Exploring Risk Premia, Pricing Kernels, and No-Arbitrage Restrictions in Option Pricing Models*

Steven Heston
University of Maryland

Kris Jacobs
University of Houston

Hyung Joo Kim
University of Houston

July 15, 2022

Abstract

The empirical literature on dynamic option valuation typically does not specify a pricing kernel. It specifies the kernel indirectly through the price of risk, or defines it implicitly as the ratio of the risk-neutral and physical probabilities. We explicitly characterize pricing kernels that satisfy absence of arbitrage. These kernels are volatility-dependent by construction and provide unique insights on the impact of stock market volatility risk on state prices. We study the implications of these kernels for pricing kernel anomalies and existing specifications of the price of risk. Different affine price-of-risk specifications correspond to pricing kernels with radically different, and sometimes implausible, economic implications. We find that it is difficult to statistically distinguish between pricing kernels with widely different economic implications and risk premia. We attribute this to the inherent statistical problem with the estimation of equity and variance risk premia. This finding extends Merton's (1980) observations 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.

*Heston: sheston@umd.edu; Jacobs: kjacobs@bauer.uh.edu; Kim: hjkim@bauer.uh.edu. We would like to thank Gurdip Bakshi, Mikhail Chernov, Hitesh Doshi, Stefano Giglio, Paola Pederzoli, Jean-Paul Renne and seminar participants at the 2022 Finance Down Under Conference, the 2022 SoFiE Conference, Syracuse University, and the University of Houston for helpful comments.

1 Introduction

The pricing kernel is the most critical concept in asset pricing. It ensures absence of arbitrage and governs the relationship between physical and risk-neutral probabilities at all times and for all return horizons. One long-standing approach to learn about 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. This literature has made enormous progress in matching the moments of asset returns, but there is as yet no consensus on the preferred model and the resulting properties of the pricing kernel.[1]

An alternative approach to identifying the pricing kernel avoids the use of consumption data. Building on the insights of Breeden and Litzenberger (1978), an extensive literature starting with Aït-Sahalia and Lo (1998), Aït-Sahalia and Lo (2000), Jackwerth and Rubinstein (1996), Jackwerth (2000) and Rosenberg and Engle (2002) estimates the pricing kernel using index returns and index option prices. Whereas the pricing kernel prices all assets, and the insights of Breeden and Litzenberger (1978) can be used to estimate the state price density for any asset, 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.[2] However, this literature has given rise to puzzles of its own, most importantly the finding that the pricing kernel is not monotonically decreasing as a function of aggregate wealth, but rather U-shaped.[3]

In light of the importance of the pricing kernel and these outstanding research questions, it is surprising that the literature seems to have overlooked an alternative approach to learn about pricing kernels. The existing empirical literature on parametric dynamic index option pricing models typically does not explicitly specify the pricing kernel. It characterizes the kernel indirectly by

[1] See Mehra and Prescott (1985) and Hansen and Jagannathan (1991) for problems with early consumption-based models. See for instance Campbell and Cochrane (1999), Bansal and Yaron (2004), Gabaix (2012) and Wachter (2013) for examples of consumption-based models that have been more successful in matching the data.

[2] See Chernov (2003) for an empirical study that estimates the pricing kernel using a cross-section of securities.

[3] 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.

specifying prices of risk or defines it implicitly as the ratio of risk-neutral and physical probabilities.[4] The motivation for this approach is presumably that the explicit mapping from the prices of risk to the pricing kernel is mechanical, and does not provide additional insights.[5]

This paper argues that the specification of the pricing kernel should be an integral part of the specification of these models and shows that some existing approaches are equivalent to implausible economic assumptions. We propose a class of pricing kernels that 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).[6] These kernels are volatility-dependent by construction and are therefore especially useful to analyze the impact of stock market volatility risk on state prices and market risk. They can be path-dependent and nest path-independent kernels (Ross, 2015) as special cases.

The main advantage of explicitly specifying the pricing kernel is that the resulting parameterization provides economic content and can be restricted to ensure that the resulting underlying returns and option prices are arbitrage-free and consistent with intuitively plausible loadings on the model's state variables. We show that small modifications to the affine price of risk specifications used in the literature correspond to different pricing kernels with radically different economic implications. Some of these kernels, and the corresponding risk specifications, give rise to economically implausible results, such as state prices that are S-shaped as a function of market returns. A kernel consistent with the completely affine price of risk specification produces very plausible results.

We find that 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. We argue that existing tests have low power to statistically distinguish different pricing kernels because the identification of the pricing kernel is equivalent to the estima-

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

[5]This is all the more noteworthy because the term structure (Ang and Piazzesi, 2003; Ang, Bekaert, and Wei, 2008; Adrian, Crump, and Moench, 2013) and currency option literatures (Chernov, Graveline, and Zviadadze, 2018) explicitly specify pricing kernels.

[6]We use the simplest possible option pricing model with a stochastic volatility factor to highlight the generality of our theoretical and empirical observations.

tion of conditional risk premia. We need a lot of data to estimate unconditional average returns on stocks or options, and even more data to estimate conditional returns. Moreover, plain vanilla option prices are sensitive to the probabilities at expiration, but are 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.

These findings provide important insights into the estimation of risk premia, which constitutes another important challenge in asset pricing. It is well known that estimating risk premia is difficult. For instance, Merton (1980) convincingly argues that very long time series of returns are required to obtain reliable estimates of the equity premium. Our findings extend the observation in Merton (1980) to joint estimation of equity and variance risk premia using the cross-section of options and the underlying returns. Our results are also consistent with Merton's in the sense that in the presence of these identification problems, we find that economic intuition and theory may be the most useful tool to identify market risk premia. Specifically, Merton (1980) advocates imposing a positivity restriction on the path of the conditional equity risk premium. Similarly, we find that imposing a negativity restriction on the market variance risk premium leads to more plausible and reliable estimates.

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.[7] We explore the most popular specifications of the price of risk in the option literature, where the equity (market) risk premium $\mu_0 + \mu_1 v(t)$ and the variance risk premium $\lambda_0 + \lambda_1 v(t)$ are modeled as affine functions of the variance $v(t)$. We refer to this specification as "affine". We also explore the important nested case with zero intercepts ($\mu_0 = \lambda_0 = 0$), which we refer to as "completely affine", adopting the terminology in Singleton (2006, p. 392) and the term structure literature. We specify pricing kernels that are consistent with these risk premium specifications and derive the parameter restrictions consistent with the martingale condition and no-arbitrage.

[7] 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), Aït-Sahalia and Kimmel (2007), Hurn, Lindsay, and McClelland (2015), and Andersen, Fusari, and Todorov (2017).

We find that the seemingly innocuous addition of an intercept in the affine specification of the risk premium has important consequences. Including an intercept requires a different specification of the pricing kernel and a different and more complex set of restrictions on the economic parameters that characterize the pricing kernel. However, while part of the literature seems to imply that $\mu_0 = \lambda_0 = 0$ is a necessary condition for no-arbitrage, we show that it is merely a sufficient condition and we characterize the no-arbitrage restrictions in the more general case.

We then study whether the restrictions on the intercept in the completely affine specification are supported by the data. While the inclusion of an intercept improves the statistical fit, we find that only the completely affine specification is strongly supported based on its economic implications. The marginal pricing kernel that defines state prices as a function of market returns (aggregate wealth) is well behaved for the completely affine specification, but the pricing kernel consistent with the affine specification leads to adjacent state prices that are very different and state prices as a function of wealth that defy economic intuition. The intuition for this finding is that the intercepts in the affine price of risk specification lead to counter-intuitive Sharpe ratios and/or implausible signs of the risk premia. We, therefore, recommend the use of the completely affine price of risk specification and argue against the more general specification, not because it allows for arbitrage but because of its unrealistic economic implications. We emphasize that this finding cannot be uncovered without explicitly considering the pricing kernel implicit in the price of risk specification.

