

## **Finance and Economics Discussion Series**

Federal Reserve Board, Washington, D.C.

ISSN 1936-2854 (Print)

ISSN 2767-3898 (Online)

# **The Pricing Kernel in Options**

**Steven Heston, Kris Jacobs, Hyung Joo Kim**

**2023-053**

Please cite this paper as:

Heston, Steven, Kris Jacobs, and Hyung Joo Kim (2023). "The Pricing Kernel in Options," Finance and Economics Discussion Series 2023-053. Washington: Board of Governors of the Federal Reserve System, <https://doi.org/10.17016/FEDS.2023.053>.

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

# The Pricing Kernel in Options\*

Steven Heston

Kris Jacobs

Hyung Joo Kim

University of Maryland

University of Houston

Federal Reserve Board

April 7, 2023

## Abstract

The empirical option valuation literature specifies the pricing kernel through the price of risk, or defines it implicitly as the ratio of risk-neutral and physical probabilities. Instead, we extend the economically appealing Rubinstein-Brennan kernels to a dynamic framework that allows path- and volatility-dependence. Because of low statistical power, kernels with different economic properties can produce similar overall option fit, even when they imply cross-sectional pricing anomalies and implausible risk premiums. Imposing parsimonious economic restrictions such as monotonicity and path-independence (recovery theory) achieves good option fit and reasonable estimates of equity and variance risk premiums, while resolving pricing kernel anomalies.

---

\*Heston: [sheston@umd.edu](mailto:sheston@umd.edu); Jacobs: [kjacobs@bauer.uh.edu](mailto:kjacobs@bauer.uh.edu); Kim: [hyungjoo.kim@frb.gov](mailto:hyungjoo.kim@frb.gov). We would like to thank Caio Almeida, David Bates, Hitesh Doshi, Bjørn Eraker, Xiaohui Gao, Stefano Giglio, Massimo Guidolin, Alex Kostakis, Paola Pederzoli, Jean-Paul Renne, seminar participants at the 2023 AFA Conference, the 2022 Finance Down Under Conference, the 2022 SoFiE Conference, the 2022 FMA Conference on Derivatives and Volatility, K.U. Leuven, Syracuse University, the Universities of Houston and Liverpool, and especially our discussants Gurdip Bakshi, Mikhail Chernov, and Jeroen Dalderop for helpful comments. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by the Federal Reserve Board or other members of its staff. Co-author Hyung Joo Kim worked on this project prior to employment at the Federal Reserve Board, while a Ph.D. candidate at the University of Houston. Data sources were obtained under purview of University of Houston licenses.

# 1 Introduction

The pricing kernel is a critical concept in asset pricing, because it determines risk premia on all securities. One approach to study the properties of the pricing kernel specifies its relation to aggregate consumption and estimates the resulting model using consumption data and returns on various assets.<sup>1</sup> Alternatively, building on the insights of Breeden and Litzenberger (1978), an extensive literature estimates the pricing kernel using index returns and index option prices.<sup>2</sup> Index options are interesting from an empirical perspective because they identify the pricing kernel under the assumption that the equity index level is equal to aggregate wealth. However, this literature has given rise to puzzling non-monotonic (U-shaped) estimates of the pricing kernel.<sup>3</sup>

This paper complements these two approaches. We specify economically intuitive pricing kernels in parametric option pricing models and estimate them using index returns and index option prices. Our approach differs from the existing parametric literature, which typically specifies prices of risk, and thereby defines the pricing kernel implicitly as the ratio of risk-neutral and physical probabilities.<sup>4</sup> Instead, we propose a class of economically motivated pricing kernels that extend the power kernels in Rubinstein (1976) and Brennan (1979) to be functions of the paths of latent variance  $v(t)$  and the index level  $S(t)$ . These kernels also nest path-independent kernels (Ross, 2015) as special cases, and are consistent with the conventional assumption of affine dynamics under the physical and risk-neutral measure in the square root stochastic volatility model (Heston, 1993).<sup>5</sup> By including separate components for volatility and stock risk, these kernels allow us to

---

<sup>1</sup>See, among others, Hansen and Singleton (1982), Mehra and Prescott (1985), Hansen and Jagannathan (1991), Campbell and Cochrane (1999), Bansal and Yaron (2004), Gabaix (2012) and Wachter (2013) for important contributions to this literature.

<sup>2</sup>See for instance Aït-Sahalia and Lo (1998), Aït-Sahalia and Lo (2000), Jackwerth and Rubinstein (1996), Jackwerth (2000) and Rosenberg and Engle (2002).

<sup>3</sup>For evidence on U-shaped pricing kernels, see for instance Jackwerth (2000), Aït-Sahalia and Lo (2000), Rosenberg and Engle (2002), Bakshi, Madan, and Panayotov (2010), Chabi-Yo (2012), Christoffersen, Heston, and Jacobs (2013), Song and Xiu (2016), and Cuesdeanu and Jackwerth (2018). Linn, Shive, and Shumway (2018) and Barone-Adesi, Fusari, Mira, and Sala (2020) on the other hand argue that the pricing kernel is well-behaved.

<sup>4</sup>For examples of this approach, see the seminal papers in this literature by Chernov and Ghysels (2000), Pan (2002), and Eraker (2004).

<sup>5</sup>For simplicity, we use the simplest possible option pricing model with a stochastic volatility factor, but our approach can be easily generalized to more complex models.

examine distinct origins of the equity and variance risk premiums.

We start our analysis with the often-used specification of an equity (market) risk premium  $\mu v(t)$  and a variance risk premium  $\lambda v(t)$ . We refer to this specification as “completely affine”, adopting the terminology in Singleton (2006, p. 392) and the term structure literature. We characterize a pricing kernel that is consistent with these risk premia, and derive the parameter restrictions consistent with the martingale conditions and absence of arbitrage. We then explore a specification with an equity risk premium  $\mu_0 + \mu_1 v(t)$  and a variance risk premium  $\lambda_0 + \lambda_1 v(t)$ . We refer to this specification as “affine”. Singleton (2006) points out that this specification is problematic because it may violate no-arbitrage conditions, and notes that it would be interesting to characterize the parameter restrictions that prevent arbitrage for this specification. Because our kernel is formulated as a function of the state variables, it is straightforward to specify such restrictions.

Our empirical analysis uses a joint likelihood based on index returns and a rich option data set, using data for the January 1996 to June 2019 period.<sup>6</sup> We estimate this joint likelihood for the pricing kernels corresponding to the completely affine and affine prices of risk, and we also estimate it subject to various restrictions on parameter values and risk premia. Because the kernels are formulated as a function of the state variables, it is relatively straightforward to derive the implications of each kernel for the “marginal” kernel which specifies state price as a function of  $S(t)$ . The marginal kernel plays a central role in empirical applications such as the pricing kernel puzzle.

A first empirical result addresses the fit and empirical content of the kernels that support completely affine and affine prices of risk. Unsurprisingly, we find that the affine price of risk specification provides significantly better fit than the nested completely affine specification. However, the improved fit due to the intercepts in the affine specification comes at the cost of implausible Sharpe ratios and/or signs of the risk premiums. Moreover, while the marginal pricing kernel for the completely affine specification is well behaved and economically plausible, the kernel in the

---

<sup>6</sup>Much of the modern option pricing literature jointly considers the time-series of observable returns and option prices. See, for instance, Pan (2002), Eraker (2004), Bates (2006), Ait-Sahalia and Kimmel (2007), Hurn, Lindsay, and McClelland (2015), and Andersen, Fusari, and Todorov (2017).

affine case implies state prices that are S-shaped as a function of wealth. When we impose additional parameter restrictions that preclude arbitrage, the marginal pricing kernel is well-behaved, but empirical fit worsens. We conclude that the affine specification corresponds to implausible economic assumptions, and that small and seemingly innocuous modifications to the price of risk specifications used in the literature correspond to different pricing kernels with radically different economic implications. Consequently, we advocate the use of the completely affine price of risk specification.

A second set of results compares the unrestricted kernel that supports the completely affine price of risk with restricted versions. We reject restrictions that the equity or variance risk premiums are zero, but we cannot reject the independence of the pricing kernel from either variance or market return shocks.<sup>7</sup> We are unable to statistically pin down the origins of these risk premiums because innovations to market returns and market variance are highly (negatively) correlated. The equity and variance risk premium each have two components, one due to variance aversion and another due to index level risk aversion. If we restrict one of the risk aversion parameters to be zero, the other parameter absorbs most of that effect. In other words, these risks largely span each other. However, restricting variance or index level risk aversion to zero implies radically different estimates of the equity and variance risk premia, as well as large differences in the state prices embodied in the marginal pricing kernel as a function of wealth. The realized time series path of the kernel without volatility risk is also substantially less variable compared to the unrestricted kernel, especially in crisis periods. A final observation is that while we cannot statistically reject the path-independence restriction used in Ross (2015), it implies an implausible estimate of the variance aversion parameter.

Our third finding sheds light on the pricing kernel puzzle – the finding that the marginal pricing kernel is U-shaped – in the existing literature. We show that U-shaped marginal kernels can result from an underlying pricing kernel that is a monotonic function of volatility and the stock price. Therefore, U-shaped pricing kernels are not anomalous nor do they constitute an asset pricing

---

<sup>7</sup>Note that the hypothesis that the variance-aversion parameter equals zero amounts to the absence of an independent variance risk premium, which amounts to logarithmic utility in the Merton (1973) ICAPM.

puzzle.

Our fourth finding addresses estimation of risk premiums. Following Breeden and Litzenberger’s (1978) insight that the risk-neutral density can be inferred from option prices, financial economists have emphasized fitting options and returns jointly to identify risk premia. We find that these data have low power to distinguish different pricing kernels, because identifying pricing kernels is equivalent to the estimation of conditional risk premia, and it is difficult to estimate average returns over short periods.<sup>8</sup> Different parameter restrictions lead to widely different Sharpe ratios and equity and variance risk premia, but do not translate into large decreases in the likelihood. Merton (1980) convincingly argues that very long time series of returns are required to obtain reliable estimates of the equity premium. Our findings extend Merton’s observation to joint estimation of equity and variance risk premia. We also reinforce Merton’s conclusion that economic restrictions increase power to identify market risk premia. While Merton (1980) advocates imposing a positivity restriction on the path of the conditional equity risk premium, we find that imposing a negativity restriction on the market variance risk premium leads to more plausible and reliable estimates. Our findings also confirm the results in Bakshi, Crosby, and Gao (2022) that some option model parameters are hard to identify because of (dark matter) unspanned risks that affect risk premiums.

Our paper is related to several other strands of literature besides those on the estimation of parametric option pricing models and the pricing kernel puzzle. Several studies use consumption-based models to analyze how preferences and pricing kernels impact index option prices.<sup>9</sup> Some of these studies use the recursive preferences of Kreps and Porteus (1978), Epstein and Zin (1989) and Duffie and Epstein (1992), which result in stochastic volatility of index returns. Our proposed pricing kernels are extensions of the power utility of Rubinstein (1976). While consumption is not a state variable in our setup, our approach provides a direct relation with existing empiri-

---

<sup>8</sup>This statement is specific to plain vanilla option prices, which are sensitive to the probabilities at expiration but not very informative about the path-dependent properties of the pricing kernel. Pricing kernels with widely different economic implications can therefore produce similar values for European options.

<sup>9</sup>See, for instance, Garcia, Luger, and Renault (2003), Eraker and Shaliastovich (2008), Drechsler (2013), Shaliastovich (2015), Eraker and Yang (2019), and Seo and Wachter (2019). Liu, Pan, and Wang (2005) and Eraker and Wu (2017) use related models with the dividend payout rate and cash flow respectively as the state variable.

cal implementations of (reduced-form) parametric dynamic option pricing models. It is therefore straightforward to implement using option data, which allows us to explore the impact of stock index volatility on the pricing kernel.

From an empirical perspective, a related paper is Chernov (2003), who reverse engineers the pricing kernel based on options on various securities. Chernov (2003) also studies the time path of the realized pricing kernel to learn about state variables and the relation between the pricing kernel and economic conditions. Ghosh, Julliard, and Taylor (2017) also explore the relation between the pricing kernel and business cycle fluctuations, but do not use options to estimate the kernel. Brennan, Liu, and Xia (2006) specify and estimate pricing kernels with multiple state variables. Beason and Schreindorfer (2022) analyze the implications of option data for macro-finance models. Dew-Becker and Giglio (2022) study the implications of synthetic puts for the properties of the marginal pricing kernel.

The paper proceeds as follows. Section 2 discusses the data. Section 3 reviews the Heston (1993) stochastic volatility model and discusses our estimation approach based on returns and options data. Section 4 specifies the class of pricing kernels that connect the risk-neutral and physical dynamics. Section 5 presents the estimation results and Section 6 discusses their economic implications. Section 7 concludes.

## 2 Data

Our empirical analysis uses out-of-the-money (OTM) S&P 500 call and put options with maturities between 14 and 365 days for the January 1996 to June 2019 period. We obtain the option data from OptionMetrics. We apply the following filters:

1. Discard options with implied volatility smaller than 5% or greater than 150%.
2. Discard options with volume or open interest less than ten contracts.
3. Discard options with mid price less than \$0.50 or bid price less than \$0.375 to avoid low-valued options.

4. Discard options with data errors – where bid price exceeds offer price, or a negative price is implied through put-call parity.
5. Discard options with moneyness  $< 0.75$  or  $> 1.25$ .

Then we keep the six most actively traded strike prices for each available maturity. It is important to use as long a time period as possible to identify key aspects of the model, including volatility persistence.<sup>10</sup> On the other hand, estimation using large option panels and long time series is very time-intensive. Rather than using a short time series of daily option data, we use an extended time period, but we select option contracts for one day per week only. Following several existing studies (see, e.g., Heston and Nandi, 2000; Christoffersen, Heston, and Jacobs, 2013), we use Wednesday data because it is the day of the week least likely to be a holiday. It is also less likely than other days to be affected by day-of-the-week effects. These steps result in a dataset with 62,483 option contracts. Table 1 presents descriptive statistics.

We obtain S&P 500 index returns from CRSP. We use data for the January 1990 to June 2019 period. This sample starts before the option sample to help with the identification of the return parameters under the physical measure, as in Christoffersen, Heston, and Jacobs (2013). We also use data on the VIX from January 1990 to June 2019, which we obtain from the Federal Reserve Bank of St. Louis Economic Database. The time series for the risk-free rate is proxied by the one-month Treasury Bill rate obtained from CRSP. Following existing work, options are valued using a maturity-specific risk-free rate. We apply a cubic spline interpolation to the data obtained from OptionMetrics.

### 3 Return-Based and Option-Based Parameter Estimates

We estimate the stylized affine Heston (1993) stochastic volatility model. We obtain parameter estimates for this model under the physical measure, exclusively based on returns, and under the risk-neutral measure, exclusively based on options. Then we compare the resulting estimates.

---

<sup>10</sup>See, for instance, Broadie, Chernov, and Johannes (2007) for a discussion.



### 3.1 The Model

We focus on the simplest possible stochastic volatility model with a single diffusive volatility factor. We recognize that the existing literature has clearly established that additional volatility factors, jumps in returns and variance and/or tail factors are required to improve option fit and pricing performance. However, we deliberately focus on the simplest possible model because it suffices to illustrate our main argument and we want to avoid comparisons between models and factors. Our analysis can be repeated using more general models, but at the cost of much greater complexity. We believe that most of the issues we highlight here using a simple model are even more relevant in more complex models, but we leave this analysis for future work.

We employ the Heston (1993) continuous-time stochastic square root volatility model to specify stock price dynamics as well as option prices. For option valuation, the risk-neutral stock price dynamic is sufficient. The square root stochastic volatility model specifies the risk-neutral dynamics of the spot index  $S(t)$  and its stochastic variance  $v(t)$  as follows:

$$\begin{aligned} dS(t)/S(t) &= rdt + \sqrt{v(t)}dz_1^*(t), \\ dv(t) &= \kappa^*(\theta^* - v(t))dt + \sigma\sqrt{v(t)}dz_2^*(t), \end{aligned} \tag{1}$$

where  $dz_1^*$  and  $dz_2^*$  are Wiener processes with correlation coefficient  $\rho$ . The risk-free rate  $r$  can be either constant or time-varying; this has negligible implications for our results. It is also straightforward to specify a stochastic model for the risk-free rate, but it is well-known from the existing literature that this does not have a major impact on option valuation (Bakshi, Cao, and Chen, 1997). We therefore deliberately focus on the simplest possible model. Consistent with most of the existing literature, we focus on a physical dynamic that has the same functional form as the risk-neutral dynamic:

$$\begin{aligned} dS(t)/S(t) &= [r + \mu(v(t))] dt + \sqrt{v(t)}dz_1(t), \\ dv(t) &= \kappa(\theta - v(t))dt + \sigma\sqrt{v(t)}dz_2(t), \end{aligned} \tag{2}$$

where  $\mu(v(t))$  denotes the equity premium as a function of  $v(t)$ , and  $dz_1$  and  $dz_2$  are Wiener processes under the physical measure. Note that  $\sigma$ , the variance of variance parameter, and  $\rho$ , the correlation between  $z_1$  and  $z_2$ , are assumed to be identical to the corresponding parameters in the risk-neutral dynamics. However, the long-run physical variance  $\theta$  and mean reversion  $\kappa$  differ from the long-run risk-neutral variance  $\theta^*$  and mean reversion  $\kappa^*$ . This specification is consistent with the existing literature. It represents the most general combination of physical and risk-neutral dynamics that are consistent with the affine specification and Girsanov's theorem. We analyze this mapping in more detail below in our discussion of (the) pricing kernel(s).

### 3.2 The Instantaneous Stochastic Variance and the VIX

In the Heston (1993) model, as well as in its many generalizations studied in the literature, the stochastic variance is unknown. This latency is typically addressed in estimation by using filtering- or simulation-based techniques (see, e.g., Eraker, Johannes, and Polson, 2003; Eraker, 2004; Bates, 2006; Christoffersen, Jacobs, and Mimouni, 2010). It is well-known that the implementation of such techniques is computationally very demanding, especially when using long time series and large cross-sections of option prices in estimation.

