(\tau)\tau}\int_0^{K_t} w_1(k)P(k; S_t, \tau)\,dk;\end{aligned}$$

$$\begin{aligned}E_t^Q[R_t^2(\tau)] &= \left[\ln\left(\frac{K_t}{S_t}\right)\right]^2 + 2\left(\frac{F_t(\tau) - K_t}{K_t}\right)\ln\left(\frac{K_t}{S_t}\right) \\ &\quad + e^{r_t(\tau)\tau}\int_{K_t}^{\infty} w_2(k; S_t)C(k; S_t, \tau)\,dk \\ &\quad + e^{r_t(\tau)\tau}\int_0^{K_t} w_2(k; S_t)P(k; S_t, \tau)\,dk.\end{aligned}$$

They then give rise to the risk-neutral variance as a portfolio of options by noting that $\sigma^2_{Qt}(\tau) = E_t^Q[R_t^2(\tau)] - (E_t^Q[R_t(\tau)])^2$.

Appendix D. Cumulative Return Moments Under NGARCH(1,1)

Under the NGARCH(1,1) model in Equations (5)–(6), the conditional variance of the cumulative return over τ-days can be derived as follows:

$$\begin{aligned}\sigma^2_{Pt}(\tau) &= E_t^P\left[\ln\left(\frac{S_{t+\tau}}{S_t}\right) - E_t^P\left(\ln\left(\frac{S_{t+\tau}}{S_t}\right)\right)\right]^2 \\ &= E_t^P\left[\sum_{i=1}^{\tau}\ln\left(\frac{S_{t+i}}{S_{t+i-1}}\right) - E_t^P\left(\sum_{i=1}^{\tau}\ln\left(\frac{S_{t+i}}{S_{t+i-1}}\right)\right)\right]^2 \\ &= E_t^P\left[\mu\tau + \sum_{i=1}^{\tau}\sigma_{t+i}\varepsilon_{t+i} - E_t^P\left(\mu\tau + \sum_{i=1}^{\tau}\sigma_{t+i}\varepsilon_{t+i}\right)\right]^2 \\ &= E_t^P\left[\sum_{i=1}^{\tau}\sigma_{t+i}\varepsilon_{t+i}\right]^2 \\ &= \sum_{i=1}^{\tau}E_t^P(\sigma^2_{t+i}).\end{aligned}$$

It is clear that $\sigma^2_{Pt}(1) = \sigma^2_{t+1}$. When $\tau \geq 2$, then,

$$\begin{aligned}\sigma^2_{Pt}(\tau) &= \sigma^2_{t+1} + \sum_{i=1}^{\tau-1}E_t^P[\beta_0 + \beta_1\sigma^2_{t+i} + \beta_2\sigma^2_{t+i}(\varepsilon_{t+i} - \eta)^2] \\ &= \sigma^2_{t+1} + \sum_{i=1}^{\tau-1}\{\beta_0 + [\beta_1 + \beta_2(1+\eta^2)]E_t^P(\sigma^2_{t+i})\}. \qquad \text{(D1)}\end{aligned}$$

Let $\lambda = \beta_1 + \beta_2(1+\eta^2)$. Recursively apply conditional expectations to Equation (6) to yield

$$E_t^P(\sigma^2_{t+i}) = \frac{\beta_0}{1-\lambda} + \lambda^{i-1}\left(\sigma^2_{t+1} - \frac{\beta_0}{1-\lambda}\right).$$

Plugging the above result into Equation (D1) gives rise to

$$\begin{aligned}\sigma^2_{Pt}(\tau) &= \sigma^2_{t+1} + \sum_{i=1}^{\tau-1}[\beta_0 + \lambda E_t^P(\sigma^2_{t+i})] \\ &= \sigma^2_{t+1} + \sum_{i=1}^{\tau-1}\left[\frac{\beta_0}{1-\lambda} + \lambda^i\left(\sigma^2_{t+1} - \frac{\beta_0}{1-\lambda}\right)\right] \\ &= \left(\sum_{i=0}^{\tau-1}\lambda^i\right)\sigma^2_{t+1} + \left(\sum_{i=1}^{\tau-1}(1-\lambda^i)\right)\frac{\beta_0}{1-\lambda} \\ &= \frac{1-\lambda^\tau}{1-\lambda}\sigma^2_{t+1} + \frac{(\tau-1)\beta_0}{1-\lambda} - \frac{\lambda(1-\lambda^{\tau-1})\beta_0}{(1-\lambda)^2}.\end{aligned}$$

Note that the above formula also applies to the case of $\tau = 1$.

References

Amihud Y (2002) Illiquidity and stock returns: Cross-section and time-series effects. *J. Financial Markets* 5:21–56.