Next we estimate the option pricing model subject to various restrictions on the parameters of the pricing kernel that underlies the completely affine specification of the price of risk. We test for the significance of the parameters that capture the pricing kernel's dependence on variance and market returns. 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. We find that imposing these strong economic restrictions lead to radically different estimates of the equity and variance risk premia. This is also reflected in large differences in the state prices embodied in the marginal pricing kernel as a function of wealth. We also use

our estimates to emphasize the critical role of the volatility state variable by inspecting the realized time series path of the pricing kernel, i.e., the path obtained by inserting the realizations of the state variables. By comparing the path of the unrestricted pricing kernel with the paths of kernels that restrict some of the state variables, we can gauge the importance of these state variables and restrictions. We find that the kernel without volatility risk is substantially less variable compared to the unrestricted kernel, especially in crisis periods. The path-independent kernel and the affine price of risk specification result in kernels that are very implausible.

Surprisingly, however, when inspecting the likelihoods based on option prices and the underlying returns, the strong economic implications of these parameter restrictions do not translate into large decreases in the likelihood. 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. This finding illustrates fundamental limitations faced by the option pricing literature. 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. Our findings suggest that this approach yields risk premium estimates that are not very precise, and that theoretical restrictions may enhance power. Our explicit characterization of the pricing kernel is critical to arrive at these conclusions, because it shows that the option data alone do not provide sufficient statistical power to distinguish economically different risk premia.

Finally, our results suggest that U-shaped pricing kernels may not be anomalous nor constitute an asset pricing puzzle. These shapes are implied by physical and risk-neutral stochastic volatility dynamics in conjunction with pricing kernels that are entirely consistent with rational economic behavior, and are due to the pricing of volatility risk in these models. Provided that the option valuation model contains volatility as a state variable, a monotonic pricing kernel will only obtain if volatility risk is close to zero or its price is very small, which is inconsistent with the existing evidence. We conclude that the study of pricing kernels as a function of index returns is overly restrictive when the variance is a relevant state variable.

We emphasize that we are not the first to study the impact of the pricing kernel on option prices. Several studies use consumption-based models to analyze how preferences and pricing kernels impact index option prices.[8] 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). The disadvantage of our approach is that it does not specify the theoretical relation with consumption as a state variable, but the advantage is that it provides a direct relation with existing empirical implementations of (reduced-form) parametric dynamic option pricing models. It is therefore relatively straightforward to implement and estimate using option data, which allows us to fully characterize and explore the impact of stock index volatility on the pricing kernel.

From an empirical perspective, a closely 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. Our contribution is distinct because our empirical approach is guided by the specification of a class of parametric pricing kernels that are consistent with a specific option pricing model. 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. Beason and Schreindorfer (2020) analyze the implications of option data for macro-finance models.

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.

[8] 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.

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 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.

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, in order to be able to identify key aspects of the model including volatility persistence.[9] 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 period is longer than 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 rates obtained from CRSP. Following existing work, options are

[9] See, for instance, Broadie, Chernov, and Johannes (2007) for a discussion.

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.

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 for 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. Our prior is that most of the issues we highlight here using a simple model are even more relevant in more complex models, but we keep 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:

$$\begin{aligned} d\log S(t) &= \left[r - \frac{1}{2}v(t)\right]dt + \sqrt{v(t)}dz_1^*(t), \\ dv(t) &= \kappa^*(\theta^* - v(t))dt + \sigma\sqrt{v(t)}dz_2^*(t), \end{aligned} \tag{1}$$

where dz_1^* and dz_2^* are Wiener processes with correlation coefficient ρ. The risk-free rate r can be either constant or time-varying; this has negligible implications for our results. It is also straight-

forward 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:

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

where μ is the parameter that identifies the equity premium, and dz_1 and dz_2 are Wiener processes under the physical measure. Note that σ, the variance of variance parameter, and ρ, 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 θ and mean reversion κ differ from the long-run risk-neutral variance θ^* and mean reversion κ^*. 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.[10] We use the fact that

[10] See Bates (2000) and Andersen, Fusari, and Todorov (2015) for alternative approaches.

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:

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

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

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

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:

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

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 based on returns and the estimation based on 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 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 the Euler discretization to equation (2), which results in:

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

where $R(t+\Delta) = S(t+\Delta)/S(t)$ represents the gross return and $\Delta = 1/252$.[11] 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 respectively 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:

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

[11] 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.

where $f(\log R(t+\Delta), v(t+\Delta)|v(t))$ is the conditional density of 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 η_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 Θ, we solve the following optimization problem:

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

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 τ is expressed by a quasi closed form up to a numerical integration, and it is given by

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

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 $f_\tau^{CH}(\phi|S(t), v(t)) = E_t^*\left[e^{i\phi \log S(t+\tau)}\right]$ 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).[12] The price of a put option with the same strike price and maturity can be obtained through put-call parity.

[12] When $\log S(t)$ and $v(t)$ are characterized by

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

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.

We use vega-weighted option pricing errors. Let O_i^{Mkt} and O_i^{Mod} denote the market and the model prices of the ith 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 ν_i^{Mkt} is the Black-Scholes vega of option i.[13]

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. Θ^* is then estimated by solving the following optimization problem:

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

the characteristic function solution is given by

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

where

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

[13]This approach is often used when the levels of asset prices are different but implied volatilities are comparable. See for example Carr and Wu (2007), Trolle and Schwartz (2009), and Christoffersen, Heston, and Jacobs (2013).

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, κ is much larger than κ^*, while θ is much smaller than θ^*. The risk-neutral long-run variance exceeds the physical long-run variance, and risk-neutral persistence exceeds physical persistence.

The option-based kurtosis parameter σ is larger than the return-based estimate of σ, and the option-based skewness parameter ρ is more negative than the return-based ρ. The distribution implied by the option data is thus more fat-tailed and skewed than the physical distribution. The finding on σ 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 σ when estimating on options. Existing findings on ρ are mixed. Typically the estimates from returns are not very different from the option-based estimates.

The estimate of μ 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 σ, ρ, η_0, and η_1 to be the same. However, we can use the parameters θ and θ^* 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 Pricing Kernels

In this section, we characterize the class of pricing kernels implied by the physical and risk-neutral dynamics in equations (1) and (2). We first characterize a class of exponential-affine pricing kernels. This exponential-affine form is a function of $S(t)$ and the historical path of $v(t)$. In special cases, the exponential-affine kernel is path-independent (Ross, 2015), as it depends only on the current value of $v(t)$. We subsequently study a more general class of pricing kernels that is helpful to understand commonly used specifications of the price of risk.

4.1 A Class of Exponential-Affine Pricing Kernels

Most of the existing literature is not explicit about the pricing kernel that links the risk-neutral and physical dynamics in equations (1) and (2). Appendix A shows that the following class of pricing kernels is consistent with these dynamics:

$$M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^{\gamma} \exp\left(\beta t + \eta \int_0^t v(s)ds + \xi(v(t) - v(0))\right), \tag{11}$$

where γ and ξ are the index level and variance preference parameters. We refer to β as the time-preference parameter and to η as the path-dependence parameter, respectively. Note that the dynamics (1) and (2) imply that return means and variances are linear in $v(t)$. The logarithm of the pricing kernel (11) is also linear in $log(S(t))$ and (the path of) $v(t)$. Henceforth, we therefore refer to this as the exponential-affine pricing kernel.

This pricing kernel nests several important special cases. First, if we set η equal to zero, then the pricing kernel is a function of time t and the state variables $S(t)$ and $v(t)$, but does not depend on the history of the state variables. Ross (2015) refers to this property as "transition-independence" or "path-independence".