To alleviate this computational burden, we follow a different approach.<sup>11</sup> We use the fact that the stochastic variance  $v(t)$  can be represented as a linear function of  $\text{VIX}^2(t)$ . This directly follows from the model specification: When  $v(t)$  follows a CIR process,  $\text{VIX}^2(t)$  is a linear function of  $v(t)$ . Specifically, the model-implied  $\text{VIX}^2(t)$  is given by:

$$\begin{aligned} \text{VIX}^2(t) &= \frac{1}{\Delta_{1m}} E_t^* \left[ \int_t^{t+\Delta_{1m}} v(u) du \right] \\ &= \theta^* + \frac{e^{-\kappa^* \Delta_{1m}} - 1}{-\kappa^* \Delta_{1m}} (v(t) - \theta^*), \end{aligned} \tag{3}$$

---

<sup>11</sup>See Bates (2000) and Andersen, Fusari, and Todorov (2015) for alternative approaches.

where  $\Delta_{1m} \approx 30/365$ . Rearranging equation (3) yields

$$v(t) = \frac{\text{VIX}^2(t) - \theta^*(1 - w)}{w}, \quad (4)$$

where  $w = (1 - \exp(-\kappa^* \Delta_{1m})) / (\kappa^* \Delta_{1m})$ . In implementation, we can add a measurement error because equations (3) and (4) use the model-implied  $\text{VIX}^2(t)$ . Equation (3) in conjunction with the measurement error yields a measurement equation which can be used to filter the latent state variable. Jones (2003), Cheung (2008), and Chernov, Graveline, and Zviadadze (2018) use this measurement equation and a Bayesian framework with Markov chain Monte Carlo methods to estimate option pricing models. We further simplify the setup: We do not use the measurement equation, but relax the restrictions on the coefficients in equation (4) and omit the measurement error. Specifically, we assume:

$$v(t) = \eta_0 + \eta_1 \text{VIX}^2(t). \quad (5)$$

We then use equation (5) in the valuation formula for all options in the sample. As a result, options are a function not only of the stochastic  $v(t)$ , but also of the observable VIX. This implementation follows Aït-Sahalia and Kimmel (2007), who use it in a sample which contains a single short-maturity at-the-money option at each time  $t$ .

We next discuss the details of this estimation approach when using returns and when using options. Our use of the VIX as a proxy for the stochastic variance has implications for both estimation exercises.

### 3.3 Return-Based Estimation

The main purpose of the assumption that the stochastic variance is an affine function of VIX is to alleviate the computational burden when estimating the model using option data. However, this assumption also has implications for the return-based estimation. Since we observe the total return of the stock index and VIX at each time  $t$ , we can formulate the joint likelihood function of the return and  $\text{VIX}^2$  to estimate the physical parameters. In most existing estimations, the variance is

instead filtered from the underlying returns, and the VIX is not used in estimation.

To characterize the likelihood function, we first apply Ito's lemma and the Euler discretization to equation (2), which results in:

$$\begin{aligned}\log R(t + \Delta) &= \left[ r + \mu(v(t)) - \frac{1}{2}v(t) \right] \Delta + \epsilon_R(t + \Delta), \\ v(t + \Delta) - v(t) &= \kappa(\theta - v(t))\Delta + \epsilon_v(t + \Delta),\end{aligned}\tag{6}$$

where  $R(t + \Delta) = S(t + \Delta)/S(t)$  represents the gross return and  $\Delta = 1/252$ .<sup>12</sup> The errors  $\epsilon(t + \Delta) = (\epsilon_R(t + \Delta), \epsilon_v(t + \Delta))'$  follow a joint normal distribution, and their mean and variance-covariance matrix are given by

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{\Sigma}(t) = \begin{pmatrix} v(t) & \sigma\rho v(t) \\ \sigma\rho v(t) & \sigma^2 v(t) \end{pmatrix} \Delta.$$

The joint log-likelihood function is given by:

$$\begin{aligned}\log L^R &= \sum_{t=1}^{T-1} \log f(\log R(t + \Delta), \text{VIX}^2(t + \Delta) | \text{VIX}^2(t)) \\ &= \sum_{t=1}^{T-1} \log f(\log R(t + \Delta), v(t + \Delta) | v(t)) \times J(t + \Delta) \\ &= \sum_{t=1}^{T-1} -\log(2\pi) - \frac{1}{2} \log |\mathbf{\Sigma}(t)| - \frac{1}{2} \epsilon(t + \Delta)' \mathbf{\Sigma}^{-1}(t) \epsilon(t + \Delta) + \log \eta_1,\end{aligned}$$

where  $f(\log R(t + \Delta), v(t + \Delta) | v(t))$  is the conditional density of the discretized  $\log R(t + \Delta)$  and  $v(t + \Delta)$ ,  $J(t + \Delta)$  is the Jacobian between  $\text{VIX}^2(t + \Delta)$  and  $v(t + \Delta)$ , which is given by  $\eta_1$  from equation (5), and  $t$  represents time measured in days. Let  $\Theta = \{\mu, \kappa, \theta, \sigma, \rho, \eta_0, \eta_1\}$  be the set of physical parameters. To estimate  $\Theta$ , we solve the following optimization problem:

$$\max_{\Theta} \log L^R.\tag{7}$$

---

<sup>12</sup>Note that  $\log R(t + \Delta)$  is the daily log return between  $t$  and  $t + \Delta$  while  $v(t)$  is the annualized variance at time  $t$ .

### 3.4 Option-Based Estimation

The risk-neutral parameters for the dynamic in equation (1) can be estimated in various ways, but each implementation requires an option valuation technique. We follow the fast Fourier implementation of Carr and Madan (1999). The price of a call option with its strike price  $K$  and maturity  $\tau$  is expressed by a quasi closed form up to a numerical integration, and it is given by

$$C(S(t), v(t), t) = \frac{e^{-\alpha k}}{\pi} \int_0^\infty \text{Re} \left[ e^{-iuk} \psi(u) \right] du, \quad (8)$$

where  $k$  is the natural log of  $K$ . The function  $\psi(u)$  is the Fourier transform of a modified call price, which is the call price multiplied by  $e^{\alpha k}$  for  $\alpha > 0$ . We found that  $\alpha = 4$  works well. The function  $\psi(u)$  is calculated as follows:

$$\psi(u) = \frac{e^{-r\tau} f_\tau^{CH}(u - i(\alpha + 1) | S(t), v(t))}{(\alpha + iu)(\alpha + 1 + iu)},$$

where  $i$  is the imaginary unit, and  $f_\tau^{CH}(\phi | S(t), v(t)) = E_t^* [e^{i\phi \log S(t+\tau)}]$  is the risk-neutral conditional characteristic function of  $\log S(t + \tau)$ . The closed-form expression of  $f_\tau^{CH}(\phi | S(t), v(t))$  follows Heston (1993).<sup>13</sup> The price of a put option with the same strike price and maturity can be obtained through put-call parity. Note that the option pricing formula in equation (8) does not account for dividends. We follow the existing literature and use a future-dividend-adjusted index price. Specifically, we use  $S(t)e^{-q\tau}$ , where  $q$  is the dividend yield at time  $t$ .

---

<sup>13</sup>When  $\log S(t)$  and  $v(t)$  are characterized by

$$\begin{aligned} d \log S(t) &= [r + uv(t)]dt + \sqrt{v(t)}dz_1(t), \\ dv(t) &= (a - bv(t))dt + \sigma\sqrt{v(t)}dz_2(t), \end{aligned}$$

the characteristic function solution is given by

$$f_\tau^{CH}(\phi | S(t), v(t)) = e^{C + Dv(t) + i\phi \log S(t)}, \quad (9)$$

where

$$\begin{aligned} C &= r\phi i\tau + \frac{a}{\sigma^2} \left\{ (b - \rho\sigma\phi i + d)\tau - 2 \log \left[ \frac{1 - ge^{d\tau}}{1 - g} \right] \right\}, \quad D = \frac{b - \rho\sigma\phi i + d}{\sigma^2} \left[ \frac{1 - e^{d\tau}}{1 - ge^{d\tau}} \right], \\ g &= \frac{b - \rho\sigma\phi i + d}{b - \rho\sigma\phi i - d}, \quad \text{and} \quad d = \sqrt{(\rho\sigma\phi i - b)^2 - \sigma^2(2u\phi i - \phi^2)}. \end{aligned} \quad (10)$$

We use vega-weighted option pricing errors. Let  $O_i^{Mkt}$  and  $O_i^{Mod}$  denote the market and model prices of the  $i^{th}$  option, respectively. Both  $O_i^{Mkt}$  and  $O_i^{Mod}$  represent call option prices if  $F/K < 1$  and put option prices if  $F/K > 1$ . Define the vega-weighted option pricing errors as

$$\epsilon_{o,i} = \frac{O_i^{Mkt} - O_i^{Mod}}{\nu_i^{Mkt}},$$

where  $\nu_i^{Mkt}$  is the Black-Scholes vega of option  $i$ .<sup>14</sup>

Maximum likelihood estimation requires a distributional assumption. Following most of the existing literature, we assume that  $\epsilon_{o,i}$  follows a normal distribution, *i.e.*  $\epsilon_{o,i} \sim N(0, s_o^2)$ , where  $s_o^2$  is the sample variance of the errors. Option valuation errors are assumed to be independent and identically distributed (i.i.d.).

The set of risk-neutral parameters to be estimated is denoted by  $\Theta^* = \{\kappa^*, \theta^*, \sigma, \rho, \eta_0, \eta_1\}$ . Let  $N$  be the total number of options data.  $\Theta^*$  is then estimated by solving the following optimization problem:

$$\max_{\Theta^*} \log L^O,$$

where the option log-likelihood function,  $\log L^O$ , is given by:

$$\log L^O = -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log s_o^2 - \frac{1}{2s_o^2} \sum_{i=1}^N \epsilon_{o,i}^2.$$

### 3.5 Parameter Estimates

Table 2 presents the estimation results. Panel A presents the (physical) parameters estimated from returns, and Panel B presents the (risk-neutral) parameters estimated from options.

The physical parameter estimates are based on the stock index return and VIX data. The risk-neutral parameters are estimated exclusively based on options data. Both physical and risk-neutral parameter estimates are economically plausible. Consistent with findings in the existing literature,

---

<sup>14</sup>These errors are often used because option prices are very different but implied volatilities are in a more narrow range. See for example Carr and Wu (2007), Trolle and Schwartz (2009), and Christoffersen, Heston, and Jacobs (2013).

$\kappa$  is much larger than  $\kappa^*$ , while  $\theta$  is much smaller than  $\theta^*$ . The risk-neutral long-run variance exceeds the physical long-run variance, and risk-neutral persistence exceeds physical persistence.

The option-based kurtosis parameter  $\sigma$  is larger than the return-based estimate of  $\sigma$ , and the option-based skewness parameter  $\rho$  is more negative than the return-based  $\rho$ . The distribution implied by the option data is thus more fat-tailed and skewed than the physical distribution. The finding on  $\sigma$  is mostly consistent with the existing literature. Bakshi, Cao, and Chen (1997), Eraker (2004), and Christoffersen, Jacobs, and Mimouni (2010) also obtain higher estimates of  $\sigma$  when estimating on options. Existing findings on  $\rho$  are mixed. Typically the estimates from returns are not very different from the option-based estimates.

The estimate of  $\mu$  in Table 2 implies an average yearly equity premium  $\mu v(t)$  of 8.19% for the January 1990 to June 2019 sample period, close to the sample average of 8.24%. It is not possible to infer the path of the model-implied variance risk premium using the estimated parameters because the physical and risk-neutral estimations do not constrain the parameter estimates of  $\sigma$ ,  $\rho$ ,  $\eta_0$ , and  $\eta_1$  to be the same. However, we can use the parameters  $\theta$  and  $\theta^*$  to compare the long-run means of the physical and risk-neutral stochastic variances. Taking the square root, we find that the model-implied long-run expectation of the stock index physical (risk-neutral) yearly volatility is 17.6% (31.4%). Figure 1 shows the time path of the option-based variance, as well as the time path of the difference between the return-based and option-based variance.

## 4 A Taxonomy of Pricing Kernels

We characterize various pricing kernels implied by the physical and risk-neutral dynamics in equations (1) and (2). Specifically, we analyze pricing kernels consistent with completely affine and affine equity and variance risk premiums. We discuss the differences between these kernels and the economic implications of these differences. We highlight several special cases, including the power utility kernel of Rubinstein (1976) and the path-independent kernel of Ross (2015).

## 4.1 The Pricing Kernel by Girsanov's Theorem

The existing literature on the Heston (1993) square root model typically restricts the physical and risk-neutral dynamics to have the same functional form, as in equations (1) and (2), but is not always explicit about the relation between the two dynamics, and by extension about the implied risk premiums and pricing kernels. We discuss two (nested) assumptions on the risk premiums that allow the physical and risk-neutral dynamics to have the same functional form.

### 4.1.1 Completely Affine Prices of Risk

Heston (1993) restricts the variance risk premium to be proportional to the variance, i.e., equal to  $\lambda v(t)$ . Because Heston (1993) does not perform (joint) estimation of the model parameters, he does not explicitly address the structure of the equity premium. However, several seminal estimation exercises (see, e.g., Bates, 2000; Pan, 2002; Ait-Sahalia and Kimmel, 2007) impose a similar proportional structure on the equity premium, i.e., equal to  $\mu v(t)$ . We therefore start our analysis with this specification of the equity and variance risk premium. Henceforth, we refer to this specification, with equity and variance risk premiums that are proportional to the stochastic variance, as *completely affine* risk premiums. Note that in this case, the relation between the risk-neutral and physical parameters in equations (1) and (2) are given by  $\kappa = \kappa^* - \lambda$  and  $\theta = \kappa^* \theta^* / \kappa$ .

Using Girsanov's Theorem, Online Appendix A shows that the following pricing kernel is consistent with the risk-neutral dynamic (1), the physical dynamic (2) and completely affine equity and variance risk premiums.

$$M(t) = M(0) \exp \left( -rt - \frac{1}{2} \int_0^t \text{Var} \left[ \frac{dM(s)}{M(s)} \right] - \int_0^t \pi_1 \sqrt{v(s)} dz_1(s) - \int_0^t \pi_2 \sqrt{v(s)} dz_2(s) \right), \quad (11)$$

where  $\text{Var} \left[ \frac{dM(s)}{M(s)} \right] = (\pi_1^2 + \pi_2^2 + 2\rho\pi_1\pi_2) v(s) ds$ , and  $\pi_1$  and  $\pi_2$  are price of risk parameters, which are related to the risk premium parameters as follows:

$$\mu = \pi_1 + \rho\pi_2 \quad \text{and} \quad \lambda = \rho\sigma\pi_1 + \sigma\pi_2, \quad (12)$$



### 4.1.2 Affine Prices of Risk

We now consider a generalization of the completely affine equity and variance risk premiums in Section 4.1.1. Specifically, we consider kernels that are consistent with equity and variance risk premiums that are linear (affine) in the stochastic variance but also contain some nonzero intercepts; that is,  $\mu_0 + \mu_1 v(t)$  and/or  $\lambda_0 + \lambda_1 v(t)$ . In this case, the relation between the risk-neutral and physical parameters is given by  $\kappa = \kappa^* - \lambda_1$  and  $\theta = (\kappa^* \theta^* + \lambda_0) / \kappa$ . See Chernov and Ghysels (2000) and Eraker (2004) for examples.

We can once again use Girsanov's Theorem to characterize the pricing kernel which is consistent with these risk premiums, the risk-neutral dynamics (1) and the physical dynamics (2). This gives:

$$M(t) = M(0) \exp \left( -rt - \frac{1}{2} \int_0^t \text{Var} \left[ \frac{dM(s)}{M(s)} \right] - \int_0^t \left( \frac{\pi_{1,0}}{\sqrt{v(s)}} + \pi_{1,1} \sqrt{v(s)} \right) dz_1(s) - \int_0^t \left( \frac{\pi_{2,0}}{\sqrt{v(s)}} + \pi_{2,1} \sqrt{v(s)} \right) dz_2(s) \right), \quad (13)$$

where the parameters that characterize the affine risk premium parameters can once again be expressed in terms of these price of risk parameters, as follows:

$$\mu_0 = \pi_{1,0} + \rho\pi_{2,0}, \quad \mu_1 = \pi_{1,1} + \rho\pi_{2,1}, \quad \lambda_0 = \rho\sigma\pi_{1,0} + \sigma\pi_{2,0}, \quad \lambda_1 = \rho\sigma\pi_{1,1} + \sigma\pi_{2,1}, \quad (14)$$

and

$$\text{Var} \left[ \frac{dM(s)}{M(s)} \right] = \left[ (\pi_{1,0}^2 + \pi_{2,0}^2 + 2\rho\pi_{1,0}\pi_{2,0}) \frac{1}{v(s)} + (\pi_{1,1}^2 + \pi_{2,1}^2 + 2\rho\pi_{1,1}\pi_{2,1}) v(s) + 2(\pi_{1,0}\pi_{1,1} + \pi_{2,0}\pi_{2,1} + \rho\pi_{1,0}\pi_{2,1} + \rho\pi_{2,0}\pi_{1,1}) \right] ds. \quad (15)$$

One critical difference between the pricing kernels in equations (11) and (13) is that (13) may allow for arbitrage when the variance reaches zero (Singleton, 2006). This can be addressed by imposing the Feller (1951) condition(s). We discuss this in more detail below.

## 4.2 A General Pricing Kernel

The pricing kernels in equations (11) and (13) contain complicated, opaque stochastic integrals that lack a clear economic interpretation. The purpose of this section is to simplify and interpret these kernels using (generalizations of) economically motivated pricing kernels from the finance literature. These include the seminal power utility kernel of Rubinstein (1976) and Brennan (1979) and the transition-independent kernel of Ross (2015).

We use the principles of monotonicity and parsimony to discuss a taxonomy of economic pricing kernels, where our exposition progresses from relatively complex kernels with more parameters to simpler monotonic kernels with fewer parameters. Monotonicity is based on the straightforward economic rationale that we expect marginal preferences to be decreasing in equity returns and increasing in variance. Parsimony appeals to Occam’s razor as a pragmatic tool. Complicated models suffer from a “curse of dimensionality”, and are often statistically indistinguishable from simpler models. Therefore, we prefer to condense complicated models to simpler ones with fewer parameters.