Amihud Y, Hurvich CM (2004) Predictive regressions: A reduced-bias estimation method. *J. Financial Quant. Anal.* 39:813–841.

Arnott RD, Bernstein PL (2002) What risk premium is "normal"? *Financial Analysts J.* 58:64–85.

Bakshi G, Madan DB (2006) A theory of volatility spread. *Management Sci.* 52:1945–1956.

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

BBC News (2008) Global shares tumble on US fears. (January 22), http://news.bbc.co.uk/2/hi/business/7199552.stm.

Bliss RR, Panigirtzoglou N (2004) Option-implied risk aversion estimates. *J. Finance* 59:407–446.

Britten-Jones M, Neuberger A (2000) Option price, implied price processes, and stochastic volatility. *J. Finance* 55:839–866.

Campbell JY, Cochrane JH (2000) Explaining the poor performance of consumption-based asset pricing models. *J. Finance* 55:2863–2878.

Carr P, Madan DB (2001) Optimal positioning in derivative securities. *Quant. Finance* 1:19–37.

Christensen BJ, Prabhala NR (1998) The relation between implied and realized volatility. *J. Financial Econom.* 50:125–150.

Damodaran A (2008) Equity risk premiums (ERP): Determinants, estimation and implications. Working paper, Stern School of Business, New York University, New York.

Deuskar P (2006) Extrapolative expectations: Implications for volatility and liquidity. AFA 2007 Chicago Meetings Paper, http://ssrn.com/abstract=891539.

Duan JC (1997) Augmented GARCH(p, q) process and its diffusion limit. *J. Econometrics* 79:97–127.

Economist, The (2002) WorldCom's bankruptcy mess. (July 22), http://www.economist.com/node/1245862.

Elton EJ (1999) Expected return, realized return, and asset pricing tests. *J. Finance* 54:1199–1220.

Engle R, Ng V (1993) Measuring and testing of the impact of news on volatility. *J. Finance* 48:1749–1778.

Fama EF, French KR (1996) Multifactor explanations of asset pricing anomalies. *J. Finance* 51:55–84.

Fama EF, French KR (2002) The equity premium. *J. Finance* 57:637–659.

Fama EF, MacBeth JD (1973) Risk, return and the equilibrium: Empirical tests. *J. Political Econom.* 82:607–636.

Fernandez P (2009a) Market risk premium used in 2008 by professors: A survey with 1,400 answers. Working paper, IESE Business School, Barcelona, Spain.

Fernandez P (2009b) The equity premium in 150 textbooks. Working paper, IESE Business School, Barcelona, Spain.

French KR, Schwert GW, Stambaugh RF (1987) Expected stock returns and volatility. *J. Financial Econom.* 19:3–29.

Graham JR, Harvey CR (2007) The equity risk premium in January 2007: Evidence from the global CFO outlook survey. *ICFAI J. Financial Risk Management* 4:46–61.

Grossman SJ, Miller MH (1988) Liquidity and market structure. *J. Finance* 43:617–637.

Grossman SJ, Shiller RJ (1981) The determinants of the variability of stock market prices. *Amer. Econom. Rev.* 71: 222–227.

Hansen LP, Singleton KJ (1983) Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *J. Political Econom.* 91:249–268.

Harvey CR, Siddique A (2000) Conditional skewness in asset pricing tests. *J. Finance* 55:1263–1295.

Jackwerth JC, Rubinstein M (1996) Recovering probability distributions from option prices. *J. Finance* 51:1611–1631.

Jiang GJ, Tian YS (2005) The model-free implied volatility and its information content. *Rev. Financial Stud.* 18:1305–1342.

Kang W, Yeo WY (2008) Liquidity beyond the best quote: A study of the NYSE limit order book. WFA 2008 Conference Paper, http://ssrn.com/abstract=888650.

Kostakis A, Panigirtzoglou N, Skiadopoulos G (2011) Market timing with option-implied distributions: A forward-looking approach. *Management Sci.* 57:1231–1249.

Kraus A, Litzenberger RH (1976) Skewness preference and the valuation of risk assets. *J. Finance* 31:1085–1100.

Malik S, Pitt MK (2011) Particle filters for continuous likelihood evaluation and maximisation. *J. Econometrics* 165: 190–209.

Merton RC (1980) On estimating the expected return on the market: An exploratory investigation. *J. Financial Econom.* 8: 323–361.

Newey K, West KD (1987) A simple, positive semi-definite, heteroskadasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–708.

Santa-Clara P, Yan S (2010) Crashes, volatility and the equity premium: Lessons from S&P 500 options. *Rev. Econom. Statist.* 92:435–451.

Welch I (2000) Views of financial economists on the equity premium and on professional controversies. *J. Bus.* 73:501–537.