Another economically important special case occurs when setting ξ equal to zero. This gives:

$$M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^{\gamma} \exp\left(\beta t + \int_0^t \eta v(s) ds\right). \tag{12}$$

In this case, the local fluctuations in the pricing kernel (11) come exclusively from changes in the spot price $S(t)$, and there is no additional variance premium. This case is a dynamic version of Rubinstein's (1976) power utility pricing kernel.[14] The martingale condition implies that the expected growth of the pricing kernel must equal the opposite of the interest rate. This restricts β and η, effectively leaving only a single index level preference parameter γ to price risk.

The existing literature typically specifies that the risk-neutral and physical parameters in equations (1) and (2) are related as follows:

$$\kappa = \kappa^* - \lambda, \tag{13}$$

$$\theta = \frac{\kappa^* \theta^*}{\kappa}, \tag{14}$$

where λ is the variance risk premium parameter. Assuming the pricing kernel in equation (11), the structure of the variance risk premium parameter λ and the equity risk premium parameter μ can be inferred by computing the instantaneous equity and variance risk premia. The risk premia can be computed using $-E\left[(dS/S)(dM/M)\right]$ and $-E\left[dv(dM/M)\right]$, which results in:

$$\text{Equity risk premium:} \qquad (-\gamma - \xi\sigma\rho)v(t), \tag{15}$$

$$\text{Variance risk premium:} \qquad (-\rho\sigma\gamma - \sigma^2\xi)v(t), \tag{16}$$

[14] In equation (12), γ is usually interpreted as a risk aversion parameter. We use the terminology "index level preference parameter" because in the presence of additional state variables in equation (11), the risk aversion terminology may cause confusion.

or equivalently the following expressions for μ and λ:

$$\mu = -\gamma - \xi\sigma\rho, \tag{17}$$

$$\lambda = -\rho\sigma\gamma - \sigma^2\xi. \tag{18}$$

Alternatively, rearranging equations (17) and (18) provides the expressions for the risk preference parameters γ and ξ:

$$\gamma = -\mu - \rho\sigma\xi, \tag{19}$$

$$\xi = \frac{\mu\sigma\rho - \lambda}{\sigma^2(1-\rho^2)}. \tag{20}$$

This leaves two more parameters, β and η, to be determined in equation (11). Using the fact that $e^{rt}M(t)$ is a martingale, and that the drift term of a martingale process has to be zero, Appendix A shows that β and η are given by:

$$\beta = -(1+\gamma)r - \xi\kappa\theta, \tag{21}$$

$$\eta = -\left(\mu - \frac{1}{2}\right)\gamma + \xi\kappa - \frac{1}{2}\left(\gamma^2 + 2\gamma\xi\sigma\rho + \xi^2\sigma^2\right). \tag{22}$$

In summary, the relation between the risk-neutral and physical parameters and the pricing kernel is as follows. Recall that σ and ρ are restricted to be the same under both measures. Given the remaining risk-neutral parameters and the risk preference parameters γ and ξ, we can obtain the physical parameters κ and θ in conjunction with the risk premium parameters μ and λ by using equations (13), (14), (17), and (18). The remaining pricing kernel parameters β and η can then be obtained from equations (21) and (22).

What are our priors on the signs of these parameters? First consider the four pricing kernel parameters, γ, ξ, β, and η. We have very clear priors regarding γ and ξ. Economies with higher index returns and lower variance indicate good times, and marginal utility decreases in good times. We therefore expect $\gamma < 0$ and $\xi > 0$. Economic intuition is not very informative on the sign (and

by extension the magnitude) of η.

We also have priors and/or empirical evidence on the risk premium parameters μ and λ. It is well known that it is difficult to measure the index equity premium μ precisely, especially over a short horizon, but the evidence is overwhelming that the point estimates are positive over sufficiently long horizons, and this is consistent with economic intuition. Estimates of the index variance risk premium show that it is also highly time-varying, but existing measurements are almost exclusively negative over long horizons.

Given this intuition on the risk premium parameters (μ and λ) and pricing kernel parameters (γ and ξ), what do equations (17) and (18) imply? Recall that the ρ parameter captures the correlation between index return and variance innovations. Our estimates in Table 2 are negative, and this is consistent with extensive existing non-parametric and parametric evidence. The σ parameter captures variance of variance and is positive by construction. Given these signs, equation (17) implies that larger γ (more negative) and ξ (more positive) increase the index equity premium. This simply captures that investors require higher risk compensation when they are more risk-averse towards wealth and variance risk. Equation (18), on the other hand, shows that larger γ and ξ also lead to a more negative λ. The higher investors' risk aversion with respect to wealth and variance, the more they are willing to pay to insure against high volatility.

One of the nested pricing kernels we consider in detail below is the "power utility" pricing kernel in equation (11), which imposes $\xi = 0$. Note how this further restricts the risk premia in equations (17) and (18), given our finding in Table 2 that $\sigma > 0$ and $\rho < 0$. Most critically, while with nonzero ξ the equity and variance risk premia are positive and negative respectively if γ and ξ are consistent with economic intuition, it is in principle possible for their signs to be the same for some parameter combinations. However, when $\xi = 0$, the equity and variance risk premia are mechanically restricted to have opposite signs. There are, of course, additional more subtle restrictions incurred by restricting $\xi = 0$. We analyze those in detail below.

4.2 Completely-Affine Prices of Risk

In the literature on the estimation of stochastic volatility option pricing models, the mapping between the two measures is usually implemented by restrictions on risk premia, and most existing implementations assume that risk premia are linear (affine) in the variance. Singleton (2006, pp. 392-396) provides a detailed discussion of this issue. Consistent with the terminology in the bond pricing literature (Duffee, 2002), Singleton (2006) refers to $\pi\sqrt{v(t)}$ as the "completely affine" form of the price of risk. Note that in this model, this specification of the price of risk implies a risk premium of $\pi\sqrt{v(t)}\sqrt{v(t)} = \pi v(t)$.

For our purpose, note that when estimating using stock price dynamics and option prices, the completely-affine price of risk specification is isomorphic to the exponential-affine pricing kernel specification under certain implementation assumptions. To see this, note that given the parameters characterizing the physical dynamics, the risk premium parameters μ and λ can be thought of as being determined by the two parameters γ and ξ from equations (17) and (18). Therefore, the exponential-affine kernel and the completely affine price of risk specification are isomorphic to each other when these are the only restrictions imposed in estimation. However, this effectively ignores the information in equations (21) and (22). These equations embody the observable implications of the time-preference parameter β and the path-dependence parameter η, and are also obtained from a martingale restriction and thus directly implied by theory.

In implementation, equation (21) depends on the risk-free rate. However, when pricing options, typically a different risk-free rate is used for every option maturity. As for the risk-free rate used in the physical dynamic, it does not restrict the risk premium parameters μ and λ. Effectively therefore, the information in equation (21) is ignored when implementing a completely affine price of risk. Equation (22) is also ignored; effectively the γ and ξ parameters that are isomorphic to μ and λ imply an η parameter which is left free in the implementation.

Ideally the information in the risk-free rate and equation (21) can be exploited in estimation. However, since one of our main purposes is to compare different specifications of the pricing kernel and the price of risk, we follow existing implementations and do not incorporate this information.

In most of our estimations, we also do not use equation (22) and therefore effectively leave η as a free parameter. One exception is that we do investigate the case $\eta = 0$ case because of our interest in the path-independent kernel of Ross (2015). In this case, we take into account the impact of the $\eta = 0$ restriction on the γ and ξ parameters while also accounting for equations (17) and (18).

4.3 More General Pricing Kernels

We now consider pricing kernels outside the exponential-affine class in equation (11), while maintaining compatibility with the risk-neutral and physical dynamics (1) and (2). Specifically, we study the following class of pricing kernels:

$$M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^{\gamma}\left(\frac{v(t)}{v(0)}\right)^{\alpha}\exp\left(\beta t + \int_0^t \eta(v(s))ds + \xi(v(t) - v(0))\right). \tag{23}$$

To facilitate comparisons with the exponential-affine kernel in equation (11), there is some abuse of notation in equation (23). While η represents the coefficient of the path-dependent component in equation (11), $\eta(v(s))$ indicates a path-dependent function of $v(s)$ in equation (23).