All kernels we consider are formulated in terms of the state variables  $S(t)$  and  $v(t)$ . The most general pricing kernel we consider is given by:

$$M(t) = M(0) \left( \frac{S(t)}{S(0)} \right)^{-\gamma} \left( \frac{v(t)}{v(0)} \right)^{\alpha} \exp \left( \beta t + \int_0^t \eta(v(s)) ds + \xi(v(t) - v(0)) \right). \quad (16)$$

Note that this is not the most general possible kernel with state variables  $S(t)$  and  $v(t)$ . We start our analysis with this kernel because it provides insight into the affine equity and variance risk premium specification, or equivalently into the pricing kernel in equation (13). Specifically, Online Appendix B shows that the mapping between the preference parameters in equation (16) and the reduced-form parameters characterizing the affine risk premiums  $\mu_0 + \mu_1 v(t)$  and  $\lambda_0 + \lambda_1 v(t)$  is as follows:

$$\mu_0 = -\alpha\sigma\rho, \quad \mu_1 = \gamma - \rho\sigma\xi, \quad \lambda_0 = -\alpha\sigma^2, \quad \lambda_1 = \gamma\rho\sigma - \xi\sigma^2. \quad (17)$$

A brief discussion of the parameter count in these kernels is in order. The affine risk premium

specification contains four parameters  $(\mu_0, \mu_1, \lambda_0, \lambda_1)$ . The kernel (16) characterizes these four parameters in terms of the parameters in the physical dynamic and the three preference parameters  $\gamma$ ,  $\alpha$  and  $\xi$ . We therefore refer to (16) as a three-parameter kernel.

One may argue that the kernel (16) contains two additional preference parameters,  $\beta$  and  $\eta$ . But using the martingale restriction on the kernel itself, Online Appendix B shows that the time-preference parameter  $\beta$  and the path-dependence function  $\eta(v(t))$  are given by:

$$\beta = -(1 - \gamma)r + \gamma\mu_0 - \xi\kappa\theta + \alpha\kappa + \gamma\alpha\sigma\rho - \xi\alpha\sigma^2, \quad (18)$$

$$\eta(v(t)) = \left[ \left( \mu_1 - \frac{1}{2} \right) \gamma + \xi\kappa - \frac{1}{2} (\gamma^2 - 2\gamma\xi\sigma\rho + \xi^2\sigma^2) \right] v(t) - \left[ \kappa\theta - \frac{1}{2}\sigma^2 + \frac{1}{2}\alpha\sigma^2 \right] \frac{\alpha}{v(t)}. \quad (19)$$

Restrictions (18)-(19) are required to match the risk-free rate, which is an affine function of the variance. Restriction (18) constrains the intercept in this affine function by constraining the  $\beta$  parameter, and restriction (19) constrains the slope. Equation (19) stipulates that given the physical parameters and the three preference parameters  $\gamma$ ,  $\alpha$  and  $\xi$ , there is a (time-varying) path-dependence parameter  $\eta(v(t))$  that imposes no-arbitrage. The implications of (not) imposing the restrictions (18) and (19) are outside the scope of our study.<sup>15</sup> We merely want to point out that for the purpose of the mapping between the kernel (16) and the affine risk premiums, the kernel effectively uses three preference parameters to characterize the four parameters  $\mu_0, \mu_1, \lambda_0$ , and  $\lambda_1$ .

From an economic perspective, we note two important features of the kernel (16). First, while it contains an integral that depends on the history of the variance, it does not depend on the history of the stock price. This represents a substantial simplification from the general stochastic integrals in (11) and (13). Second, the kernel can be seen as extending the power utility of Rubinstein (1976) and Brennan (1979) to include the pricing of variance risk. However, the kernel contains two parameters ( $\alpha$  and  $\xi$ ) that govern the pricing of variance, and this is a potential source of problems. Any economically plausible pricing kernel should produce a monotonic variance premium. Because

---

<sup>15</sup>The restrictions (18) and (19) should in principle be imposed in empirical work *in addition* to the affine risk premiums  $\mu_0 + \mu_1 v(t)$  and  $\lambda_0 + \lambda_1 v(t)$ , but usually these restrictions are ignored. This amounts to stating that one has no prior on the magnitude of the path-dependence parameter, and that one is willing to accept the risk-free rate implied by the values of  $\mu_0, \mu_1 v(t), \lambda_0, \lambda_1$ , even if the resulting risk-free rates are economically implausible.

$\alpha$  affects the intercept of the variance premium and  $\xi$  affects the slope (see equation (17)), the pricing kernel can be a nonmonotonic function of variance if these parameters have opposite signs. In this case, investors sometimes like higher variance, while at other times they prefer lower variance. This will likely have strange implications for the cross-section of option returns.

### 4.3 A Monotonic Kernel

To prevent these implausible nonmonotonicities, we can restrict  $\alpha$  in equation (16) to be zero, which gives the following simpler two-parameter specification:

$$M(t) = M(0) \left( \frac{S(t)}{S(0)} \right)^{-\gamma} \exp \left( \beta t + \eta \int_0^t v(s) ds + \xi(v(t) - v(0)) \right). \quad (20)$$

This specification is easier to interpret economically, because one parameter  $\gamma$  determines equity risk-aversion, and another parameter  $\xi$  governs variance-aversion. Moreover, Online Appendix C shows that given the risk-neutral dynamics (1), the physical dynamics (2), no-arbitrage, and completely affine equity and variance risk premiums  $\mu v(t)$  and  $\lambda v(t)$ , the two pricing kernel parameters  $\gamma$  and  $\xi$  determine the equity premium  $\mu$  and variance premium  $\lambda$  parameters, as follows:

$$\text{Equity risk premium:} \quad \mu v(t) = (\gamma - \xi \sigma \rho) v(t), \quad (21)$$

$$\text{Variance risk premium:} \quad \lambda v(t) = (\rho \sigma \gamma - \sigma^2 \xi) v(t) = (\rho \sigma \mu - (1 - \rho^2) \sigma^2 \xi) v(t). \quad (22)$$

Equations (21) and (22) indicate that due to the correlation  $\rho$  between stock returns and variance, both the equity premium  $\mu$  and variance premium  $\lambda$  depend on both of the pricing parameters  $\gamma$  and  $\xi$ . The variance premium  $\lambda$  may be negative because the pricing kernel parameter  $\xi$  is positive, or because variance is negatively correlated with equity returns ( $\rho < 0$ ) and the equity premium is positive (i.e., variance has a negative beta). We can test these hypotheses separately, by restricting  $\xi$  or  $\lambda$  to be zero. Similarly, we can test the hypothesis that the equity risk-aversion  $\gamma$  is zero separately from the hypothesis that the equity premium  $\mu$  itself is zero. Finally, note that for this pricing kernel the no-arbitrage restrictions (18) and (19) reduce to:

$$\beta = -(1 - \gamma)r - \xi\kappa\theta, \quad (23)$$

$$\eta = \left(\mu - \frac{1}{2}\right)\gamma + \xi\kappa - \frac{1}{2}(\gamma^2 - 2\gamma\xi\sigma\rho + \xi^2\sigma^2). \quad (24)$$

Henceforth we refer to the kernel in equation (20) as the exponential-affine pricing kernel. Next we discuss two important special cases of this kernel.

#### 4.4 The Power Utility Pricing Kernel

A first economically important special case emanates from setting  $\xi$  equal to zero in equation (20). This imposes absence of variance preference, and gives:

$$M(t) = M(0) \left(\frac{S(t)}{S(0)}\right)^{-\gamma} \exp\left(\beta t + \int_0^t \eta v(s) ds\right). \quad (25)$$

The local fluctuations in the pricing kernel (25) now result exclusively from changes in the spot price  $S(t)$ , and there is no additional premium due to variance aversion. The martingale condition once again implies that the expected growth of the pricing kernel must equal the opposite of the interest rate. This restricts  $\beta$  and  $\eta$ , effectively leaving only a single index level preference parameter  $\gamma$  to price risk.

For convenience we will refer to the kernel in (25) as the power utility pricing kernel. However, a critical point is that the seminal work by Rubinstein (1976) and Brennan (1979) specifies economically motivated pricing kernels for economies with constant variance. While the kernel (25) is related to the power utility pricing kernels of Rubinstein (1976) and Brennan (1979), it highlights that we cannot simply adopt those static kernels in the presence of stochastic volatility. Specifically, the kernel contains a path dependent component  $\eta \int_0^t v(s) ds$  even in the absence of variance preference.

## 4.5 A Path Independent Pricing Kernel

Finally, following Ross (2015), we consider eliminating the dependence of the pricing kernel on the historical path of  $v(t)$ . This type of pricing kernel follows from time-separable expected utility or from more general Epstein-Zin preferences. Specifically, if we restrict  $\xi$  to solve the quadratic equation

$$\gamma^2 - (1 + 2\rho\sigma\xi)\gamma + \sigma^2\xi^2 + 2\kappa^*\xi = 0,$$

then the pricing kernel  $M(t)$  is a separable function of the contemporaneous stock price  $S(t)$  and contemporaneous variance  $v(t)$ :

$$M(t) = M(0) \left( \frac{S(t)}{S(0)} \right)^{-\gamma} \exp(-((1 - \gamma)r + \xi\kappa\theta)t + \xi(v(t) - v(0))).$$

This special case effectively sets  $\eta = 0$  in equation (24), which is a joint restriction on the equity risk aversion parameter  $\gamma$  and variance risk aversion  $\xi$ , and therefore also reduces the pricing kernel to a single parameter. We test this restriction in our empirical application.

## 5 Pricing Kernels: Parameter Estimates

We estimate the stochastic volatility model subject to the restrictions imposed by the pricing kernels in the previous section. Our results are based on the joint likelihood composed of returns and options. We use the empirical fit for different sets of restrictions to test the more general pricing kernels against more parsimonious versions. We also compare and analyze the risk premia and other economic implications for different restricted and unrestricted pricing kernels.

### 5.1 Joint Maximum Likelihood Estimation with No-Arbitrage Restrictions

Rather than separately estimating the physical and risk-neutral dynamics as in Table 2, we now jointly estimate both dynamics subject to various specifications of the pricing kernels. In our implementation, we either use a no-arbitrage condition based on a specific structure of the pricing

kernel, as in equation (20) for example, or we directly impose a structure on the parameters that characterize the risk premia such as  $\mu$  and  $\lambda$  in the completely affine case.

For kernels nested in the exponential-affine specification in equation (20), given a set of parameters  $\Theta^{PK1} = \{\gamma, \xi, \kappa, \theta, \sigma, \rho, \eta_0, \eta_1\}$ , we can obtain the risk premium parameters  $\mu$  and  $\lambda$  via equations (21) and (22). Likewise, for the more general kernel in equation (16), given  $\Theta^{PK2} = \{\gamma, \xi, \alpha, \kappa, \theta, \sigma, \rho, \eta_0, \eta_1\}$ , the risk premium parameters  $\mu_0, \mu_1, \lambda_0$ , and  $\lambda_1$  can be obtained from equation (17). We also investigate specifications that restrict the risk premium parameters instead. For the affine price of risk specification, the parameter vector is  $\Theta^{PR} = \{\mu_0, \mu_1, \lambda_0, \lambda_1, \kappa, \theta, \sigma, \rho, \eta_0, \eta_1\}$ . Regardless of the parameterization, estimation is based on the sum of the return and option log-likelihoods. That is, we solve the following optimization problem:

$$\max_{\Theta} \log L^R + \log L^O,$$

where  $\Theta$  can refer to  $\Theta^{PK1}$ ,  $\Theta^{PK2}$ , or  $\Theta^{PR}$ .

Table 3 presents estimation results for the exponential-affine kernel (20) and six restricted cases. Column (1) reports on the unrestricted exponential-affine pricing kernel in equation (20). In columns (2) and (3) we have two restricted estimation exercises based on risk premium restrictions:  $\mu = 0$  and  $\lambda = 0$  respectively. Next, we have several cases based on preference parameter restrictions. In column (4) we impose  $\gamma = 0$  and in column (5) we impose  $\xi = 0$ . The restriction  $\xi = 0$  corresponds to the power utility pricing kernel in equation (25). We also test the joint restrictions  $\gamma = \xi = 0$  in column (6).<sup>16</sup> Finally, we impose  $\eta = 0$  in column (7). This case provides insight into the literature on recovery theory (see, e.g., Ross, 2015; Borovička, Hansen, and Scheinkman, 2016; Qin and Linetsky, 2016). Table 4 reports on the affine specification of the price of risk in column (8), the more general kernel in equation (16) in column (9), and the special case of this kernel with  $\xi = 0$  in column (10). Online Appendix D summarizes the implementation of the restrictions used in these tables.

---

<sup>16</sup>From equations (21) and (22), the joint restriction  $\mu = \lambda = 0$  is equivalent to the joint restriction  $\gamma = \xi = 0$ .

## 5.2 Parameter Estimates: The Exponential-Affine Pricing Kernel

Table 3 reports the results of the joint MLE estimation. We report robust standard errors for all parameters, except for parameters that are determined by the other estimates, see Online Appendix D. Column (1) presents the estimates for the MLE estimation of the exponential-affine pricing kernel in equation (20).<sup>17</sup> We can compare the risk-neutral estimates in column (1) with the option-based estimates in Table 2. The risk-neutral mean-reversion  $\kappa^* = \kappa + \lambda$  from joint estimation in Table 3 is 1.114, somewhat higher than the 0.0986 estimate in Table 2. The risk-neutral long-run variance  $\theta^* = \kappa\theta/\kappa^*$  in Table 3 is 0.0877, somewhat lower than the 0.0986 estimate in Table 2. This finding is not surprising, because the option-implied variance typically exceeds the average variance implied by returns.

We can also compare the physical parameters in column (1) of Table 3 with the return-based estimates in Table 2. The physical mean reversion  $\kappa$  in Table 3 is 2.926, substantially smaller than the 4.146 estimate in Table 2. The physical long-run variance  $\theta$  in Table 3 is 0.033, similar to the 0.031 estimate in Table 2. The most important difference is that the return-based estimates of  $\rho$  and  $\sigma$  in Table 2 are smaller (in absolute value) compared to Table 3. Recall that in joint estimation the  $\rho$  and  $\sigma$  parameters are the same under both measures. The estimate of  $\rho$  from joint estimation is a bit larger in absolute value compared to the one in Panel B of Table 2, and the estimate of  $\sigma$  is a bit smaller, but the differences are relatively minor.

The estimate of the index level preference parameter (“risk aversion”)  $\gamma$  in column (1) is 1.393 and the estimate of the variance preference parameter  $\xi$  is 1.947. The signs of these preference parameters are consistent with economic intuition. The point estimate of  $\mu$  in column (1) of Table 3 is similar but slightly lower than the estimate in Table 2. The variance risk premium parameter  $\lambda$  is estimated to be negative in column (1) of Table 3, which is also intuitively plausible. The negative  $\lambda$  implies  $\kappa > \kappa^*$  and  $\theta < \theta^*$ . The negative  $\lambda$  also means that the ordinal relations between the physical and risk-neutral parameters are the same as in Table 2, which reports estimates from

---

<sup>17</sup>Recall that this kernel is isomorphic to the completely affine price of risk specification with equity premium  $\mu v(t)$  and variance risk premium  $\lambda v(t)$ .



returns only and options only.

### 5.3 Restrictions on the Exponential-Affine Pricing Kernel

Columns (2)-(7) in Table 3 report on estimation subject to various restrictions on the exponential-affine pricing kernel. When we impose restrictions on the risk premium parameters  $\mu$  and  $\lambda$  in columns (2) and (3), the resulting negative estimates of the risk preference parameters are inconsistent with economic intuition. The implied risk premium parameters  $\mu$  and  $\lambda$  and the physical parameters also substantially differ from the estimates in column (1). Columns (4) and (5) directly restrict risk preferences instead. The remaining unrestricted risk preference parameter estimates have the same sign as in column (1), but are larger in absolute value. The other parameters are very similar to column (1). Column (6) imposes the joint restriction  $\gamma = \xi = 0$ , which is equivalent to setting the risk premium parameters equal to zero ( $\mu = \lambda = 0$ ). These restrictions seem to impact the estimates of the drift parameters  $\kappa$  and  $\theta$ , but not the parameters  $\rho$  and  $\sigma$ , which determine skewness and kurtosis.

Column (7) tests the path-independence restriction, characterized by  $\eta = 0$ . In columns (1)-(6) and the other restricted estimations, the martingale property implies the value of  $\eta$ , and the resulting estimates are different from zero. The  $\eta = 0$  restriction results in a negative estimate of the variance preference parameter  $\xi$ , which is not economically plausible.

The most important finding is that while the pricing kernel parameters differ across columns (1)-(7), the risk-neutral parameter estimates  $\kappa^*$ ,  $\theta^*$ ,  $\sigma$ , and  $\rho$  remain remarkably similar across columns. As a result, the vega-weighted option RMSEs are nearly identical. However, the physical parameters  $\kappa$  and  $\theta$  differ, which explains the large differences in the pricing kernel parameters  $\gamma$ ,  $\xi$ , and  $\eta$ , or equivalently the risk premium parameters  $\mu$  and  $\lambda$ .

Panel A of Table 5 reports p-values for tests of the restricted pricing kernels against the unrestricted exponential-affine kernel in column (1). While the restrictions that one or both of the risk premium parameters ( $\mu$  and or  $\lambda$ ) are equal to zero in columns (2), (3) and (6) are rejected at the 5% level, this is not the case when we set one of the preference parameters ( $\gamma$  or  $\xi$ ) equal to zero

in columns (4) or (5). The reason is the robustly high (in absolute value) correlation between the innovations to aggregate variance and equity risk. When we set the representative agent’s equity risk aversion or variance aversion equal to zero, the other parameter simply takes over the task of the restricted parameter. This is confirmed by the parameter estimates in columns (4) and (5) of Table 3, where respectively  $\xi$  and  $\gamma$  increase when the other parameter is set to zero, and the overall likelihood is little affected. This finding is consistent with the recent findings of Dew-Becker and Giglio (2022) that the entire index variance risk premium is due to a negative stock beta.