We specify this kernel as a generalization of the exponential-affine kernel in equation (11). As a result, it is overparameterized because both the α and ξ parameters capture variance risk aversion. When the variance risk aversion parameter α is equal to zero, it corresponds to the exponential-affine kernel in equation (11).

The martingale property shows that under the kernel in equation (23), the model-implied equity and variance risk premia are still linear, but not necessarilly proportional to $v(t)$:

$$\text{Equity risk premium:} \qquad -\alpha\sigma\rho - (\gamma + \rho\sigma\xi)v(t), \tag{24}$$

$$\text{Variance risk premium:} \qquad -\alpha\sigma^2 - (\gamma\rho\sigma + \xi\sigma^2)v(t). \tag{25}$$

Note how the variance aversion parameter α is scaled by $\sigma\rho$ in equation (24) and by σ^2 in equation (25). Due to the parameterization of the pricing kernel, these are also the scaling factors on $\xi v(t)$.

$$
\begin{aligned}
d\log S(t) &= \left[r - \frac{1}{2}v_t + \mu_0 + \mu_1 v(t)\right] dt + \sqrt{v(t)}dz_1(t), \\
dv(t) &= \kappa(\theta - v(t))dt + \sigma\sqrt{v(t)}dz_2(t),
\end{aligned} \tag{26}
$$

where

$$
\kappa = \kappa^* - \lambda_1, \quad \theta = (\kappa^*\theta^* + \lambda_0)/\kappa, \tag{27}
$$

$$
\mu_0 = -\alpha\sigma\rho, \quad \mu_1 = -\gamma - \rho\sigma\xi, \quad \lambda_0 = -\alpha\sigma^2, \quad \lambda_1 = -\gamma\rho\sigma - \xi\sigma^2. \tag{28}
$$

The time-preference parameter and the path-dependence function $\eta(v(t))$ are given by:

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

Note that for the exponential-affine pricing kernels in Section 4.1, the instantaneous equity risk premium in equation (15) can be written as $\mu v(t)$ and the instantaneous variance risk premium in equation (16) as $\lambda v(t)$. In the case of the pricing kernel in equation (23), the equity and variance risk premia in equation (24) and (25) conform to $\mu_0 + \mu_1 v(t)$ and $\lambda_0 + \lambda_1 v(t)$, respectively. Also note that for the kernels in equation (23), the path-dependent component $\eta(v(s))$ is a function of $v(s)$ which is not necessarily linear.

Finally, similar to the terminology used in Section 4.2, we can also express these equity and variance risk premia per unit of risk, which we previously referred to as the prices of risk. We refer to $\frac{\pi_0}{\sqrt{v(t)}} + \pi_1\sqrt{v(t)}$ as the "affine" specification of the price of risk, corresponding to affine risk premia, namely $\mu_0 + \mu_1 v(t)$ for the equity risk premium and $\lambda_0 + \lambda_1 v(t)$ for the variance risk premium.

We conclude that just as the completely-affine specification of the risk premia is intimately related to the pricing kernel in equations (11), the affine specification of the risk premia is related to the kernel in equation (23). These kernels can therefore be used to provide economic content for these assumptions on the price of risk used in risk neutralization. In Section 5.4 below, we compare the results of an implementation with affine risk premia to implementations that explicitly use the parameters characterizing the pricing kernel.

4.4 An Alternative Representation of the Pricing Kernel

We now discuss a related setup used by the existing literature. For expositional clarity we rewrite the physical dynamic in equation (2) as follows:

$$\begin{aligned} d\log S(t) &= \left[r + \left(\mu - \frac{1}{2}\right) v(t)\right] dt + \sqrt{v(t)} dz_s(t), \\ dv(t) &= \kappa(\theta - v(t))dt + \sigma\sqrt{v(t)}(\rho dz_s(t) + \sqrt{1-\rho^2} dz_v(t)), \end{aligned} \tag{29}$$

where dz_s and dz_v are independent Wiener processes. The risk-neutral dynamic has the same functional form:

$$\begin{aligned} d\log S(t) &= \left[r - \frac{1}{2} v(t)\right] dt + \sqrt{v(t)} dz_s^*(t), \\ dv(t) &= \kappa^*(\theta^* - v(t))dt + \sigma\sqrt{v(t)}(\rho dz_s^*(t) + \sqrt{1-\rho^2} dz_v^*(t)), \end{aligned} \tag{30}$$

where dz_s^* and dz_v^* are also independent Wiener processes. Now consider the following pricing kernel:

$$\frac{dM}{M} = -r dt - \Pi_{s,t} dz_{s,t} - \Pi_{v,t} dz_{v,t}, \tag{31}$$

where $\Pi_t = (\pi_{s,t}, \pi_{v,t})'$ are usually referred to as the market prices of risk. The risk premiums then have the form $\Sigma\sqrt{\Omega_t}\Pi_t$ (Singleton, 2006), where

$$\Sigma = \begin{pmatrix} 1 & 0 \\ \rho\sigma & \sigma\sqrt{1-\rho^2} \end{pmatrix}, \quad \Omega_t = \begin{pmatrix} v_t & 0 \\ 0 & v_t \end{pmatrix}, \quad \Pi_t = \begin{pmatrix} \pi_s\sqrt{v_t} \\ \pi_v\sqrt{v_t} \end{pmatrix}, \tag{32}$$

Note that this follows directly from the Euler equation and that the structure of Σ follows from the dynamic in (2). However, the structure of Π_t in (32) embodies an additional assumption, and is equivalent to the completely affine price of risk assumption in Section 4.2. Indeed, from (32) the equity and variance risk premia are given by:

$$\text{ERP} = \pi_s v_t, \quad \text{VRP} = (\rho\sigma\pi_s + \sigma\sqrt{1-\rho^2}\pi_v)v_t, \tag{33}$$

equivalent to the completely affine prices $\text{ERP} = \mu v_t$ and $\text{VRP} = \lambda v_t$. If instead we assume $\Pi_t = (\frac{\pi_{s,0}}{\sqrt{v_t}} + \pi_{s,1}\sqrt{v_t}, \frac{\pi_{v,0}}{\sqrt{v_t}} + \pi_{v,1}\sqrt{v_t})'$, we obtain affine equity and variance risk premia $\text{ERP} = \mu_0 + \mu_1 v_t$ and $\text{VRP} = \lambda_0 + \lambda_1 v_t$, where $\mu_0 = \pi_{s,0}$, $\mu_1 = \pi_{s,1}$, $\lambda_0 = \rho\sigma\pi_{s,0} + \sigma\sqrt{1-\rho^2}\pi_{v,0}$, and $\lambda_1 = \rho\sigma\pi_{s,1} + \sigma\sqrt{1-\rho^2}\pi_{v,1}$.

We conclude that while the setup in this section specifies a pricing kernel, it does not capture any additional economic intuition above and beyond affine restrictions on the price of risk. Our approach differs because it specifies the pricing kernel as a function of the state variables and thereby clarifies the interactions between the economic magnitudes of index level and index variance aversion and the parameters of the stochastic volatility dynamic in determining the risk premia. It also highlights the role of path dependence in an economy with stochastic volatility. Moreover, the approach outlined in this section also ignores the relation between the magnitudes of the equity and variance risk premia and the riskless rate embodied in equation (21), as well as the information in equation (22).[15]

[15] This information in the riskfree rate can of course be combined with the specification in equation (32), see Chernov (2003) for an example.

5 Pricing Kernels: Parameter Estimates

We estimate the stochastic volatility model subject to the restrictions imposed by the pricing kernels in Sections 4.1 and 4.3. 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 (11) for example, or we directly impose a structure on the parameters μ and λ that characterize the risk premia.

For kernels nested in the exponential-affine specification in equation (11), given a set of parameters $\Theta^{PK1} = \{\gamma,\ \xi,\ \kappa,\ \theta,\ \sigma,\ \rho,\ \eta_0,\ \eta_1\}$, we can obtain the risk premium parameters μ and λ via equations (17) and (18). Likewise, for the more general kernel in equation (23), given $\Theta^{PK2} = \{\gamma,\ \xi,\ \alpha,\ \kappa,\ \theta,\ \sigma,\ \rho,\ \eta_0,\ \eta_1\}$, the risk premium parameters μ_0, μ_1, λ_0, and λ_1 can be obtained from equation (28). The risk-neutral parameters κ^* and θ^* are then implied by equations (13) and (14) for the exponential-affine case and by equation (27) for the more general specification in equation (23).

We also investigate specifications that cannot be implemented by explicitly considering the pricing kernel parameters. In these cases, we restrict the risk premium parameters instead. Specifically, for the affine price of risk specification, given the parameters $\Theta^{PR} = \{\mu_0,\ \mu_1,\ \lambda_0,\ \lambda_1,\ \kappa,\ \theta,\ \sigma,\ \rho,\ \eta_0,\ \eta_1\}$, equations (27) and (28) provide the risk-neutral parameters.

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 Θ can refer to Θ^{PK1}, Θ^{PK2}, or Θ^{PR}.

Table 3 presents estimation results for eight different cases, which we now discuss. Column (1) reports on the unrestricted exponential-affine pricing kernel in equation (11). In columns (2) and (3) we have two restricted estimation exercises based on risk premium restrictions: $\mu = 0$ and $\lambda = 0$. Next, we have several interesting cases based on preference parameter restrictions. In column (4) we have $\gamma = 0$ and in column (5) we have $\xi = 0$. The restriction of $\xi = 0$ corresponds to the power utility pricing kernel in equation (12). We also test the joint restrictions $\gamma = \xi = 0$ in column (6).[16] Finally, we also test $\eta = 0$ in column (7), as well as $\eta = \xi = 0$ in column (8). These latter two cases provide insight into the literature on recovery theory (see, e.g., Ross, 2015; Borovička, Hansen, and Scheinkman, 2016; Qin and Linetsky, 2016). It is worthwhile to note that we do not impose any restrictions on the time-preference component β, which is required to describe the discounting factor.

Column (9) in Table 4 reports on the more general kernel in equation (23) and column (10) on the special case of this kernel with $\xi = 0$. In the latter case, the free parameters are γ and α. Finally, column (11) in Table 4 reports on the implementation with the affine price of risk.

The following table summarizes the restrictions for columns (2)-(10) in Tables 3 and 4. As mentioned above, for the case in column (11) there are no restrictions. We implement these cases by letting either γ or ξ be a free parameter, while the other one is implied by the restriction(s). The resulting mapping between γ and ξ is reported in the last column.

[16] Note that from equations (17) and (18), the joint restriction $\mu = \lambda = 0$ is equivalent to the joint restriction $\gamma = \xi = 0$.

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

5.2 The Exponential-Affine Pricing Kernel

Panel A of Table 3 reports the results of the joint MLE estimation. We report robust standard errors for the parameters in Θ^{PK1}, Θ^{PK2}, or Θ^{PR} except for the parameter that is fixed according to the last column in the table above. Column (1) presents the estimates for the MLE estimation of the exponential-affine pricing kernel in equation (11). We can compare the risk-neutral estimates in column (1) with the option-based estimates in Table 2. The risk-neutral mean-reversion κ^* from joint estimation in Table 3 is 1.114, somewhat higher than the 0.986 estimate in Table 2. The risk-neutral long-run variance θ^* in Table 3 is 0.0877, somewhat lower than the 0.0986 estimate in Table 2. These findings are not surprising, because the option-implied variance typically exceeds the average variance implied by returns. Recall that the ρ and σ parameters are the same under both measures. The estimate of ρ from joint estimation is a bit larger in absolute value compared to the one in Panel B of Table 2, and the estimate of σ is a bit smaller, but the differences are relatively minor.

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 κ in Table 3 is 2.926, substantially smaller than the 4.146 estimate in Table 2. The physical long-run variance θ in Table 3 is 0.033, similar to 0.031 in Table 2. The most important difference is that the return-based estimates of ρ and σ in Table 2

are substantially smaller (in absolute value) compared to Table 3. Recall that they are also smaller than the option-based estimates in Table 2.

The estimate of the index level preference parameter ("risk aversion") γ in column (1) is -1.393 and the estimate of the variance preference parameter ξ is 1.947. Both signs are consistent with economic intuition. The point estimate of μ in column (1) of Table 3 is similar but slightly lower than the estimate in Table 2. The variance risk premium parameter λ is estimated to be negative in column (1) of Table 3, which is also intuitively plausible. The negative λ implies $\kappa > \kappa^*$ and $\theta < \theta^*$. These ordinal relations between the physical and risk-neutral parameters are the same as in Table 2, although the values of the physical parameters differ somewhat.

5.3 Restrictions on the Exponential-Affine Pricing Kernel

Columns (2)-(8) 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 μ and λ in columns (2) and (3), the resulting estimates of the risk preference parameters are inconsistent with economic intuition. We observe either a positive γ or a negative ξ. The implied risk premium parameters μ and λ and the physical parameters also strongly deviate from the estimates in column (1). Columns (4) and (5) present the results when risk preferences are restricted. In these cases, the remaining unrestricted risk preference estimates have the same sign as in column (1), but they are larger in absolute value. The other parameters are very similar to the parameters in 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 κ and θ, but not the parameters ρ and σ, which determine skewness and kurtosis.

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

Column (8) imposes $\eta = \xi = 0$. This is another very strong restriction, which implies that $\gamma = -1$ from equations (19) and (22), or equivalently $\mu = 1$. Similar to the results in column (6), these strong joint restrictions impact the estimates of the drift parameters κ and θ.

Note that while the parameters characterizing the pricing kernel are very different in columns (1)-(8), the risk-neutral parameter estimates κ^*, θ^*, σ, and ρ remain remarkably similar across columns. The remaining physical parameters κ and θ differ, which explains the large differences in the pricing kernel parameters γ, ξ, and η, or equivalently the risk premium parameters μ and λ.

The bottom rows present the log-likelihood, as well as the p-values for tests of the restricted pricing kernels against the unrestricted exponential-affine kernel in column (1). Surprisingly, while imposing restrictions on the exponential-affine kernel results in very different implied risk premia and preference parameters, the resulting differences in the log-likelihoods are small, and in some cases, the likelihood ratio tests do not indicate rejection at conventional significance levels. Even for the double restriction $\gamma = \xi = 0$ in column (6), which is extreme from an economic perspective, the likelihood does not change by a large amount. We discuss the implications of these findings in more detail in Section 6.

5.4 More General Pricing Kernels and Affine Prices of Risk

In this section, we discuss results based on the more general kernel in equation (23), as well as results based on the completely affine and affine specifications of the price of risk. Columns (9)-(10) in Table 4 present results for the pricing kernel in equation (23). The physical mean reversion in column (9) in Table 4 is larger compared to that for the exponential-affine kernel in columns (1)-(8) in Table 3, and the risk-neutral mean-reversion is similar. The other parameters are very similar. The log-likelihoods are again very similar to column (1).