#### 5.4 More General Pricing Kernels and Affine Prices of Risk

Table 4 presents results based on the more general kernel in equation (16) and the affine price of risk specification. Columns (9)-(10) present results for the pricing kernel in equation (16). The risk-neutral mean reversion is similar to that for the exponential-affine kernel in columns (1)-(7) in Table 3, while the risk-neutral mean-reversion is higher. The other parameters are very similar. Option RMSE is also similar to Table 3.

Column (8) in Table 4 reports on the affine specification of the price of risk with an equity premium  $\mu_0 + \mu_1 v(t)$  and a variance risk premium  $\lambda_0 + \lambda_1 v(t)$ . Note that the risk-neutral parameter estimates  $\theta^*$ ,  $\sigma$ , and  $\rho$  are again very similar to the estimates in Table 3, but the risk-neutral mean reversion  $\kappa^*$  is slightly smaller. The physical mean reversion is also larger, indicating that this specification allows for a larger wedge between the physical and risk-neutral mean reversion.

The joint log-likelihood in column (8) exceeds the log-likelihood in column (1) of Table 3, as well as the one in column (9). The likelihood ratio tests in Table 5 show that these differences are statistically significant at the 1% level. It seems that including an (unrestricted) intercept in the risk premium specification provides a better fit to the data. This improvement in fit mainly derives from a better fit of the return data. Note that in column (9), the more general kernel in equation (16) also contains an intercept in the risk premiums, but this intercept is subject to the restrictions in equation (17).

We conclude that the affine risk premium specification has important implications for model

fit. To the best of our knowledge, while this specification is frequently used in the option valuation literature, the implications of relaxing these restrictions are not documented or discussed in detail. Our results are consistent with findings in the bond pricing literature, where the completely affine specification is outperformed by more general specifications. Below we investigate the economic implications of these improvements in fit.

### 5.5 Imposing the Feller Condition(s)

Finally, while the estimates in column (9) impose the no-arbitrage restrictions emanating from the martingale condition in Section 4.2, they ignore that the more general kernel in equation (16) may also lead to zero state prices if the variance process reaches zero. Another necessary condition for no-arbitrage in this case is therefore that the Feller (1951) no-arbitrage condition holds under both measures, so  $2\kappa\theta > \sigma^2$  and  $2\kappa^*\theta^* > \sigma^2$ . Singleton (2006, p. 326) and Heston, Loewenstein, and Willard (2006) discuss how this same issue arises due to the presence of the intercept in the price of risk, which we show is equivalent to the specification of the kernel in equation (16).

Table A.1 in the Online Appendix repeats the estimation in columns (8) and (9) of Table 4 but now imposes the Feller conditions. These additional no-arbitrage restrictions lower the likelihood and worsen the option fit. The estimate of equity risk aversion  $\gamma$  now has the wrong sign in column (2). Moreover, the intercept  $\lambda_0$  is driven to zero in column (1) and both intercepts are zero in column (2). In other words, when we impose the Feller conditions in column (2), the estimation converges to the completely affine case, which reaffirms the economic sensibility of that kernel.

Next we investigate the economic implications of the specifications in Table 3 and Table 4. We show that pricing kernels that are difficult to distinguish statistically may have very different economic implications.

## 6 Pricing Kernels: Economic Implications

We first show that while model fit for most columns in Tables 3 and 4 is similar, risk premia and Sharpe ratios greatly vary with the specification of the pricing kernel. We then show how to estimate

the marginal pricing kernel that plots the pricing kernel as a function of log index returns. We show that this marginal pricing kernel is not necessarily downward sloping, which means that existing findings of non-monotonic kernels may not constitute puzzles. We also show how to construct the time series path of the pricing kernel and find that it also widely varies, dependent on the restrictions we impose on the pricing kernel. The path associated with the affine price of risk specification is not plausible. Lastly, we offer some observations on the relative importance of the pricing kernel components associated with the index return and variance state variables.

## 6.1 Risk Premia

A striking finding in Tables 3 and 4 is that the risk-neutral parameters are very similar across columns, and consequently the option RMSEs are identical. It is well known that in joint estimation, the option sample dominates the joint log likelihood because there are many options each day and only one return. Tables 3 and 4 show that this is not a problem in our sample because the component of the likelihood due to returns is large enough to create meaningful differences in the joint log likelihood. Recall that our return likelihood differs from most existing approaches because we exploit the model implication that the latent variance is a function of the VIX.<sup>18</sup>

We now turn to the economic implications of the different pricing kernel restrictions rather than the statistical fit. We find diametrically opposite results, namely that the different pricing kernels result in widely different economic implications. Panel B of Table 5 presents descriptive statistics for the equity and variance risk premia. Recall from equations (21), (22), and (17) that these risk premia are linear functions of the variance. This implies that the higher moments are the same across columns. However, the first two moments are very different. The annualized equity premium for the exponential-affine kernel in column (1) is 8.32%, which is very reasonable given the sample estimate of 8.24%. The average equity risk premium for the more general kernel in column (9) is very similar.

---

<sup>18</sup>We also repeated the estimation in Tables 3 and 4 with an option likelihood that is scaled back by the number of options so that the sample on each day effectively consists of one option and one return. Results for this estimation exercise are very similar.

In columns (2) and (6), the equity premium is zero by construction. Some of the other economic restrictions such as  $\lambda = 0$  in column (3),  $\gamma = 0$  in column (4), and  $\eta = 0$  in column (7) lead to substantially lower equity premia. The variance risk premium for the exponential-affine pricing kernel in column (1) of Panel B is equal to  $-0.0605$ , which amounts to  $-24.60\%$  in annual standard deviation terms. For the 1990 to 2019 period, the sample mean of the estimated volatility is  $18.28\%$ , and for the VIX it is  $20.73\%$  per year. The implied variance risk premium is therefore slightly larger than the index variance itself. This finding is similar to Bollerslev, Tauchen, and Zhou (2009). The results on the more general kernel in column (9) of Table 4 and the affine price of risk specification in column (8) yield similar estimates, but note that the signs of the (implied) risk premium parameters in Table 4 mean that the equity and variance risk premium do not always have the economically plausible signs. Imposing the restrictions in columns (2), (5) and (7) results in smaller (in absolute value) variance risk premia.

We conclude that while different pricing kernels lead to nearly identical option fit and risk-neutral parameters, as well as small differences in the joint log likelihood, they result in very different economic implications and risk premia. Bates (2003) argued that it is difficult to distinguish between option models based on option fit because misspecified models can fit options relatively well at the cost of overfitting and unreasonable out-of-sample implications. He therefore advocates joint estimation, which makes it easier to differentiate between models. While our estimation exercise in Tables 3 and 4 does not compare models with different dynamics and/or state variables, it does compare the specification of the pricing kernel, which is part of the model specification. Our results show that even when using a joint likelihood based on returns and options, it is difficult to statistically distinguish models. They are consistent with Bates' observation in the sense that misspecification instead shows up in implausible economic implications.

Another important conclusion is that the unrestricted exponential-affine pricing kernel in column (1) of Table 3 is economically plausible. Restricted versions of this kernel often result in implausible economic implications, as measured by risk premia and preference parameters. While these restrictions do not lead to a deterioration in option fit and are sometimes not statistically

rejected, these tests may lack power. An equivalent way to characterize our findings is that option pricing models can imply very different economic implications and risk premia, which cannot be distinguished statistically even when using returns and options in estimation. This finding is, to the best of our knowledge, novel and surprising, but it is consistent with the findings of Merton (1980) on the equity premium.

## 6.2 Sharpe Ratios

The parameter estimates allow us to retrieve the conditional mean and standard deviation of the daily market return. For exponential-affine pricing kernels, they are given by

$$E_t(R(t + \Delta)) = 1 + [r + \mu v(t)]\Delta, \quad (26)$$

$$\sigma_t(R(t + \Delta)) = \sqrt{v(t)}\Delta. \quad (27)$$

For the more general pricing kernels in equation (16) and the affine price of risk specification, the conditional mean of the daily market return is given by

$$E_t(R(t + \Delta)) = 1 + [r + \mu_0 + \mu_1 v(t)]\Delta. \quad (28)$$

We calculate the time series of the daily Sharpe ratios for the various kernel specifications based on these model-implied conditional means and standard deviations. The last row in Panel B of Table 5 reports the time-series averages of these daily Sharpe ratios. The unrestricted exponential-affine pricing kernel in column (1) implies an average daily Sharpe ratio of 0.0259, which is equivalent to an annualized Sharpe ratio of 0.411. This is close to the sample average of 0.439. The more general kernel in column (9) yields a Sharpe ratio close to the sample average, and the Sharpe ratio remains plausible when setting  $\xi = 0$  in columns (5) and (10). However, the average Sharpe ratios are not plausible when imposing some of the other restrictions. Importantly, the affine price of risk specification in column (8), which is successful in matching the average equity premium (see the discussion in Section 6.1), yields a Sharpe ratio which is too high. Moreover, the time series of

the Sharpe ratio in this case (not reported) is very different from the time series associated with columns (1) and (9). The time variation in the Sharpe ratio is much smaller in the case of the affine specification. Moreover, due to the intercept in the risk premium, the lowest observed Sharpe ratio in our sample period is 0.449, which is unrealistic and much higher than in columns (1) and (9).

### 6.3 A Multivariate Representation of the Estimated Pricing Kernel

Figure 2 provides a scatterplot of the estimated exponential-affine kernel in equation (20) as a function of the log stock return  $\log R(t)$  and the daily change in variance  $v(t) - v(t - 1)$ , based on the MLE estimates in column (1) of Table 3. The two bottom pictures illustrate the univariate relations. Figure 2 illustrates that the log return and the change in variance are negatively and positively correlated respectively with the log pricing kernel. This is, of course, due to the positive estimates of the index return preference  $\gamma$  and the variance preference  $\xi$ . The implied correlation coefficient  $\rho$  between the log return and the change in variance is clearly highly negative. Stock price increases or log return increases, therefore, reduce the pricing kernel directly through the channel of the positive  $\gamma$  and indirectly through the channel of the positive  $\xi$  combined with the negative  $\rho$ . An increase of the (change in) variance increases the pricing kernel directly and indirectly through the same mechanism.

Note that the univariate scatterplot on the left does not illustrate the marginal or the conditional distribution of the pricing kernel as a function of (log) returns. At each point in time, the relation between the pricing kernel and returns is captured for a different level of volatility. We therefore proceed to a more formal analysis of the relation between the pricing kernel and the state variables.

### 6.4 The Marginal Distribution of the Pricing Kernel

Following Jackwerth and Rubinstein (1996) and Jackwerth (2000), an extensive literature has investigated the shape of the pricing kernel. This literature is mainly motivated by the power utility pricing kernel in equation (25). Based on the kernel in equation (25), we expect a downward sloping log pricing kernel as a function of the log index return. However, starting with Jackwerth

(2000), several studies document that both the conditional and the unconditional pricing kernels are not downward sloping as a function of index returns (aggregate wealth), but instead are U-shaped or characterized by an even more nonlinear function (see, for example, the literature review in Cuesdeanu and Jackwerth, 2018). This finding is often referred to as a puzzle or an anomaly.

The literature has proposed several potential explanations for this anomaly. For instance, Bakshi, Madan, and Panayotov (2010) argue that a U-shaped kernel naturally arises from heterogeneous expectations. We instead propose that the pricing kernel will not generally be downward sloping when viewed as a function of aggregate wealth. Instead, given our knowledge of option prices, a sensible pricing kernel should contain several state variables. To the extent that these state variables are correlated with the index return, graphing state prices as a function of the log return may not result in a downward sloping function even if the kernel in equation (20) conforms to economic theory with respect to aggregate wealth, that is, if  $\gamma > 0$  in equation (20). A U-shaped pricing kernel, therefore, arises naturally and should not be thought of as an anomaly.

We investigate this conjecture empirically by characterizing the shape of the pricing kernel in the aggregate wealth dimension under different restrictions on the pricing kernel, i.e., for the different columns in Tables 3 and 4. We refer to the univariate pricing kernel in the aggregate wealth dimension as the marginal distribution of the pricing kernel. It is well-known that it is not straightforward to reliably estimate the pricing kernel. We need to characterize both the physical and risk-neutral distribution. Characterizing the risk-neutral distribution is relatively straightforward because of the abundance of option data on a given day, but characterizing the physical distribution is more challenging. The literature contains several non-parametric and parametric approaches. We proceed parametrically based on the parameter estimates in Tables 2, 3, and 4.

Given the parametric stock return dynamics in equations (1) and (2), we can generate the probability density of returns under both the physical and risk-neutral measures. For simplicity, we let the initial stock price  $S(0) = 1$  or equivalently  $\log S(0) = 0$ . Following Heston (1993), the



cumulative distribution function of log returns can be calculated as follows:

$$Pr(\log R(\tau) \leq x) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty Re \left[ \frac{e^{-i\phi x} f_\tau^{CH}(\phi \mid S(0) = 1, v(0) = v)}{i\phi} \right] d\phi,$$

where the characteristic function  $f_\tau^{CH}$  follows equation (9). We can calculate this characteristic function under both the physical and risk-neutral dynamics. By taking the first derivative of the cumulative distribution function with respect to  $x$ , we get the probability distribution function:

$$\begin{aligned} Pr(\log R(\tau) = x) &= \frac{dPr(\log R(\tau) \leq x)}{dx} \\ &= -\frac{1}{\pi} \int_0^\infty Re \left[ -e^{-i\phi x + C + Dv} \right] d\phi, \end{aligned} \quad (29)$$

where  $C$  and  $D$  are given in equation (10).

We calculate the numerical integral in equation (29) with the physical and risk-neutral dynamics of log returns to find the probability densities  $P(\log R(\tau) | v(0) = \theta)$  and  $Q(\log R(\tau) | v(0) = \theta)$ , respectively. We set  $v(0)$  at its unconditional mean level  $\theta$  and fix the risk-free rate at  $\bar{r} = 0.0261$  (2.61%), the sample mean for our 1990-2019 sample.

The pricing kernel is, by definition, the ratio of risk-neutral density to physical density, multiplied by the risk-free discount factor. We thus express the  $\tau$ -maturity log pricing kernel as follows:

$$\log M(\tau) - \log M(0) = -\bar{r}\tau + \log Q(\log R(\tau) | v(0) = \theta) - \log P(\log R(\tau) | v(0) = \theta). \quad (30)$$

By changing the value of  $\log R(\tau)$ , we can therefore generate the log pricing kernel as a function of the log index return. We repeat this process for various physical and risk-neutral parameter vectors, corresponding to the estimates for various restrictions on the pricing kernel in Tables 3 and 4. The right panels in Figures 3 and 4 report the results. For convenience, we express the x-axis in standardized log returns (the log return in deviation from its mean and divided by its standard

deviation).<sup>19</sup> Results are qualitatively similar across maturities, but not surprisingly the results are more pronounced for longer maturities. Figures 3 and 4 present the results for the implied six-month pricing kernels. For comparison, Figure A.1 in the Online Appendix presents additional results for one-month pricing kernels.

## 6.5 Recovering Pricing Kernel Dynamics

Most of the literature studies (differences in) pricing kernels using the equivalent of the right-side panels in Figures 3 and 4, that is, as a univariate function of the log return (aggregate wealth). However, our proposed pricing kernels (20) and (16) contain additional state variables. We now provide additional perspective on (differences between) pricing kernels by illustrating the impact of these state variables on the path of the pricing kernels.

The left panels in Figures 3 and 4 are obtained by inserting the observed realized returns and variances at each time  $t$  into the expression for the pricing kernel. These figures can be thought of as the (daily) time series path of the realized pricing kernel conditional on the realized values of the state variables. It is straightforward to generate this time path for all unrestricted and restricted specifications of the pricing kernel.<sup>20</sup> See Chernov (2003) for a related exercise.

## 6.6 Marginal Pricing Kernels and Pricing Kernel Dynamics: Empirical Results

Figure 3 presents results for five pricing kernels. The time path of the pricing kernel is on the left, and the marginal kernel is on the right. Panel A is based on the parameter estimates in Table 2, which do not impose any economic restrictions. Panel B of Figure 3 is based on the estimates

---

<sup>19</sup>Since we already have the physical probability density, we can calculate the expectation and variance of the log return as:

$$\begin{aligned}
 E(\log R(\tau)|v(0) = \theta) &= \int_{-\infty}^{\infty} \log R \times P(\log R(\tau)|v(0) = \theta) d\log R \\
 Var(\log R(\tau)|v(0) = \theta) &= \int_{-\infty}^{\infty} (\log R)^2 \times P(\log R(\tau)|v(0) = \theta) d\log R - E^2(\log R(\tau)|v(0) = \theta).
 \end{aligned}$$

<sup>20</sup>However, the computation is more complex in the case where the physical and risk-neutral parameters are independently estimated and for the affine price of risk specification. Online Appendix E discusses these cases.

in column (1) of Table 3, i.e., the exponential-affine pricing kernel. Recall that the completely affine price of risk specification is isomorphic to the unrestricted exponentially-affine pricing kernel. Panels C and D report on two special cases of column (1). Panel C imposes  $\gamma = 0$  and Panel D imposes  $\xi = 0$ , i.e., the power utility pricing kernel in equation (25). Finally, Panel E imposes path independence ( $\eta = 0$ )

We start with the results in Panels B, C and D. The differences between these three kernels are striking. Consider the marginal pricing kernels in the right column. By definition, the kernel in Panel D is linear in the log return space, and because the estimate of  $\gamma$  in column (5) of Table 3 is positive, it is downward sloping, consistent with the intuition of decreasing marginal utility of wealth. For the exponential-affine kernel in Panel B, state prices are no longer linearly downward sloping as a function of log aggregate wealth. Because the estimate of  $\xi$  is positive (and  $\rho$  is negative), we now have a convex function. However, the estimate of  $\xi$  is not high enough to generate a U-shaped pricing kernel. In Panel C on the other hand, column (4) of Table 3 indicates that we get a higher estimate of  $\xi$ , and we obtain a U-shaped pricing kernel.<sup>21</sup>

The dynamics of the time paths of the kernel (left column) for Panels C and D display some similarities with the exponential-affine kernel in Panel B. The realized kernels fluctuate a lot over time, especially in the 2008-2009 financial crisis period. There is also substantial variation around the 1998 LTCM crisis and the early 2000 recession. The most important difference between Panels B and D is that the time variation in the pricing kernel in Panel B is more pronounced in the second half of the sample. This is due to the fact that this time period contains more volatility outliers, as can be seen from Figure 1. Comparing Panels B and C, the outliers in Panel C are more pronounced due to the larger estimate of  $\xi$ .