Now consider the completely affine specification of the price of risk with an equity premium $\mu v(t)$ and a variance risk premium $\lambda v(t)$. It turns out that the estimation results are identical to the results of the exponential-affine specification in column (1). This finding obtains mechanically because the specifications are isomorphic due to the number of (free) parameters in both specifica-

tions. For the exponential-affine kernel in column (1), the free parameters are γ, ξ, and the physical parameters. For the completely affine price of risk specification, the free parameters are μ, λ, and the physical parameters. But recall that γ and ξ are restricted by μ and λ through equations (19) and (20) and vice versa. In summary, we use the same number of free parameters in estimation and the parameters are subject to the same restrictions (19) and (20). This may seem confusing because the pricing kernel contains two extra parameters, β and η. However, these parameters are calibrated from the risk-free rate under the martingale condition and do not affect MLE estimation.

We now turn to the affine specification of the price of risk with an equity premium $\mu_0 + \mu_1 v(t)$ and a variance risk premium $\lambda_0 + \lambda_1 v(t)$. The results are presented in column (11) in Table 4. Note that the risk-neutral parameter estimates κ^*, θ^*, σ, and ρ are once again very similar to the estimates in Table 3. The most important difference is that the risk-neutral mean reversion (κ^*) 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 (11) exceeds the log-likelihood in column (1). While the improvement in log-likelihood may seem relatively small, the difference is statistically significant using a likelihood ratio test. Moreover, the improvement in likelihood is relevant given the small differences in columns (1)-(10). 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, as the option log likelihood is not very different from the one in column (1). The results for the more general kernel in equation (23) in columns (9) and (10) also imply an intercept in the risk premiums, but the likelihood in column (11) is higher than the one in columns (9) and (10).

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 affine specification can outperform the completely affine specification. However, the improvement in fit in column (11)

compared to column (1) ignores the no-arbitrage restrictions imposed in columns (9) and (10).

Finally, while the estimates in column (9) impose the no-arbitrage restrictions emanating from the martingale condition in Section 4.3, they ignore that the more general kernel in equation (23) 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 $2\kappa\theta = 2\kappa^*\theta^* > \sigma^2$. 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 (23). Column (12) shows that this additional no-arbitrage restriction clearly affects model fit and it also affects the parameter estimates, most notably the estimate of ξ.

We next proceed by investigating the economic implications of these no-arbitrage restrictions. We show that the economic implications of (relaxing) this restriction are critical. We also show that while some pricing kernels are difficult to distinguish statistically, they 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 typically not downward sloping, which means that existing findings of non-monotonic kernels do not constitute puzzles.

We 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 return and variance.

6.1 Risk Premia

The most striking finding from Tables 3 and 4 is that the risk-neutral parameters are very similar across columns and that the differences in the log likelihoods are also small and sometimes statistically insignificant. 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.[17]

We now turn to the economic implications of the different pricing kernel restrictions rather than the statistical fit. We find a diametrically opposite result, namely that the different pricing kernels result in widely different economic implications. Panel B of Tables 3 and 4 present descriptive statistics for the equity and variance risk premia. Recall from equations (15), (16), (24), and (25) 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.

The restrictions in columns (2) and (6) imply that the equity premium is zero. However, some of the other strong economic restrictions such as $\lambda = 0$ in column (3) and $\gamma = 0$ in column (4) also lead to substantially lower equity premia. Setting the variance preference parameter ξ equal to zero in column (5) and setting the path dependence parameter η equal to zero in column (7) results in higher equity risk 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. Rather

[17] 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.

than compare this to a sample estimate of the variance risk premium, which requires additional assumptions, note that 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 in the estimation indicates that its size is as large as or slightly larger than the size of the index variance itself. This finding is similar to the one in 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 (11) yield similar estimates. The other restrictions result 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 based on 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.

We conclude that the unrestricted exponential-affine pricing kernel in column (1) of Table 3 has economic implications that are intuitively plausible. Restricted versions of this kernel often result in implausible economic implications, as measured by risk premia and preference parameters. The restrictions do not result in a much worse fit as measured by statistical criteria, but this may be due to the fact that these tests 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. Next we explore the implications of our findings for the literature that characterizes

the shape of the pricing kernel.

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

$$E_t(R(t+\Delta)) = 1 + [r + \mu v(t)]\Delta, \tag{34}$$

$$\sigma_t(R(t+\Delta)) = \sqrt{v(t)\Delta}. \tag{35}$$

For the more general pricing kernels in equation (23) and the affine price of risk specification, the conditional mean of the daily market return is given by

$$E_t(R(t+\Delta)) = 1 + [r + \mu_0 + \mu_1 v(t)]\Delta. \tag{36}$$

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 Tables 3 and 4 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 a yearly Sharpe ratio of 0.411. This is close to the sample average of 0.439. The more general kernel in column (9) also yields an average Sharpe ratio that is close to the market 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 (11), 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 yearly Sharpe ratio in our sample is 0.449, which is unrealistic and much

higher compared to the kernels in columns (1) and (9).

6.3 The Pricing Kernel Puzzle

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 (12). Based on the kernel in equation (12), we expect a downward sloping log pricing kernel as a function of the log index return.

Starting with Jackwerth (2000), several studies document that both the conditional and the unconditional pricing kernel 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 (11) conforms to economic theory with respect to aggregate wealth, that is, if $\gamma < 0$ in equation (11). A U-shaped pricing kernel, therefore, arises naturally and should not be thought of as an anomaly. We now investigate this 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.

6.4 A Multivariate Representation of the Estimated Pricing Kernel

We start our empirical investigation in the simplest possible way, by depicting the pricing kernel as a function of the index return and variance innovations. Figure 2 provides a scatterplot of the

estimated exponential-affine kernel in equation (11), based on the MLE estimates in column (1) of Table 3, as a function of the log stock return $\log R(t)$ and the daily change in variance $v(t)-v(t-1)$. 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 negative estimate of the index return preference γ and the positive estimate of the variance preference ξ. The implied correlation coefficient ρ 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 negative γ and indirectly through the channel of the positive ξ combined with the negative ρ. 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 now proceed to a more formal analysis of the relation between the pricing kernel and the state variables.

6.5 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 estimates in Tables 2, 3, and 4.

Specifically, we proceed as follows. Given the parametric stock return dynamics in equations (1) and (2), we can generate the probability densities 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_τ^{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:

$$\begin{aligned} Pr(\log R(\tau) = x) &= \frac{dPr(\log R(\tau) \leq x)}{dx} \\ &= -\frac{1}{\pi} \int_0^\infty Re \left[-e^{-i\phi x + C + Dv} \right] d\phi, \end{aligned} \tag{37}$$

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

We calculate the numerical integral in equation (37) 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 θ 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:

$$\log M(\tau) - \log M(0) = -\bar{r}\tau + \log Q(\log R(\tau)|v(0) = \theta) - \log P(\log R(\tau)|v(0) = \theta). \tag{38}$$

By changing the values 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 column of Figures 3 and 4 reports the results. For convenience, we express the x-axis in standard deviations of the log return from the expected log return.[18] Results are

[18]Since we already have the physical probability density, we can calculate the expectation and variance of the log

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 appendix presents additional results for one-month pricing kernels.

6.6 Recovering Pricing Kernel Dynamics

Most of the literature studies (differences in) pricing kernels using the equivalent of the right-side graphs in Figures 3 and 4, that is, as a univariate function of the log return (aggregate wealth). However, our proposed pricing kernels (11) and (23) 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 column in Figures 3 and 4 is 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. See Chernov (2003) for a related exercise.

Unfortunately it is not straightforward to perform a similar exercise 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:

$$\log M(t+\Delta)-\log M(t) = -r\Delta+\log Q(\log R(t+\Delta), v(t+\Delta)|v(t))-\log P(\log R(t+\Delta), v(t+\Delta)|v(t)), \tag{39}$$

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

---

return as:

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

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 (37), because we need to use the realization of both $\log R(t+\Delta)$ and $v(t+\Delta)$.[19] We can also find a quasi closed form expression (up to numerically computing double integrals for the joint probability). Appendix C provides the details.

6.7 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 of Figure 3 is based on the estimates 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 (12).

The differences between the results for 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 γ in column (5) of Table 3 is negative, it is downward sloping, consistent with the intuition of decreasing marginal utility of wealth. For the exponential-affine kernel in Panel A, state prices are no longer linearly downward sloping as a function of log aggregate wealth. Because the estimate of ξ is positive (and ρ is negative), we now have a convex function. However, the estimate of ξ does not seem to be 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 ξ, and we obtain a U-shaped pricing kernel.[20]

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 A. The realized kernels fluctuate a lot over

[19] The marginal probability of the log return is the expectation of the joint probability with respect to the variance $v(t+\Delta)$.

[20] 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.

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 A and D is that the time variation in the pricing kernel in Panel A 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 A and C, the outliers in Panel C are more pronounced due to the larger estimate of ξ.

Panel B 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. Both sets of 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 has a W-shape. This is consistent with some existing findings (Cuesdeanu and Jackwerth, 2018), although a U-shape is more common.[21]

The dramatic differences between the marginal pricing kernel in Panel B and the ones in Panels A, C and D are confirmed by the time path of the pricing kernel.[22] The only thing the path in Panel B 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 B. More importantly, the outliers in the pricing kernel are of an entirely different order of magnitude compared to the exponential-affine pricing kernel. 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.

Figure 4 reports on the kernels corresponding to the affine specification. Panel A reports on the pricing kernel in equation (23), corresponding to the estimates in column (9) of Table 4. Panel B imposes $\xi = 0$, as in column (10) of Table 4. Panel C corresponds to the affine price of risk specification in column (11) of Table 4. Recall that these estimates are obtained under the affine

[21]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, and to some extent, this is obscured in Figure 3 because we do not report confidence intervals, and we impose a parametric structure.

[22]Note that the $y-$axis in Panel B is scaled differently.

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 C of Figure 4 is very different from the ones in Panels A and B. Note also the difference with the exponentially-affine kernel in Panel A of Figure 3; the shape is actually more similar to the unrestricted pricing kernel in Panel B of Figure 3. This finding is surprising given that Panel B of Figure 3 does not impose any restrictions across measures, while Panel C of Figure 4 does.

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

Finally, we discuss two more technical cases. First, Panel D of Figure 4 corresponds to column (12) in Table 4. It imposes the Feller condition on the kernel in equation (23). A comparison with Panel A of Figure 4 indicates that this restriction strongly affects the economic implications, consistent with the differences in parameter estimates reported in Table 4. Second, kernels that incorporate path-dependence (*i.e.* nonzero η) 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.[23]

Our main conclusions from Figures 3 and 4 are as follows: 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)

[23]This implication results from the zero η in conjunction with the negative estimate of ξ.

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; 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 C 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.

6.8 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 ($\xi[v(t) - v(t-1)]$). We present 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 the most 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. Economic intuition specifies how pricing kernels relate to relevant state variables and suggests that kernels should be well-behaved and smooth as a function of these state variables.

Options play an important role in the empirical analysis of pricing kernels, because they can be used to estimate the risk-neutral probabilities that are required to identify the pricing kernel. Equity index options are particularly valuable, because the return on the underlying can be thought of as an approximation to the return on aggregate wealth. These insights therefore provide a motivation for the study of index option pricing that transcends the more narrow question of derivatives pricing. They highlight the importance of derivatives pricing for asset pricing and for macro-finance and economics more in general.

In light of this, it is surprising that the existing literature on parametric dynamic index option pricing models typically does not explicitly specify the pricing kernel. It characterizes the kernel indirectly by specifying prices of risk or defines it implicitly as the ratio of risk-neutral and physical probabilities. We propose a class of pricing kernels that are consistent with the conventional assumption of affine dynamics under the physical and risk-neutral measure in the square root stochastic volatility model. These kernels are volatility-dependent by construction and are therefore especially useful to analyze the impact of stock market volatility risk on state prices and market risk.

We estimate the resulting models subject to various restrictions on the pricing kernel, using index returns and option prices. We show that affine risk premia can produce kernels with counter-intuitive economic properties and that 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 find that 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. Existing tests have low power to statistically distinguish different pricing kernels because the identification of the pricing kernel is equivalent to the estimation of conditional risk premia. These findings extend Merton's (1980) observations 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. We need a lot of data to estimate unconditional average returns

on stocks or options, and even more data to estimate conditional returns. Moreover, plain vanilla option prices are sensitive to the probabilities at expiration, but are 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.

The analysis in this paper is based on a single-factor diffusion model. State-of-the-art models in the option valuation literature contain multiple volatility factors, as well as jumps in returns and volatility. It would be interesting to characterize and analyze the pricing kernels that are consistent with these models and no-arbitrage. One important question for future research is whether these models facilitate identification of risk premia, and whether pricing kernels in high-dimensional models are meaningfully different from pricing kernels in simpler models. A related question is whether other factors such as tail factors or factors related to intermediary risk can help identify risk premia. Finally, our empirical analysis uses a large cross-section of options, but it is limited to plain-vanilla index option contracts. This leaves open the possibility that other option contracts may facilitate the estimation and identification of pricing kernels and equity and variance risk premia.

References

Adrian, Tobias, Richard K. Crump, and Emanuel Moench, 2013, Pricing the Term Structure with Linear Regressions, *Journal of Financial Economics* 110, 110–138.

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.

Ang, Andrew, Geert Bekaert, and Min Wei, 2008, The Term Structure of Real Rates and Expected Inflation, *Journal of Finance* 63, 797–849.

Ang, Andrew, and Monika Piazzesi, 2003, A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, *Journal of Monetary Economics* 50, 745–787.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical Performance of Alternative Option Pricing Models, *Journal of Finance* 52, 2003–2049.

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, 2020, On Sources of Risk Premia in Representative Agent Models, *Working Paper, Available at SSRN 3452743.*

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.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model Specification and Risk Premia: 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.

Drechsler, Itamar, 2013, Uncertainty, Time-Varying Fear, and Asset Prices, *Journal of Finance* 68, 1843–1889.

Duffee, Gregory R., 2002, Term Premia and Interest Rate Forecasts in Affine Models, *Journal of Finance* 57, 369–443.

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, Available at SSRN 3520256.*

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.

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.

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

Panel A: Return-Based Physical Parameters							
μ	κ	θ	σ	ρ	η_0	η_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							
μ^*	κ^*	θ^*	σ	ρ	η_0	η_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