Panel A of Figure 3 reports on the pricing kernel implied by the estimates in Table 2, i.e., using physical estimates obtained from returns and risk-neutral estimates obtained (separately) from options. These estimates do not impose restrictions on the pricing kernel, but they do of course impose a parametric structure on the probability distributions. The resulting pricing kernel

---

<sup>21</sup>Song and Xiu (2016) also emphasize how volatility risk can explain U-shaped pricing kernels. They use data on S&P 500 and VIX options and nonparametric estimation techniques.

has a W-shape. This is consistent with some existing findings (Cuesdeanu and Jackwerth, 2018), although a U-shape is more common.<sup>22</sup>

The dramatic differences between the marginal pricing kernel in Panel A and the ones in Panels B, C and D are confirmed by the time path of the pricing kernels in the left panels.<sup>23</sup> The only thing the path in Panel A has in common with the two others is that it varies most during the financial crisis. However, the differences between the financial crisis and the rest of the sample are much less pronounced in Panel A. More importantly, the outliers in the pricing kernel are of an entirely different order of magnitude compared to the exponential-affine pricing kernel in Panel B. We conclude that the time-series patterns in Panel B are implausible. This is due to the fact that the underlying P and Q estimates are not linked by economic assumptions.

Finally, the kernel that incorporates path-dependence (*i.e.* nonzero  $\eta$ ) generate very different paths compared to the path independent case ( $\eta = 0$ ) in Panel E of Figure 3 and strongly differ during periods when the stock variance abruptly increases or decreases. Under path-independence, the pricing kernel is not highly sensitive to changes in variance.<sup>24</sup>

Panel A in Figure 4 reports on the affine price of risk specification in column (8) of Table 4. Panel B reports on the more general pricing kernel in equation (16), corresponding to the estimates in column (9) of Table 4. Panel C imposes  $\xi = 0$ , as in column (10) of Table 4. Recall that the estimates in Panel A are obtained under the affine restriction that the equity premium is equal to  $\mu_0 + \mu_1 v(t)$  and the variance risk premium is equal to  $\lambda_0 + \lambda_1 v(t)$ . Existing results often impose these restrictions instead of the no-arbitrage restrictions, presumably because different no-arbitrage restrictions all result in the risk premia being linear in the variance. Our results suggest that this assumption is not innocuous. The marginal pricing kernel in Panel A is very different from the ones in Panels B and C. Note also the difference with the exponentially-affine kernel in Panel B of Figure 3; the shape is actually more similar to the unrestricted pricing kernel in Panel A of

---

<sup>22</sup>It is important to note that the estimation of the pricing kernel is much more precise for low returns. For high returns, confidence intervals are typically much larger due to more limited option information. We do not report confidence intervals in Figure 3, and moreover this issue is also impacted by imposing a parametric structure.

<sup>23</sup>Note that the  $y$ -axis in Panel B is scaled differently.

<sup>24</sup>This implication results from the zero  $\eta$  in conjunction with the negative estimate of  $\xi$ .

Figure 3. This finding is surprising given that Panel A of Figure 3 does not impose any restrictions across measures, while Panel A of Figure 4 does.

While the implied pricing kernels as a function of the index returns in Panel A of Figure 3 and Panel A of Figure 4 are similar, this is not the case for the corresponding time paths in the left column. Panel A of Figure 3 displays many more extreme positive and negative outliers throughout the sample. However, the time-series pattern in Panel A of Figure 4 also does not seem plausible. For instance, it is volatile during 1993-1996 and 2004-2007, whereas the fluctuations for the exponential-affine pricing kernel in Panel B of Figure 3 occur mainly in the financial crisis.

Panel D of Figure 4 is based on the parameter estimates in Table A.1 in the Online Appendix. It imposes the Feller condition on the kernel in equation (16). A comparison with Panel B of Figure 4 indicates that this restriction strongly affects the economic implications. Most notably, we now obtain a strongly U-shaped pricing kernel. However, the upward sloping part of this kernel is due to the counterintuitive negative sign of  $\gamma$ . This provides additional evidence that the exponential-affine kernel is preferable to the more general kernel (16).

Figures 3 and 4 yield six important conclusions: 1) We confirm that the unrestricted pricing kernel (based on the estimates in Table 2) is (highly) nonlinear as a function of aggregate wealth; 2) It is straightforward to write down economically meaningful pricing kernels (models) that can generate log pricing kernels that are U-shaped or non-monotonic in log aggregate wealth; 3) The exponential-affine pricing kernel has very plausible economic implications and is preferable to the more general kernel (16); 4) The affine price of risk specification provides a better fit to the option data, but its time path and other economic implications are implausible; 5) Imposing no-arbitrage restrictions is critically important, as evidenced for instance by the differences between Panels A and B of Figure 4; 6) Even when the log likelihoods in Tables 3 and 4 are similar, the time paths of the pricing kernels and the marginal kernels can be very different. While it seems to be difficult to statistically distinguish between models and/or pricing kernels, it may be easier to do so on economic grounds. This finding extends the observations of Merton (1980) on the estimation of the market equity premium to joint estimation of equity and variance risk premia using the cross-section

of options and the underlying returns.

## 6.7 Time-Variation in Pricing Kernel Components

In our final empirical exercise, we decompose the pricing kernel into its different components: the one related to the stock return ( $\gamma \log R(t)$ ) and to the change in variance ( $\xi[v(t) - v(t-1)]$ ). Figure 5 presents results for the exponential-affine specification in column (1) of Table 3. We omit the path-dependent component for convenience. It is small and very persistent. Figure 5 shows how the different components of the pricing kernel account for the overall variation of the unrestricted exponential-affine pricing kernel. We plot the cumulative log pricing kernel and its components over the sample period to emphasize the difference in the means of the components. Both the return and variance components are large at some points in time, due to their large standard deviations. However, Figure 5 shows that the cumulative index return component is much larger than the cumulative variance component, due to the fact that the mean of the return component is much larger than the mean of the variance component.

## 7 Conclusion

The pricing kernel is a critical concept in asset pricing. It governs the relationship between physical and risk-neutral probabilities at all times and for all return horizons, and ensures absence of arbitrage. Options can be used to estimate risk-neutral probabilities that are required to identify the pricing kernel, highlighting the importance of derivatives for asset pricing and financial economics.

This paper proposes pricing kernels that generalize the economically appealing kernels in Rubinstein (1976) and Brennan (1979). Economic intuition specifies how pricing kernels relate to relevant state variables and suggests that kernels should be monotonic, well-behaved and smooth functions of these state variables. The proposed kernels can be path-dependent and nest path-independent kernels (Ross, 2015) as special cases. They are volatility-dependent by construction and therefore characterize the impact of stock market volatility risk on state prices and risk premiums. We use these kernels to study the economic implications of the affine and completely affine price of risk

specifications. We show that affine risk premia can produce kernels with counter-intuitive economic properties and pricing kernels that are non-monotonic in index returns are not anomalous. A kernel consistent with the completely affine price of risk specification produces very plausible results.

We also estimate the resulting models subject to various restrictions on the pricing kernel. We reject the hypotheses of a zero equity or variance premium, but due to the negative correlation between index returns and variance we cannot statistically pin down the source of these risk premiums. More generally, we find that the data do not provide sufficient power: It is difficult to statistically distinguish between pricing kernels, even when they embody very different economic assumptions and generate widely different equity and variance risk premia and Sharpe ratios. These findings extend Merton's (1980) observations on the estimation of the market equity premium.

The analysis in this paper is based on a single-factor diffusion model. An obvious question is whether pricing kernels in more complex models are meaningfully different from pricing kernels in simpler models. Finally, our implementation uses plain-vanilla index options. It is possible that other option contracts may facilitate the estimation and identification of pricing kernels, equity and variance risk premia.

## References

- Aït-Sahalia, Yacine, and Robert Kimmel, 2007, Maximum Likelihood Estimation of Stochastic Volatility Models, *Journal of Financial Economics* 83, 413–452.
- Aït-Sahalia, Yacine, and Andrew W. Lo, 1998, Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, *Journal of Finance* 53, 499–547.
- Aït-Sahalia, Yacine, and Andrew W. Lo, 2000, Nonparametric Risk Management and Implied Risk Aversion, *Journal of Econometrics* 94, 9–51.
- Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015, Parametric Inference and Dynamic State Recovery from Option Panels, *Econometrica* 83, 1081–1145.
- Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2017, Short-Term Market Risks Implied by Weekly Options, *Journal of Finance* 72, 1335–1386.
- Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical Performance of Alternative Option Pricing Models, *Journal of Finance* 52, 2003–2049.
- Bakshi, Gurdip, John Crosby, and Xiaohui Gao, 2022, Dark Matter in (Volatility and) Equity Option Risk Premiums, *Operations Research* 70, 3108–3124.
- Bakshi, Gurdip, Dilip Madan, and George Panayotov, 2010, Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels, *Journal of Financial Economics* 97, 130–154.
- Bansal, Ravi, and Amir Yaron, 2004, Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 59, 1481–1509.
- Barone-Adesi, Giovanni, Nicola Fusari, Antonietta Mira, and Carlo Sala, 2020, Option Market Trading Activity and the Estimation of the Pricing Kernel: A Bayesian Approach, *Journal of Econometrics* 216, 430–449.



- Bates, David S., 2000, Post-'87 Crash Fears in the S&P 500 Futures Option Market, *Journal of Econometrics* 94, 181–238.
- Bates, David S., 2003, Empirical Option Pricing: A Retrospection, *Journal of Econometrics* 116, 387–404.
- Bates, David S., 2006, Maximum Likelihood Estimation of Latent Affine Processes, *Review of Financial Studies* 19, 909–965.
- Beason, Tyler, and David Schreindorfer, 2022, Dissecting the Equity Premium, *Journal of Political Economy* 130, 2203–2222.
- Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22, 4463–4492.
- Borovička, Jaroslav, Lars Peter Hansen, and José A. Scheinkman, 2016, Misspecified Recovery, *Journal of Finance* 71, 2493–2544.
- Breeden, Douglas T., and Robert H. Litzenberger, 1978, Prices of State-Contingent Claims Implicit in Option Prices, *Journal of Business* 51, 621–651.
- Brennan, Michael J., 1979, The Pricing of Contingent Claims in Discrete Time Models, *Journal of Finance* 34, 53–68.
- Brennan, Michael J., Xiaoquan Liu, and Yihong Xia, 2006, Option Pricing Kernels and the ICAPM, *Working Paper, University of California, Los Angeles*.
- Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model Specification and Risk Premium: Evidence from Futures Options, *Journal of Finance* 62, 1453–1490.
- Campbell, John Y., and John H. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, *Journal of Political Economy* 107, 205–251.

- Carr, Peter, and Dilip Madan, 1999, Option Valuation Using the Fast Fourier Transform, *Journal of Computational Finance* 2, 61–73.
- Carr, Peter, and Liuren Wu, 2007, Stochastic Skew in Currency Options, *Journal of Financial Economics* 86, 213–247.
- Chabi-Yo, Fousseni, 2012, Pricing Kernels with Stochastic Skewness and Volatility Risk, *Management Science* 58, 624–640.
- Chernov, Mikhail, 2003, Empirical Reverse Engineering of the Pricing Kernel, *Journal of Econometrics* 116, 329–364.
- Chernov, Mikhail, and Eric Ghysels, 2000, A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation, *Journal of Financial Economics* 56, 407–458.
- Chernov, Mikhail, Jeremy Graveline, and Irina Zviadadze, 2018, Crash Risk in Currency Returns, *Journal of Financial and Quantitative Analysis* 53, 137–170.
- Cheung, Sam, 2008, An Empirical Analysis of Joint Time-Series of Returns and the Term-Structure of Option Prices, *Working Paper, Columbia University*.
- Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2013, Capturing Option Anomalies with a Variance-Dependent Pricing Kernel, *Review of Financial Studies* 26, 1963–2006.
- Christoffersen, Peter, Kris Jacobs, and Karim Mimouni, 2010, Volatility Dynamics for the S&P500: Evidence from Realized Volatility, Daily Returns, and Option Prices, *Review of Financial Studies* 23, 3141–3189.
- Cuesdeanu, Horatio, and Jens Carsten Jackwerth, 2018, The Pricing Kernel Puzzle: Survey and Outlook, *Annals of Finance* 14, 289–329.
- Dew-Becker, Ian, and Stefano Giglio, 2022, Risk Preferences Implied by Synthetic Options, *Working Paper, Northwestern University and Yale University*.

- Drechsler, Itamar, 2013, Uncertainty, Time-Varying Fear, and Asset Prices, *Journal of Finance* 68, 1843–1889.
- Duffie, Darrell, and Larry G Epstein, 1992, Asset Pricing with Stochastic Differential Utility, *Review of Financial Studies* 5, 411–436.
- Epstein, Larry, and Stan Zin, 1989, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* 57, 937–969.
- Eraker, Bjørn, 2004, Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, *Journal of Finance* 59, 1367–1404.
- Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The Impact of Jumps in Volatility and Returns, *Journal of Finance* 58, 1269–1300.
- Eraker, Bjørn, and Ivan Shaliastovich, 2008, An Equilibrium Guide to Designing Affine Pricing Models, *Mathematical Finance* 18, 519–543.
- Eraker, Bjørn, and Yue Wu, 2017, Explaining the Negative Returns to VIX Futures and ETNs: An Equilibrium Approach, *Journal of Financial Economics* 125, 72–98.
- Eraker, Bjørn, and Aoxiang Yang, 2019, The Price of Higher Order Catastrophe Insurance: The Case of VIX Options, *Working Paper, University of Wisconsin-Madison*.
- Feller, William, 1951, Two Singular Diffusion Problems, *The Annals of Mathematics* 54, 173–182.
- Gabaix, Xavier, 2012, An Exactly Solved Framework for Ten Puzzles in Macro-Finance, *Quarterly Journal of Economics* 127, 645–700.
- Garcia, Rene, Richard Luger, and Eric Renault, 2003, Empirical Assessment of an Intertemporal Option Pricing Model with Latent Variables, *Journal of Econometrics* 116, 49–83.
- Ghosh, Anisha, Christian Julliard, and Alex P. Taylor, 2017, What is the Consumption-CAPM Missing? An Information-Theoretic Framework for the Analysis of Asset Pricing Models, *Review of Financial Studies* 30, 442–504.

- Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of Security Market Data for Models of Dynamic Economies, *Journal of Political Economy* 99, 225–262.
- Hansen, Lars Peter, and Ken Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.
- Heston, Steven L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 327–343.
- Heston, Steven L., Mark Loewenstein, and Gregory A. Willard, 2006, Options and Bubbles, *Review of Financial Studies* 20, 359–390.
- Heston, Steven L., and Saikat Nandi, 2000, A Closed-Form GARCH Option Valuation Model, *Review of Financial Studies* 13, 585–625.
- Hurn, A. Stan, Kenneth A. Lindsay, and Andrew J. McClelland, 2015, Estimating the Parameters of Stochastic Volatility Models Using Option Price Data, *Journal of Business & Economic Statistics* 33, 579–594.
- Jackwerth, Jens Carsten, 2000, Recovering Risk Aversion from Option Prices and Realized Returns, *Review of Financial Studies* 13, 433–451.
- Jackwerth, Jens Carsten, and Mark Rubinstein, 1996, Recovering Probability Distributions from Option Prices, *Journal of Finance* 51, 1611–1631.
- Jones, Christopher S., 2003, The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, *Journal of Econometrics* 116, 181–224.
- Kreps, D., and E. Porteus, 1978, Temporal Resolution of Uncertainty and Dynamic Choice Theory, *Econometrica* 46, 185–200.
- Linn, Matthew, Sophie Shive, and Tyler Shumway, 2018, Pricing Kernel Monotonicity and Conditional Information, *Review of Financial Studies* 31, 493–531.

- Liu, Jun, Jun Pan, and Tan Wang, 2005, An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks, *Review of Financial Studies* 18, 131–164.
- Mehra, Rajnish, and Edward Prescott, 1985, The Equity Premium Puzzle, *Journal of Monetary Economics* 15, 145–161.
- Merton, Robert C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica* 41, 867–887.
- Merton, Robert C., 1980, On Estimating the Expected Return on the Market: An Exploratory Investigation, *Journal of Financial Economics* 8, 323–361.
- Ooura, Takuya, 2001, A Continuous Euler Transformation and Its Application to the Fourier Transform of a Slowly Decaying Function, *Journal of Computational and Applied Mathematics* 130, 259–270.
- Pan, Jun, 2002, The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study, *Journal of Financial Economics* 63, 3–50.
- Qin, Likuan, and Vadim Linetsky, 2016, Positive Eigenfunctions of Markovian Pricing Operators: Hansen-Scheinkman Factorization, Ross Recovery, and Long-Term Pricing, *Operations Research* 64, 99–117.
- Rosenberg, Joshua V., and Robert F. Engle, 2002, Empirical Pricing Kernels, *Journal of Financial Economics* 64, 341–372.
- Ross, Steve, 2015, The Recovery Theorem, *Journal of Finance* 70, 615–648.
- Rubinstein, Mark, 1976, The Valuation of Uncertain Income Streams and the Pricing of Options, *Bell Journal of Economics* 7, 407–425.
- Seo, Sang Byung, and Jessica A. Wachter, 2019, Option Prices in a Model with Stochastic Disaster Risk, *Management Science* 65, 3449–3469.

- Shaliastovich, Ivan, 2015, Learning, Confidence, and Option Prices, *Journal of Econometrics* 187, 18–42.
- Singleton, Kenneth, 2006, *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*. (Princeton University Press Princeton, NJ).
- Song, Zhaogang, and Dacheng Xiu, 2016, A Tale of Two Option Markets: Pricing Kernels and Volatility Risk, *Journal of Econometrics* 190, 176–196.
- Trolle, Anders B., and Eduardo S. Schwartz, 2009, Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives, *Review of Financial Studies* 22, 4423–4461.
- Wachter, Jessica A., 2013, Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?, *Journal of Finance* 68, 987–1035.

Table 1: Descriptive Statistics

Panel A: Option Data by Moneyness									
	$F/K \leq 0.94$	$0.94 < F/K \leq 0.98$	$0.98 < F/K \leq 1.02$	$1.02 < F/K \leq 1.06$	$1.06 < F/K \leq 1.10$	$F/K > 1.10$	All		
Number of contracts	6,718	9,227	15,133	11,975	7,928	11,502	62,483		
Average IV (%)	16.21	13.65	15.10	17.97	20.98	25.09	18.14		
Average price	14.64	20.28	37.67	26.70	21.73	17.36	24.76		
Panel B: Option Data by Maturity									
	$DTM \leq 30$	$30 < DTM \leq 60$	$60 < DTM \leq 90$	$90 < DTM \leq 120$	$120 < DTM \leq 180$	$DTM > 180$	All		
Number of contracts	13,969	14,483	8,482	5,374	7,143	13,032	62,483		
Average IV (%)	16.16	18.08	19.05	19.10	18.77	18.99	18.14		
Average price	9.01	14.14	21.31	27.16	32.69	50.36	24.76		
Panel C: Stock Returns (annualized)									
	1990-2019	1996-2019							
Mean (%)	9.31	8.56							
Standard deviation (%)	17.51	18.75							
Skewness	-0.26	-0.26							
Kurtosis	11.81	11.09							

Notes: Panels A and B present descriptive statistics for Wednesday closing OTM option contracts for the January 10, 1996 to June 26, 2019 period. Moneyness is defined as the implied futures price  $F = Se^{(r-q)\tau}$  divided by the strike price  $K$ . Panel C reports on the log of daily index returns for the January 1, 1990 to June 30, 2019 and January 1, 1996 to June 30, 2019 periods. Mean and standard deviation are annualized, skewness and kurtosis are computed from daily returns. Source: Authors' calculations based on OptionMetrics and CRSP.

Table 2: Return-Based and Option-Based Parameter Estimates

Panel A: Return-Based Physical Parameters							
$\mu$	$\kappa$	$\theta$	$\sigma$	$\rho$	$\eta_0$	$\eta_1$	$\log L^R$
2.6367	4.1457	0.0310	0.5309	-0.6877	-0.0068	0.8815	55,409
(1.0624)	(1.2397)	(0.0070)	(0.0242)	(0.0114)	(0.0003)	(0.0219)	
Panel B: Option-Based Risk-Neutral Parameters							
$\mu^*$	$\kappa^*$	$\theta^*$	$\sigma$	$\rho$	$\eta_0$	$\eta_1$	$\log L^O$
0	0.9860	0.0986	0.7916	-0.7452	-0.0042	0.8740	156,831
	(0.0712)	(0.0050)	(0.0059)	(0.0021)	(0.0002)	(0.0072)	

Notes: Panel A presents the physical parameters estimated using index returns and the VIX. Panel B presents the risk-neutral parameters estimated using option prices and the VIX. Robust standard errors are in parentheses.



Table 3: Joint MLE Estimation 1990-2019: Exponential-Affine Specifications

Restrictions	Exponential-Affine						
	None (1)	$\mu = 0$ (2)	$\lambda = 0$ (3)	$\gamma = 0$ (4)	$\xi = 0$ (5)	$\gamma = \xi = 0$ (6)	$\eta = 0$ (7)
Risk Preference Parameters							
$\gamma$	1.3929 (0.2266)	-1.0921	1.3940	0	2.4850 (0.2182)	0	2.2311
$\xi$	1.9474 (0.4037)	1.9472 (0.2034)	-1.4776 (0.8098)	3.4244 (0.5430)	0	0	-0.6242 (0.0511)
Implied Risk Premia Parameters							
$\mu$	2.4852	0	0.5652	1.9206	2.4850	0	1.8810
$\lambda$	-1.8115	-0.4177	0	-1.8116	-1.3937	0	-0.9211
P-Parameters							
$\kappa$	2.9252 (0.2082)	1.5314 (0.0727)	1.1141 (0.0558)	2.9252 (0.2933)	2.5076 (0.1255)	1.1139 (0.0558)	2.0352 (0.0855)
$\theta$	0.0334 (0.0024)	0.0638 (0.0023)	0.0877 (0.0030)	0.0334 (0.0033)	0.0390 (0.0019)	0.0877 (0.0030)	0.0480 (0.0012)
$\sigma$	0.7274 (0.0082)	0.7274 (0.0083)	0.7274 (0.0083)	0.7273 (0.0083)	0.7274 (0.0083)	0.7274 (0.0083)	0.7274 (0.0083)
$\rho$	-0.7711 (0.0039)	-0.7711 (0.0039)	-0.7711 (0.0039)	-0.7711 (0.0039)	-0.7711 (0.0039)	-0.7711 (0.0039)	-0.7711 (0.0039)
$\eta_0$	-0.0047 (0.0002)	-0.0047 (0.0002)	-0.0047 (0.0002)	-0.0047 (0.0002)	-0.0047 (0.0002)	-0.0047 (0.0002)	-0.0047 (0.0002)
$\eta_1$	0.8881 (0.0059)	0.8881 (0.0060)	0.8881 (0.0060)	0.8880 (0.0059)	0.8881 (0.0060)	0.8881 (0.0060)	0.8881 (0.0060)
Q-Parameters							
$\kappa^*$	1.1137	1.1137	1.1141	1.1135	1.1139	1.1139	1.1141
$\theta^*$	0.0877	0.0877	0.0877	0.0877	0.0877	0.0877	0.0877
Model Fit							
$\log L^R$	54,443.42	54,440.38	54,440.30	54,443.11	54,443.02	54,439.98	54,442.60
$\log L^O$	156,578.03	156,578.03	156,578.10	156,577.96	156,578.04	156,578.03	156,577.97
$\log L^R + \log L^O$	211,021.45	211,018.41	211,018.40	211,021.07	211,021.05	211,018.01	211,020.58
Vega-Weighted Option Price RMSE	0.0197	0.0197	0.0197	0.0197	0.0197	0.0197	0.0197

Notes: This table reports the results of joint maximum likelihood estimation. The risk preference parameters  $\gamma$  and  $\xi$  as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions in columns (1) to (7). Robust standard errors are in parentheses. We estimate the model for seven different specifications. We report the log likelihood and the vega-weighted option price RMSE for each specification.

Table 4: Joint MLE Estimation 1990-2019: Other Specifications

Restrictions	Affine POR	PK in Eq. (16)	
	None (8)	None (9)	$\xi = 0$ (10)
Risk Preference Parameters			
$\gamma$		1.4059 (0.2282)	2.5693 (0.1980)
$\xi$		2.9002 (0.2599)	0
$\alpha$		-0.0324 (0.0048)	-0.0050 (0.0004)
Implied Risk Premia Parameters			
$\mu_0$	0.0485 (0.0202)	-0.0182	-0.0028
$\mu_1$	1.0393 (1.2862)	3.0324	2.5693
$\lambda_0$	0.0172 (0.0054)	0.0172	0.0027
$\lambda_1$	-2.3297 (0.5833)	-2.3227	-1.4410
P-Parameters			
$\kappa$	3.4423 (0.5790)	3.4361 (0.1714)	2.5549 (0.1165)
$\theta$	0.0334 (0.0072)	0.0334 (0.0015)	0.0393 (0.0018)
$\sigma$	0.7272 (0.0083)	0.7273 (0.0083)	0.7274 (0.0083)
$\rho$	-0.7711 (0.0039)	-0.7711 (0.0039)	-0.7711 (0.0039)
$\eta_0$	-0.0047 (0.0002)	-0.0047 (0.0002)	-0.0047 (0.0002)
$\eta_1$	0.8880 (0.0059)	0.8880 (0.0060)	0.8881 (0.0059)
Q-Parameters			
$\kappa^*$	1.1126	1.1134	1.1139
$\theta^*$	0.0878	0.0877	0.0877
Model Fit			
$\log L^R$	54,449.61	54,443.71	54,443.03
$\log L^O$	156,577.08	156,578.00	156,578.03
$\log L^R + \log L^O$	211,026.69	211,021.71	211,021.06
Vega-Weighted Option Price RMSE	0.0197	0.0197	0.0197

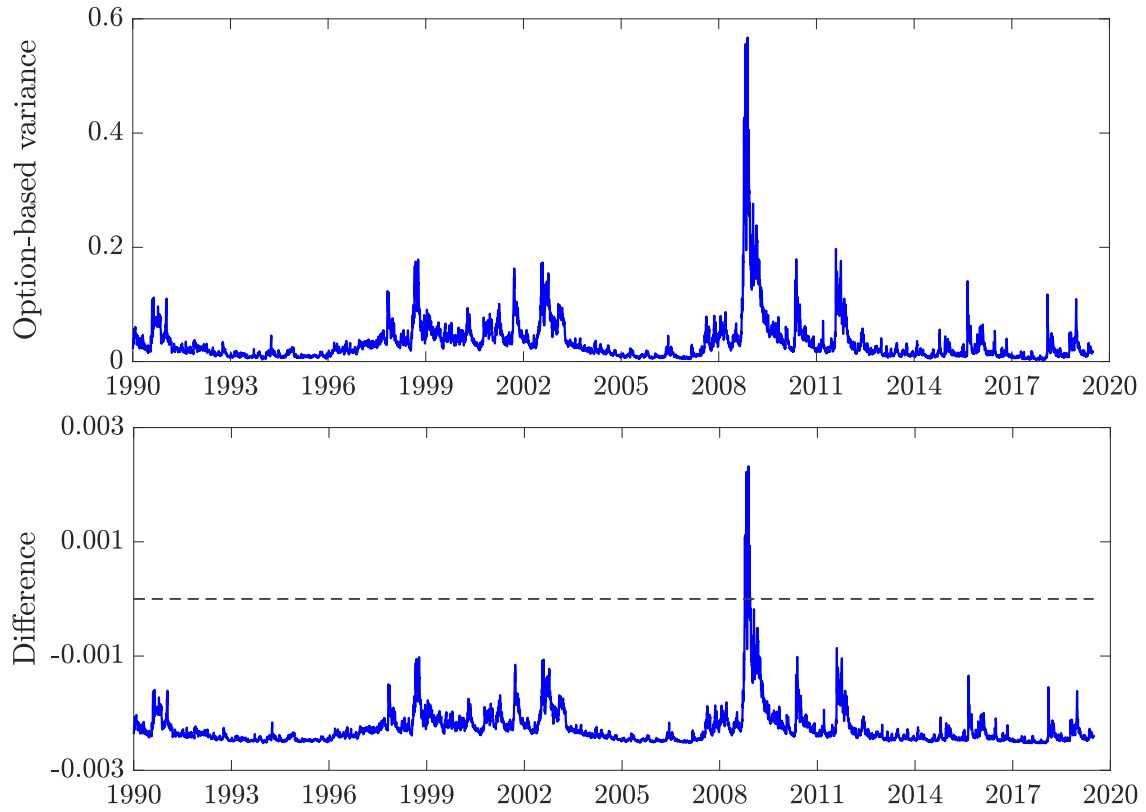
Notes: This table reports the results of joint maximum likelihood estimation. The risk preference parameters  $\gamma$ ,  $\xi$ , and  $\alpha$  as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions. Robust standard errors are in parentheses. We report on the affine price of risk (POR) specification in column (8), the more general pricing kernel (PK) from equation (16) in column (9), and the more general PK subject to  $\xi = 0$  in column (10).

Table 5: Model Comparison

Panel A. Likelihood Ratio Test: P-Values										
Restrictions	Exponential-Affine					Affine POR		PK in Eq. (16)		
	None (1)	$\mu = 0$ (2)	$\lambda = 0$ (3)	$\gamma = 0$ (4)	$\xi = 0$ (5)	$\gamma = \xi = 0$ (6)	$\eta = 0$ (7)	None (8)	None (9)	$\xi = 0$ (10)
Comparison with (1)		0.0137	0.0135	0.3833	0.3739	0.0321	0.1862			
Comparison with (8)	0.0053	0.0009	0.0009	0.0105	0.0103	0.0016	0.0066		0.0016	0.0036
Comparison with (9)	0.4708	0.0369	0.0365	0.5273	0.5192	0.0602	0.3218			0.2542
Panel B. Economic Implications										
Instantaneous Equity Premium	0.0832	0	0.0189	0.0643	0.0832	0	0.0629	0.0833	0.0833	0.0832
Instantaneous Variance Premium	-0.0606	-0.0140	0	-0.0606	-0.0466	0	-0.0308	-0.0607	-0.0606	-0.0456
Annualized Sharpe Ratio	0.4111	0	0.0937	0.3175	0.4111	0	0.3111	0.5223	0.3746	0.4048

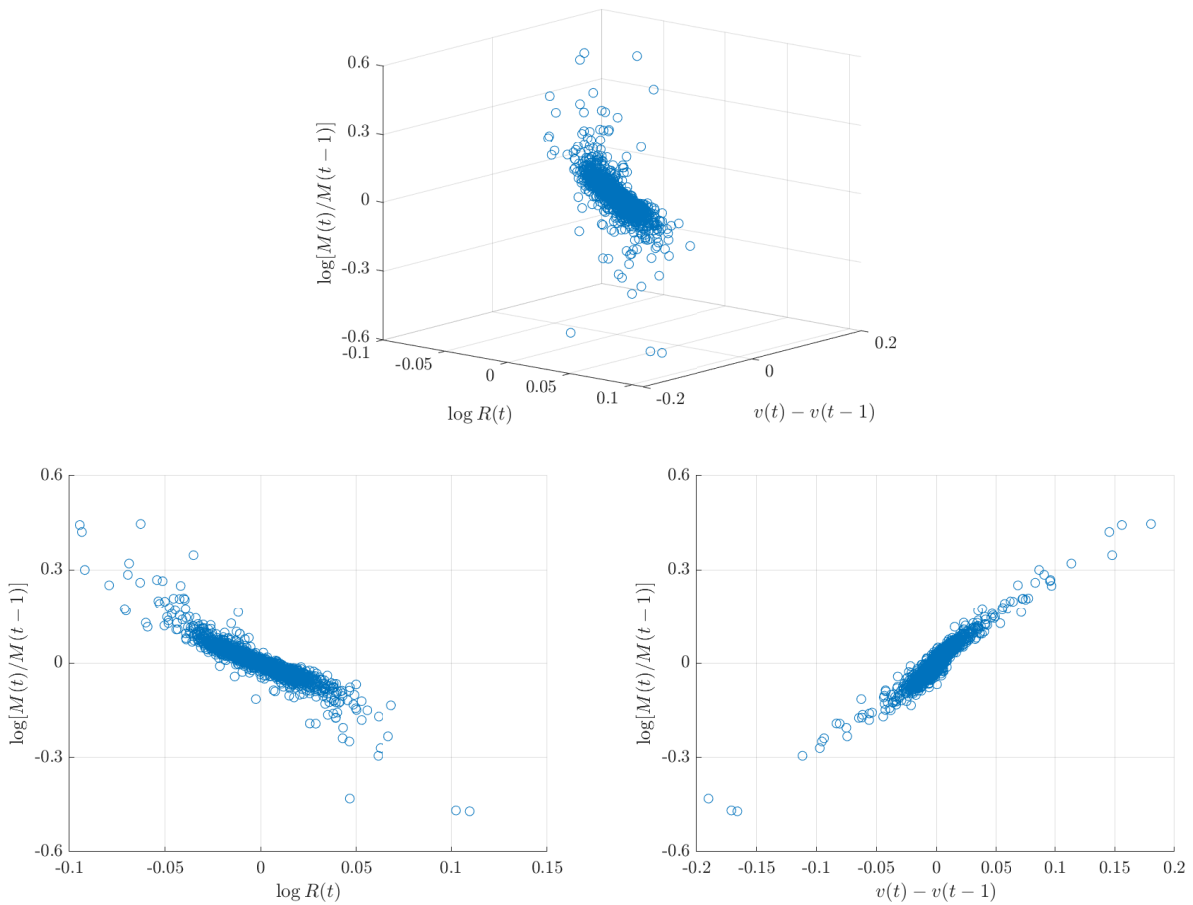
Notes: Panel A reports the P-values for the Likelihood Ratio tests against the specification in columns (1), (8), and (9). Minus two times the difference in the log-likelihood is distributed chi-square with degrees of freedom equal to the number of restrictions. Panel B presents the sample means for the implied instantaneous equity and variance risk premia and the annualized Sharpe ratio.

Figure 1: Instantaneous Variance, 1990 - 2019



Notes: We plot the time series of the stochastic variance from the option-based estimation and the difference between the return-based and the option-based variance. The return-based variance is computed as  $v(t) = \eta_0 + \eta_1 \text{VIX}^2(t)$ , the option-based variance is computed as  $v^*(t) = \eta_0^* + \eta_1^* \text{VIX}^2(t)$ , and the difference is computed as  $v(t) - v^*(t)$ , where  $\eta_0$ ,  $\eta_1$ ,  $\eta_0^*$ , and  $\eta_1^*$  are from Table 2.

Figure 2: The Log Pricing Kernel as a Function of Log Return and Variance

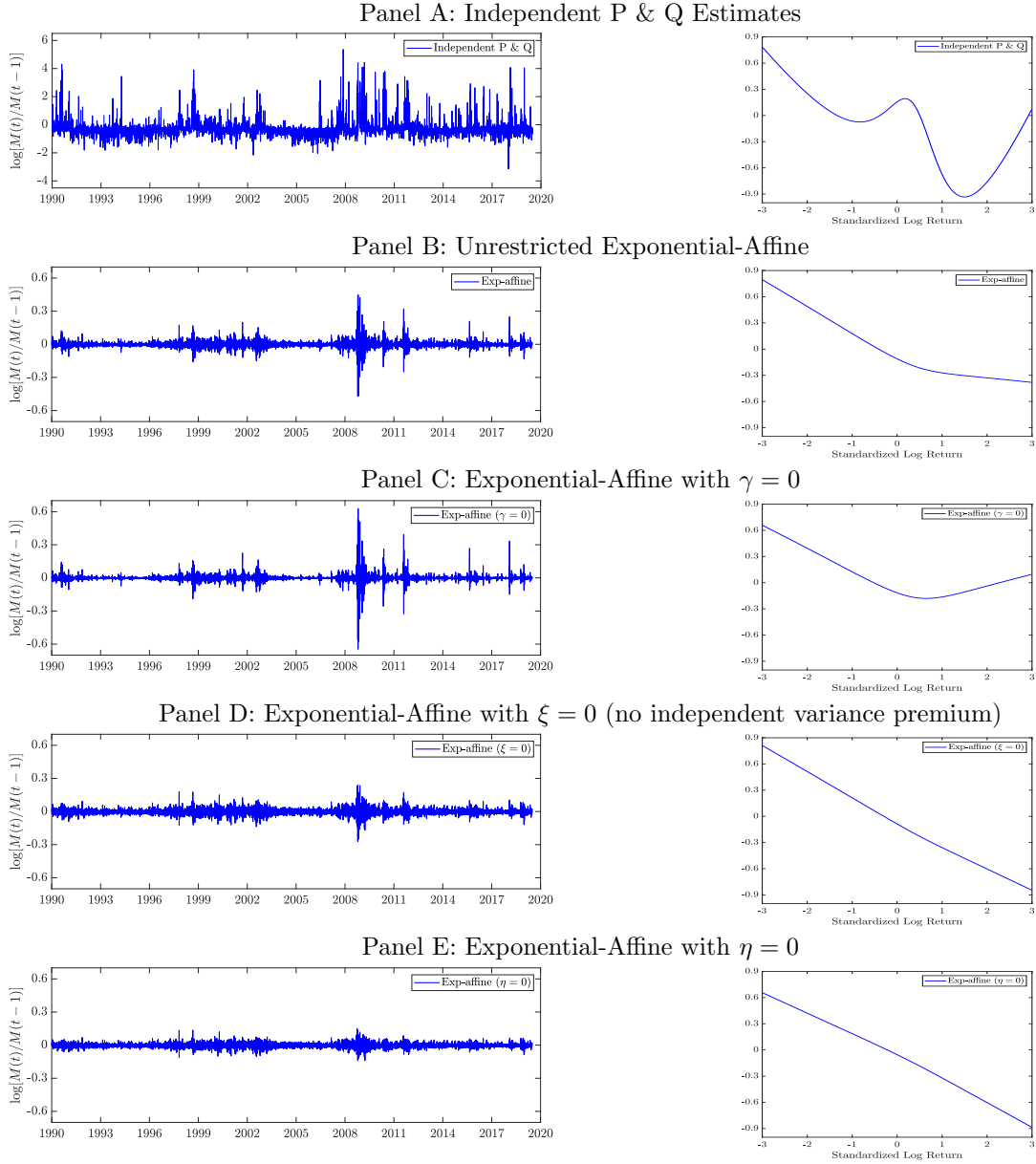


Notes: We plot the multivariate and univariate relation between the log pricing kernel, the log stock return ( $\log R(t)$ ), and the daily change in variance ( $v(t) - v(t-1)$ ). The results are based on the exponential-affine pricing kernel, using the parameters in column (1) of Table 3.

Figure 3: Exponentially-Affine Log Pricing Kernels

Daily (Realized) Log Pricing Kernels 1990-2019

Implied Six-Month Log Marginal Pricing Kernels

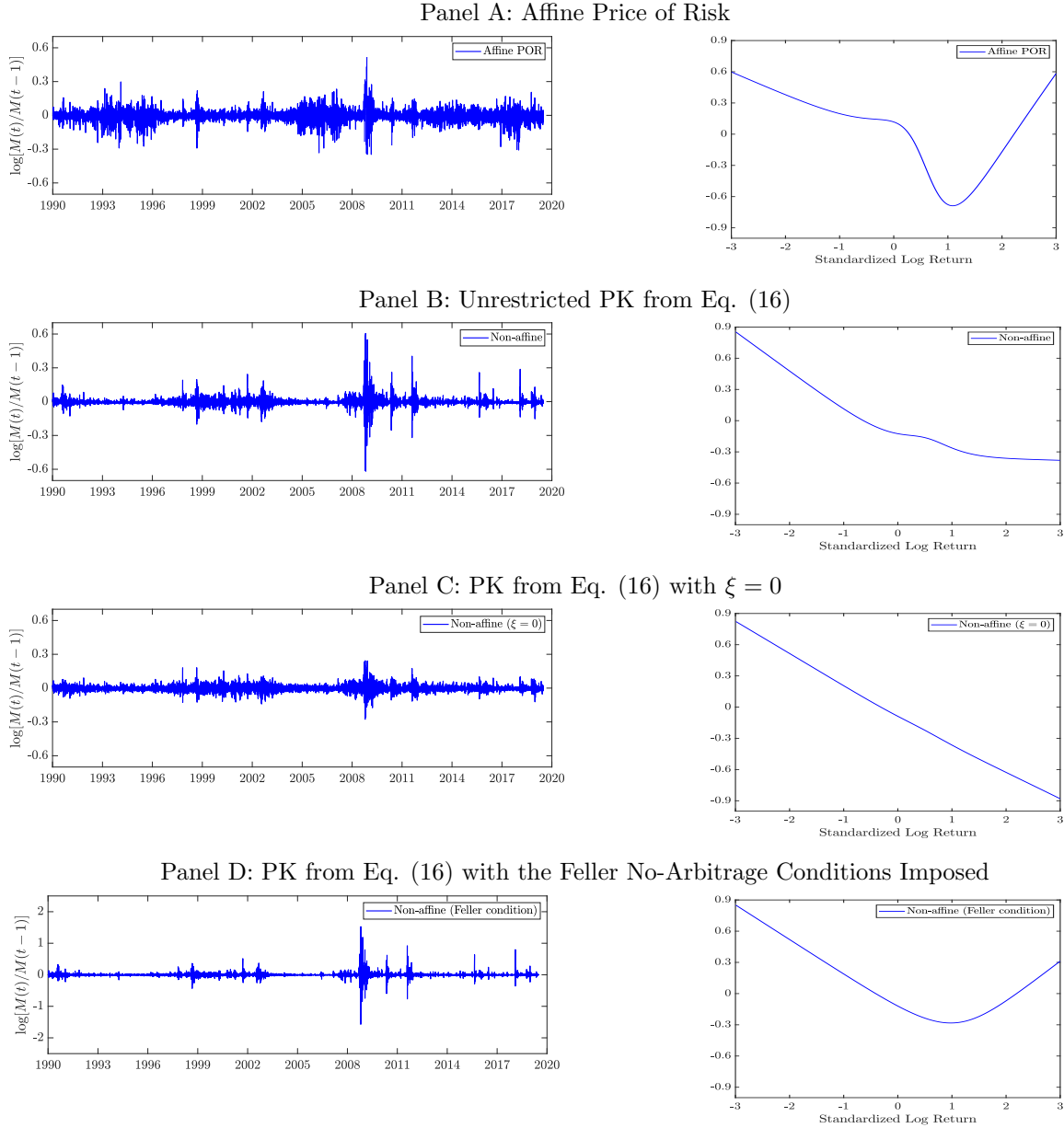


Notes: We plot the time series of the daily log pricing kernels and the implied 6-month log marginal pricing kernels for the following five specifications: independent estimation of the P- and Q-parameters (Panel A), the unrestricted exponential-affine (Panel B), the exponential-affine specifications with a restriction  $\gamma = 0$  (Panel C),  $\xi = 0$  (Panel D), and  $\eta = 0$  (Panel E). Parameter values for Panel A are from Table 2, and those for Panels B, C, D, and E are from Table 3. For the implied 6-month kernels, the  $x$ -axis represents the log return in deviation from its mean and divided by its standard deviation.

Figure 4: Other Log Pricing Kernels

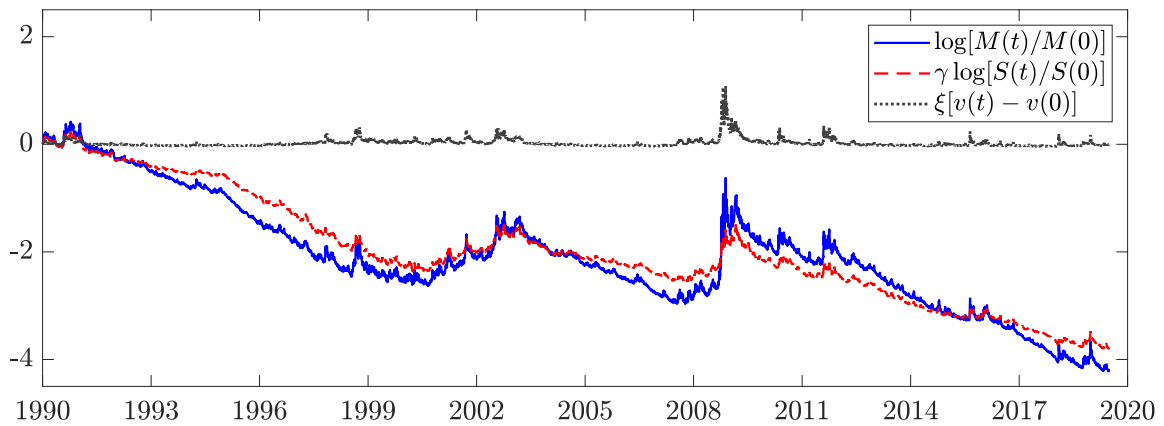
Daily (Realized) Log Pricing Kernels 1990-2019

Implied Six-Month Log Marginal Pricing Kernels



Notes: We plot the time series of the daily log pricing kernels and the implied 6-month log marginal pricing kernels for the following four specifications: the affine price of risk specification (Panel A), the unrestricted pricing kernel from equation (16) (Panel B), the kernel from equation (16) with restriction  $\xi = 0$  imposed (Panel C), and the pricing kernel from equation (16) with the Feller condition imposed (Panel D). Parameter values for Panels A, B, and C are from Table 4, and those for Panel D are from Table A.1. For the implied 6-month kernels, the  $x$ -axis represents the log return in deviation from its mean and divided by its standard deviation.

Figure 5: Cumulative Log Pricing Kernel Components 1990 - 2019



Notes: We plot the cumulative (over time) logarithm of the exponential-affine pricing kernel based on the estimates in column (1) of Table 3. This figure contains the entire pricing kernel and the components of the pricing kernel associated with the index price and its instantaneous volatility.



## Online Appendix

### A The Pricing Kernel and Risk Premia

Given the pricing kernel in equation (11) or (13), we use the fact that  $U(t)M(t)$  is a martingale for any asset  $U$  to calculate risk premia. Applying Ito's lemma to the martingale condition  $E[d(U(t)M(t))] = 0$  shows that expected returns are determined by their covariance with the pricing kernel:

$$\frac{E[d(U(t)M(t))]}{U(t)M(t)} = E\left[\frac{dU(t)}{U(t)}\right] + E\left[\frac{dM(t)}{M(t)}\right] + Cov\left[\frac{dU(t)}{U(t)}, \frac{dM(t)}{M(t)}\right] = 0. \quad (\text{A.1})$$

Inserting a bond price  $U(t) = \exp(-r(T-t))$  into equation (A.1) provides an expression for the interest rate

$$E\left[\frac{dM(t)}{M(t)}\right] = -r dt.$$

Ito's lemma then identifies the equity and variance risk premia in terms of their covariances with the pricing kernel:

$$E\left[\frac{dS(t)}{S(t)}\right] - r dt = -Cov\left[\frac{dS(t)}{S(t)}, \frac{dM(t)}{M(t)}\right], \quad (\text{A.2})$$

$$E[dv(t)] - E^*[dv(t)] = -Cov\left[dv(t), \frac{dM(t)}{M(t)}\right]. \quad (\text{A.3})$$

Using equations (A.2) and (A.3), it can be shown that the pricing kernel with the completely affine prices of risk in equation (11) is consistent with a completely affine equity risk premium of  $\mu v(t)$  and variance risk premium of  $\lambda v(t)$ , where

$$\mu = \pi_1 + \rho\pi_2 \quad \text{and} \quad \lambda = \rho\sigma\pi_1 + \sigma\pi_2. \quad (\text{A.4})$$

Likewise, the pricing kernel with the affine prices of risk in equation (13) is consistent with affine

risk premiums  $\mu_0 + \mu_1 v(t)$  and  $\lambda_0 + \lambda_1 v(t)$ , where

$$\mu_0 = \pi_{1,0} + \rho\pi_{2,0}, \quad \mu_1 = \pi_{1,1} + \rho\pi_{2,1}, \quad \lambda_0 = \rho\sigma\pi_{1,0} + \sigma\pi_{2,0}, \quad \lambda_1 = \rho\sigma\pi_{1,1} + \sigma\pi_{2,1}. \quad (\text{A.5})$$

## B No-Arbitrage Restrictions for the More General Pricing Kernel

We use the risk-neutral dynamics in equation (1) and physical dynamics in equation (B.1).

$$\begin{aligned} d \log S(t) &= \left[ r - \frac{1}{2}v(t) + \mu_0 + \mu_1 v(t) \right] dt + \sqrt{v(t)} dz_1(t) \\ dv(t) &= \kappa(\theta - v(t))dt + \sigma\sqrt{v(t)} dz_2(t) \end{aligned} \quad (\text{B.1})$$

Applying Ito's lemma to equation (16) gives

$$\begin{aligned} d \log M &= -\gamma d \log S + \alpha d \log v + \beta dt + \eta(v)dt + \xi dv \\ &= \left[ -\gamma \left( r + \mu_0 + \mu_1 v - \frac{1}{2}v \right) + \beta + \eta(v) + \xi \kappa(\theta - v) + \frac{\alpha}{v} \left( \kappa(\theta - v) - \frac{1}{2}\sigma^2 \right) \right] dt \\ &\quad - \gamma\sqrt{v} dz_1 + \left( \xi\sigma\sqrt{v} + \frac{\alpha\sigma}{\sqrt{v}} \right) dz_2, \end{aligned}$$

where we use the fact that  $d \log v = \frac{1}{v} (\kappa(\theta - v) - \frac{1}{2}\sigma^2) dt + \frac{\sigma}{\sqrt{v}} dz_2$ . Again by Ito's lemma, we have

$$\begin{aligned} \frac{dM}{M} &= \left[ -\gamma \left( r + \mu_0 + \mu_1 v - \frac{1}{2}v \right) + \beta + \eta(v) + \xi \kappa(\theta - v) + \frac{\alpha}{v} \left( \kappa(\theta - v) - \frac{1}{2}\sigma^2 \right) \right. \\ &\quad \left. + \frac{1}{2} \left( \gamma^2 v + \xi^2 \sigma^2 v + \frac{\alpha^2 \sigma^2}{v} \right) - \gamma \xi \sigma \rho v - \gamma \alpha \sigma \rho + \xi \alpha \sigma^2 \right] dt \\ &\quad - \gamma\sqrt{v} dz_1 + \left( \xi\sigma\sqrt{v} + \frac{\alpha\sigma}{\sqrt{v}} \right) dz_2. \end{aligned} \quad (\text{B.2})$$

The drift of  $M$  should be equal to  $-rMdt$ . By rearranging the drift term in equation (B.2),

$$\begin{aligned} &[r - \gamma(r + \mu_0) + \xi\kappa\theta - \alpha\kappa - \gamma\alpha\sigma\rho + \xi\alpha\sigma^2] + \beta \\ &+ \left[ -\gamma \left( \mu_1 - \frac{1}{2} \right) - \xi\kappa + \frac{1}{2} (\gamma^2 + \xi^2 \sigma^2 - 2\gamma\xi\sigma\rho) \right] v + \left[ \alpha\kappa\theta - \frac{1}{2}\alpha\sigma^2 + \frac{1}{2}\alpha^2\sigma^2 \right] \frac{1}{v} + \eta(v) = 0. \end{aligned}$$

Thus, the time-preference and the stochastic path-dependence terms are expressed as

$$\begin{aligned}\beta &= -(1-\gamma)r + \gamma\mu_0 - \xi\kappa\theta + \alpha\kappa + \gamma\alpha\sigma\rho - \xi\alpha\sigma^2 \\ \eta(v) &= \left[ \left( \mu_1 - \frac{1}{2} \right) \gamma + \xi\kappa - \frac{1}{2} (\gamma^2 - 2\gamma\xi\sigma\rho + \xi^2\sigma^2) \right] v - \left[ \kappa\theta - \frac{1}{2}\sigma^2 + \frac{1}{2}\alpha\sigma^2 \right] \frac{\alpha}{v}.\end{aligned}$$

Moreover, by equations (A.5) and (B.2), we obtain the risk premia expressed as  $\mu_0 + \mu_1 v(t)$  and  $\lambda_0 + \lambda_1 v(t)$ , where

$$\mu_0 = -\alpha\sigma\rho, \quad \mu_1 = \gamma - \rho\sigma\xi, \quad \lambda_0 = -\alpha\sigma^2, \quad \lambda_1 = \gamma\rho\sigma - \xi\sigma^2.$$

The physical dynamics in equation (B.1) are restricted according to

$$\kappa = \kappa^* - \lambda_1 \quad \text{and} \quad \theta = (\kappa^*\theta^* + \lambda_0)/\kappa.$$

## C No-Arbitrage Restrictions for the Exponential-Affine Pricing Kernel

Applying Ito's lemma to equation (20) gives

$$\begin{aligned}d \log M &= -\gamma d \log S + \beta dt + \eta v dt + \xi dv \\ &= \left[ -\gamma \left( r + \mu v - \frac{1}{2} v \right) + \beta + \eta v + \xi \kappa (\theta - v) \right] dt - \gamma \sqrt{v} dz_1 + \xi \sigma \sqrt{v} dz_2,\end{aligned}$$

where we drop the time- $t$  dependence ( $t$ ) of  $M$ ,  $S$ ,  $v$ ,  $z_1$ , and  $z_2$  for notational convenience. Again by Ito's lemma, we get

$$\begin{aligned}\frac{dM}{M} &= \left[ -\gamma \left( r + \mu v - \frac{1}{2} v \right) + \beta + \eta v + \xi \kappa (\theta - v) + \frac{1}{2} \gamma^2 v - \gamma \xi \sigma \rho v + \frac{1}{2} \xi^2 \sigma^2 v \right] dt \\ &\quad - \gamma \sqrt{v} dz_1 + \xi \sigma \sqrt{v} dz_2.\end{aligned}\tag{C.1}$$

In order to find the restrictions on  $\beta$  and  $\eta$ , we use the fact that  $e^{rt}M(t)$  is a martingale – the drift

term of  $d(e^{rt}M(t))$  must be zero. That is, from equation (C.1),

$$r - \gamma \left( r + \mu v - \frac{1}{2}v \right) + \beta + \eta v + \xi \kappa (\theta - v) + \frac{1}{2}\gamma^2 v - \gamma \xi \sigma \rho v + \frac{1}{2}\xi^2 \sigma^2 v = 0.$$

By rearranging this equation,

$$\underbrace{[\beta + (1 - \gamma)r + \xi \kappa \theta]}_{(A)} + \underbrace{\left[ \eta - \gamma \mu + \frac{1}{2}\gamma - \xi \kappa + \frac{1}{2}(\gamma^2 - 2\gamma \xi \sigma \rho + \xi^2 \sigma^2) \right]}_{(B)} v = 0. \quad (C.2)$$

Since equation (C.2) must hold for any values of  $v$ , it implies that (A) and (B) must be zero.

Therefore, we have the following restrictions on  $\beta$  and  $\eta$ :

$$\begin{aligned} \beta &= -(1 - \gamma)r - \xi \kappa \theta \\ \eta &= \gamma \mu - \frac{1}{2}\gamma + \xi \kappa - \frac{1}{2}(\gamma^2 - 2\gamma \xi \sigma \rho + \xi^2 \sigma^2). \end{aligned}$$

Furthermore, by equations (A.4) and (C.1), we find the expression for the equity and variance risk premium parameters  $\mu$  and  $\lambda$  as functions of the preference parameters:

$$\mu = \gamma - \rho \sigma \xi \quad \text{and} \quad \lambda = \gamma \rho \sigma - \xi \sigma^2.$$

We can now deduce the relations between the physical and risk-neutral parameters of the variance dynamics. With the variance risk premium  $\lambda v$ , the risk-neutral dynamics of variance should be

$$dv(t) = [\kappa(\theta - v(t)) - \lambda v(t)]dt + \sigma \sqrt{v(t)} dz_2^*(t). \quad (C.3)$$

Comparing equation (1) with (C.3) implies the following restrictions:

$$\kappa = \kappa^* - \lambda \quad \text{and} \quad \theta = \kappa^* \theta^* / \kappa.$$

## D Parameter Restrictions in the Joint Option and Returns Likelihood

The following table summarizes the restrictions used when estimating the joint likelihoods in Tables 3 and 4. As mentioned above, for the case in column (8) there are no restrictions. We implement these cases by letting either  $\gamma$  or  $\xi$  be a free parameter, while the other one is implied by the restriction(s). The resulting mapping between  $\gamma$  and  $\xi$  is reported in the last column.

Pricing Kernel	Restriction	Free parameter	Fixed parameter
(2)	$\mu = 0$	$\xi$	$\gamma = \xi\sigma\rho$
(3)	$\lambda = 0$	$\xi$	$\gamma = \sigma\xi/\rho$
(4)	$\gamma = 0$	$\xi$	$\gamma = 0$
(5)	$\xi = 0$	$\gamma$	$\xi = 0$
(6)	$\gamma = \xi = 0$	None	$\gamma = \xi = 0$
(7)	$\eta = 0$	$\xi$	$\gamma = \frac{(1+2\rho\sigma\xi) \pm \sqrt{(1+2\rho\sigma\xi)^2 - 4((\sigma\xi)^2 + 2\kappa^*\xi)}}{2}$
(10)	$\xi = 0$	$\gamma, \alpha$	$\xi = 0$

## E A Closed-Form Expression for the Joint Probability Distribution

Unfortunately it is not straightforward to compute the marginal pricing kernel for the case where the physical and risk-neutral parameters are independently estimated or when the affine price of risk specification is employed in the joint estimation. The pricing kernel can obviously not be computed directly for these two cases. Instead we calculate the daily log pricing kernel every day as follows:

$$\log M(t+\Delta) - \log M(t) = -r\Delta + \log Q(\log R(t+\Delta), v(t+\Delta)|v(t)) - \log P(\log R(t+\Delta), v(t+\Delta)|v(t)), \quad (\text{E.1})$$

where  $\Delta = 1/252$ . Note that  $Q(\log R(t+\Delta), v(t+\Delta)|v(t))$  and  $P(\log R(t+\Delta), v(t+\Delta)|v(t))$  are the risk-neutral and physical joint probability distributions of  $\log R(t+\Delta)$  and  $v(t+\Delta)$  conditional on  $v(t)$ . In this implementation, we have to use the joint probability rather than the marginal

probability of the log return described in equation (29), because we need to use the realization of both  $\log R(t+\Delta)$  and  $v(t+\Delta)$ .<sup>25</sup> We can also find a quasi closed form expression (up to numerically computing double integrals for the joint probability).

To find the expression for the joint probability distribution of the log stock return and variance, we first find the joint characteristic function and then apply the inverse Fourier transform. Let  $g_\tau^{CH}(\phi_x, \phi_v|x(t), v(t))$  denote the joint characteristic function of the log stock price (here denoted by  $x$ ) and the variance ( $v$ ).

When  $x$  and  $v$  evolve according to

$$\begin{aligned} dx(t) &= [r + uv(t)]dt + \sqrt{v(t)}dz_1(t), \\ dv(t) &= (a - bv)dt + \sigma\sqrt{v(t)}dz_2(t), \end{aligned}$$

with  $\text{corr}(z_1, z_2) = \rho$ , the characteristic function  $g_\tau^{CH}(\phi_x, \phi_v|x, v)$  must satisfy the following partial differential equation:

$$\frac{1}{2}v\frac{\partial^2 g}{\partial g^2} + \rho\sigma v\frac{\partial^2 g}{\partial x\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 g}{\partial v^2} + (r + uv)\frac{\partial g}{\partial x} + (a - bv)\frac{\partial g}{\partial v} + \frac{\partial g}{\partial t} = 0. \quad (\text{E.2})$$

See Heston (1993) for more details. Since  $g$  is the joint characteristic function, the terminal condition of the PDE is

$$g_0^{CH}(\phi_x, \phi_v|x, v) = e^{i\phi_x x + i\phi_v v}. \quad (\text{E.3})$$

Suppose  $g$  has the following functional form:

$$g_\tau^{CH}(\phi_x, \phi_v|x, v) = e^{G(\tau) + H(\tau)v + i\phi_x x}. \quad (\text{E.4})$$

By substituting equation (E.4) into equation (E.2), we get the following ordinary differential equa-

---

<sup>25</sup>The marginal probability of the log return is the expectation of the joint probability with respect to the variance  $v(t+\Delta)$ .

tions (ODEs) for  $G(\tau)$  and  $H(\tau)$ :

$$\begin{aligned} G'(\tau) &= r\phi_x i + aH(\tau), \\ H'(\tau) &= -\frac{1}{2}\phi_x^2 + \rho\sigma\phi_x i H(\tau) + \frac{1}{2}\sigma^2 H(\tau)^2 + u\phi_x i - bH(\tau), \end{aligned} \quad (\text{E.5})$$

and the terminal conditions of the ODEs are  $G(0) = 0$  and  $H(0) = i\phi_v$  inferred from equation (E.3).

This system of ODEs expressed in equation (E.5) has the following closed-form solution:

$$\begin{aligned} H(\tau) &= \frac{A_2(D_m X - iA_2 Y)\phi_v - 2iA_1(-D_m - 2A_3\phi_v + (Y-1)(D_m - 2A_3\phi_v))}{iA_2 D_m X - A_2^2 Y + 4A_1 A_3 Y - 2A_3 D_m X \phi_v}, \\ G(\tau) &= r\phi_x i \tau \\ &+ \frac{1}{4A_3} \left[ -2aA_2\tau - \frac{2ia\tau(A_2^2 - 4A_1 A_2)}{D_m} + 2ia \arctan\left(\frac{A_2 A_3 X^2 \phi_v}{A_2^2(Y-1) - A_1 A_3 Y^2 + A_3^2 X^2 + \phi_v^2}\right) \right. \\ &+ \frac{4aD_p}{D_m} \arctan\left(\frac{iA_2^2 + 2A_2 A_3 X \phi_v - 2iA_3(A_1 Y - A_3 X \phi_v^2)}{D_p(A_2 + 2iA_3\phi_v)}\right) \\ &- \frac{4iaD_p}{D_m} \operatorname{arctanh}\left(\frac{D_p}{A_2 + 2iA_3\phi_v}\right) + a \log(D_p) \\ &\left. - a \log(A_2^4(Z-1) + A_1^2 A_3^2 Y^4 + A_2^2 A_3^2 X^2 Z \phi_v^2 + A_3^4 X^4 \phi_v^4 - 2A_1 A_3 Y^2 (A_2^2(Y-1) + A_3^2 X^2 \phi_v^2)) \right], \end{aligned}$$

where

$$\begin{aligned} A_1 &= -\frac{1}{2}\phi_x^2 + u\phi_x i, \quad A_2 = \rho\sigma\phi_x i - b, \quad A_3 = \frac{1}{2}\sigma^2 \\ D_m &= \sqrt{-A_2^2 + 4A_1 A_3}, \quad D_p = \sqrt{A_2^2 - 4A_1 A_3} \\ X &= -1 + e^{iD_m\tau}, \quad Y = 1 + e^{iD_m\tau}, \quad Z = 1 + e^{2iD_m\tau}. \end{aligned}$$

To find the joint probability distribution function of  $x(t+\tau)$  and  $v(t+\tau)$ , we apply the inverse Fourier transform to the characteristic function. That is,

$$Pr(x, v|x(t), v(t)) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\phi_x x - i\phi_v v} g_{\tau}^{CH}(\phi_x, \phi_v|x(t), v(t)) d\phi_x d\phi_v \quad (\text{E.6})$$

To simplify the notation, we normalize the stock price. By letting the log stock price at time  $t$  be

$x(t) = 0$ ,  $x(t + \tau)$  represents the  $\tau$ -horizon log return. Then, equation (E.6) can be written as

$$Pr(x, v|v(t)) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\phi_x x - i\phi_v v + G(\tau|\phi_x, \phi_v) + H(\tau|\phi_x, \phi_v)v(t)} d\phi_x d\phi_v. \quad (\text{E.7})$$

Numerical integration of equation (E.7) is not tractable because the exponential function to be integrated decays very slowly, especially over the  $\phi_v$  dimension. To expedite its calculation, we use the following weight function for the  $j$ -th grid point of  $\phi_v$ ,  $w_j$ :

$$w_j = \begin{cases} \frac{1}{2} \operatorname{erfc} \left( -\frac{j}{\sqrt{N_v/2}} - \sqrt{N_v/2} \right) & \text{if } j < 0 \\ \frac{1}{2} \operatorname{erfc} \left( \frac{j}{\sqrt{N_v/2}} - \sqrt{N_v/2} \right) & \text{if } j \geq 0, \end{cases}$$

where  $2N_v$  is the number of grid points of  $\phi_v$ . This weight function forces the exponential function to decay much faster (see, for example, Ooura, 2001).<sup>26</sup> Hence, we numerically compute equation (E.7) as

$$Pr(x, v|v(t)) = \frac{1}{4\pi^2} \sum_{j=-N_v}^{N_v-1} \sum_{k=-N_x}^{N_x-1} e^{-ix\phi_{x,k} - iv\phi_{v,j} + G(\tau|\phi_{x,k}, \phi_{v,j}) + H(\tau|\phi_{x,k}, \phi_{v,j})v(t)} w_j \zeta(\phi_{x,k}, \phi_{v,j}),$$

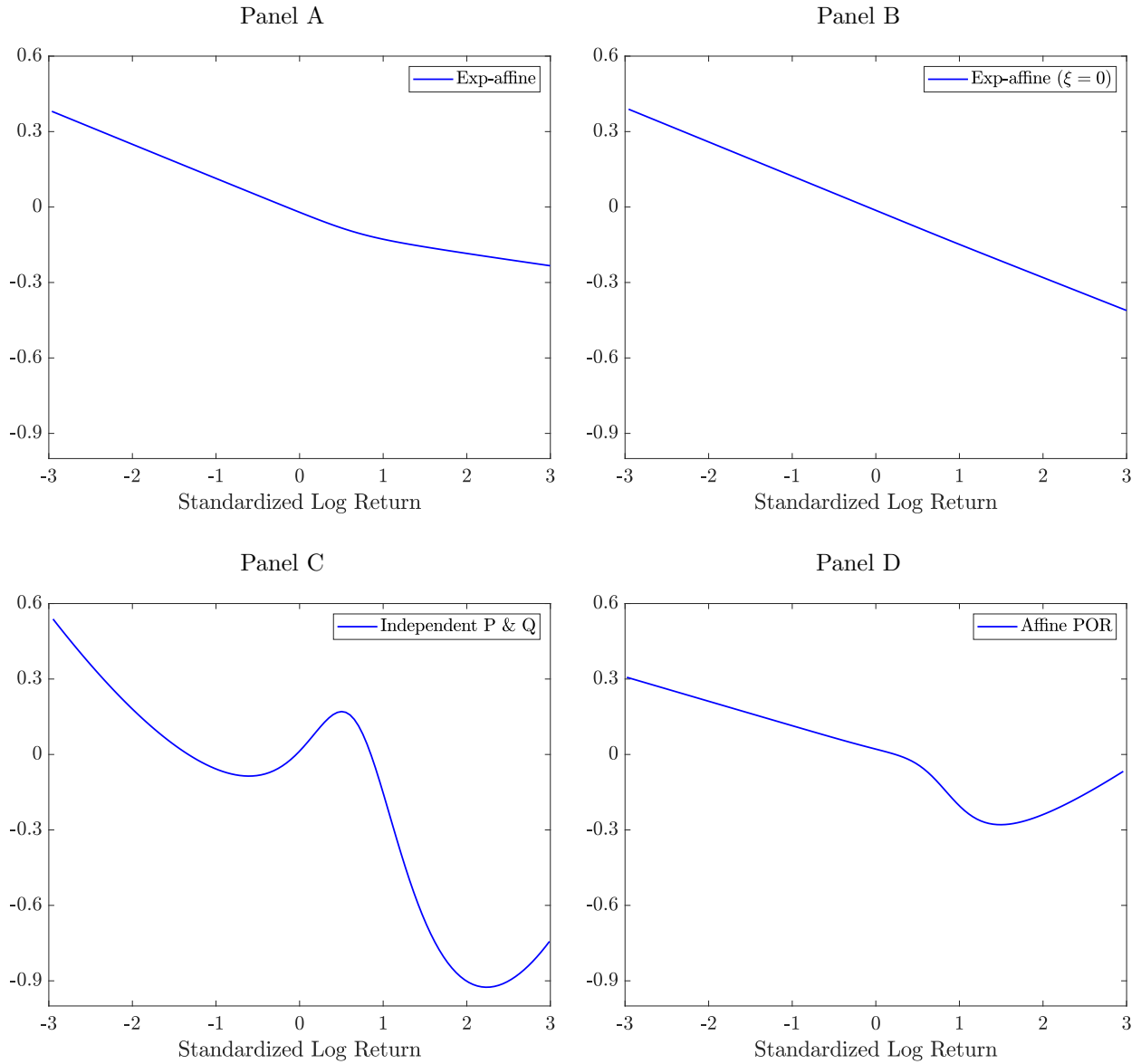
where  $\zeta(\phi_x, \phi_v)$  is an integration rule such as the trapezoidal or Gaussian quadrature method.

---

<sup>26</sup>Note that this approximation may cause an error if the true probability is near zero. This can happen when the size of  $x(t + \tau)$  or  $v(t + \tau) - v(t)$  is very large. We find that the approximation works well for our application.



Figure A.1: Implied One-Month Log Pricing Kernels



Notes: We plot the implied one-month log pricing kernels for the following four specifications: The unrestricted exponential-affine (Panel A), the exponential-affine specification with the restriction  $\xi = 0$  (Panel B), the kernel implied by independent estimation of the P- and Q-parameters (Panel C), and the affine price-of-risk specification (Panel D). Parameter values for Panels A and B are from Table 3, parameter values for Panel C are from Table 2, and parameter values for Panel D are from Table 4. The  $x$ -axis represents the log return in deviation from its mean and divided by its standard deviation.

Table A.1: Joint MLE Estimation 1990-2019. Feller Conditions Imposed

Panel A. Parameter Estimates		
	Feller Condition Imposed	
	Affine POR (1)	PK in Eq. (16) (2)
Risk Preference Parameters		
$\gamma$		-1.1661 (0.1140)
$\xi$		8.4911 (0.2829)
$\alpha$		0.0000 (0.0000)
Implied Risk Premia Parameters		
$\mu_0$	0.0294 (0.0265)	0.0000
$\mu_1$	1.6372 (1.4750)	2.4652
$\lambda_0$	0.0000 (0.0000)	0.0000
$\lambda_1$	-1.6270 (0.0406)	-1.6555
P-Parameters		
$\kappa$	3.9022 (0.0294)	3.9316 (0.0303)
$\theta$	0.0325 (0.0003)	0.0323 (0.0003)
$\sigma$	0.5036 (0.0073)	0.5037 (0.0074)
$\rho$	-0.8491 (0.0050)	-0.8491 (0.0053)
$\eta_0$	-0.0064 (0.0002)	-0.0064 (0.0002)
$\eta_1$	0.9060 (0.0059)	0.9060 (0.0060)
Q-Parameters		
$\kappa^*$	2.2752	2.2761
$\theta^*$	0.0557	0.0557
Model Fit		
$\log L^R$	53,970.97	53,970.58
$\log L^O$	149,289.28	149,287.78
$\log L^R + \log L^O$	203,260.25	203,258.36
Vega-Weighted Option Price RMSE	0.0222	0.0222
Panel B. Economic Implications		
Instantaneous Equity Premium	0.0828	0.0803
Instantaneous Variance Premium	-0.0530	-0.0539
Annualized Sharpe Ratio	0.4897	0.3976

Notes: This table reports the results of joint maximum likelihood estimation. The risk preference parameters  $\gamma$ ,  $\xi$ , and  $\alpha$  as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions. Robust standard errors are in parentheses. Column (1) reports on the affine price of risk (POR) specification and column (2) reports on the general pricing kernel (PK) from equation (16). The Feller condition is imposed under both the P- and Q-measures.