Panel A. Parameter Estimates								
	Exponential-Affine							
	No Restr.	$\mu = 0$	$\lambda = 0$	$\gamma = 0$	$\xi = 0$	$\gamma = \xi = 0$	$\eta = 0$	$\eta = \xi = 0$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Risk Preference Parameters								
γ	-1.3929	1.0921	-1.3940	0	-2.4850	0	-2.2311	-1
	(0.2266)				(0.2182)			
ξ	1.9474	1.9472	-1.4776	3.4244	0	0	-0.6242	0
	(0.4037)	(0.2034)	(0.8098)	(0.5430)			(0.0511)	
(Implied Average) β	-0.1800	-0.2448	0.1546	-0.3606	0.0387	-0.0261	0.0931	0
(Implied) η	4.9671	3.1214	-1.9492	6.9151	1.8451	0	0	0
Implied Risk Premia Parameters								
μ	2.4852	0	0.5652	1.9206	2.4850	0	1.8810	1.0000
λ	-1.8115	-0.4177	0	-1.8116	-1.3937	0	-0.9211	-0.5609
P-Parameters								
κ	2.9252	1.5314	1.1141	2.9252	2.5076	1.1139	2.0352	1.6748
	(0.2082)	(0.0727)	(0.0558)	(0.2933)	(0.1255)	(0.0558)	(0.0855)	(0.0574)
θ	0.0334	0.0638	0.0877	0.0334	0.0390	0.0877	0.0480	0.0583
	(0.0024)	(0.0023)	(0.0030)	(0.0033)	(0.0019)	(0.0030)	(0.0012)	(0.0011)
σ	0.7274	0.7274	0.7274	0.7273	0.7274	0.7274	0.7274	0.7274
	(0.0082)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)
ρ	-0.7711	-0.7711	-0.7711	-0.7711	-0.7711	-0.7711	-0.7711	-0.7711
	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)
η_0	-0.0047	-0.0047	-0.0047	-0.0047	-0.0047	-0.0047	-0.0047	-0.0047
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
η_1	0.8881	0.8881	0.8881	0.8880	0.8881	0.8881	0.8881	0.8881
	(0.0059)	(0.0060)	(0.0060)	(0.0059)	(0.0060)	(0.0060)	(0.0060)	(0.0060)
Q-Parameters								
κ^*	1.1137	1.1137	1.1141	1.1135	1.1139	1.1139	1.1141	1.1139
θ^*	0.0877	0.0877	0.0877	0.0877	0.0877	0.0877	0.0877	0.0877
Likelihood								
$\log L^R$	54,443.42	54,440.38	54,440.30	54,443.11	54,443.02	54,439.98	54,442.60	54,441.93
$\log L^O$	156,578.03	156,578.03	156,578.10	156,577.96	156,578.04	156,578.03	156,577.97	156,578.03
$\log L^R + \log L^O$	211,021.45	211,018.41	211,018.40	211,021.07	211,021.05	211,018.01	211,020.58	211,019.96
P-Value for the LR Test with (1)		0.0137	0.0135	0.3833	0.3739	0.0321	0.1862	0.2254
Panel B. Economic Implications								
Instantaneous Equity Premium	0.0832	0	0.0189	0.0643	0.0832	0	0.0629	0.0335
Instantaneous Variance Premium	-0.0606	-0.0140	0	-0.0606	-0.0466	0	-0.0308	-0.0188
Daily Sharpe Ratio	0.0259	0	0.0059	0.0200	0.0259	0	0.0196	0.0104

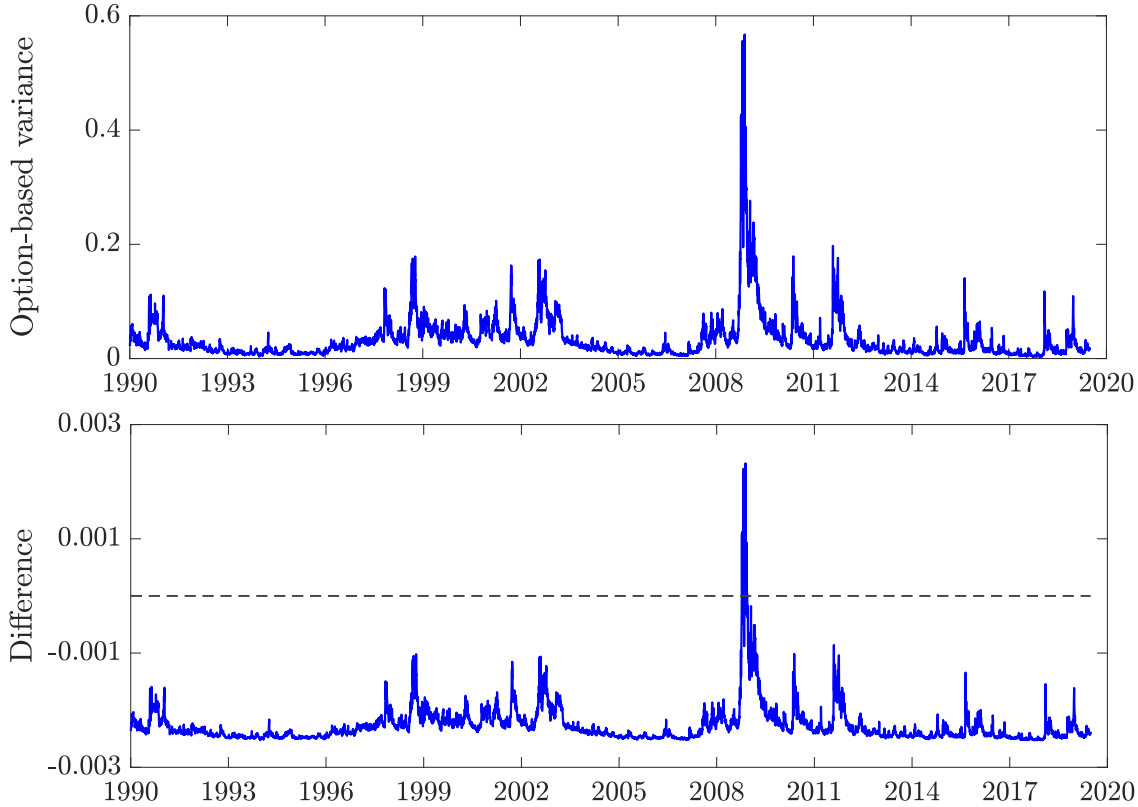
Notes: Panel A reports the results of joint maximum likelihood estimation. The risk preference parameters γ and ξ as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions in columns (1) to (8). Robust standard errors are in parentheses. We estimate the model for eight different specifications. We report the log likelihood for each specification and the P-values for the LR test against the specification in column (1). Panel B presents the sample means for the implied instantaneous equity and variance risk premia and the daily Sharpe ratio.

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

Panel A. Parameter Estimates					
	PK in Eq. (23)		Affine POR	Feller Condition Imposed	
	No Restr.	$\xi = 0$	No Restr.	PK	POR
	(9)	(10)	(11)	(12)	(13)
Risk Preference Parameters					
γ	-1.4059	-2.5693		-1.4786	
	(0.2282)	(0.1980)		(0.8652)	
ξ	2.9002	0		11.2114	
	(0.2599)			(4.3118)	
α	-0.0324	-0.0050		-0.3150	
	(0.0048)	(0.0004)		(0.0424)	
Implied Risk Premia Parameters					
μ_0	-0.0182	-0.0028	0.0485	-0.1764	-0.1094
			(0.0202)		(0.0385)
μ_1	3.0324	2.5693	1.0393	7.7565	5.6322
			(1.2862)		(4.4872)
λ_0	0.0172	0.0027	0.0172	0.1660	0.1659
			(0.0054)		(0.0204)
λ_1	-2.3227	-1.4410	-2.3297	-6.7342	-6.6525
			(0.5833)		(3.7376)
P-Parameters					
κ	3.4361	2.5549	3.4423	7.8438	7.7613
	(0.1714)	(0.1165)	(0.5790)	(1.8485)	(3.7376)
θ	0.0334	0.0393	0.0334	0.0336	0.0339
	(0.0015)	(0.0018)	(0.0072)	(0.0058)	(0.0139)
σ	0.7273	0.7274	0.7272	0.7258	0.7256
	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)
ρ	-0.7711	-0.7711	-0.7711	-0.7715	-0.7716
	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)
η_0	-0.0047	-0.0047	-0.0047	-0.0047	-0.0047
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
η_1	0.8880	0.8881	0.8880	0.8874	0.8874
	(0.0060)	(0.0059)	(0.0059)	(0.0060)	(0.0059)
Q-Parameters					
κ^*	1.1134	1.1139	1.1126	1.1096	1.1089
θ^*	0.0877	0.0877	0.0878	0.0878	0.0878
Likelihood					
$\log L^R$	54,443.71	54,443.03	54,449.61	54,430.16	54,436.11
$\log L^O$	156,578.00	156,578.03	156,577.08	156,572.56	156,571.61
$\log L^R + \log L^O$	211,021.71	211,021.06	211,026.69	211,002.72	211,007.71
P-Value for the LR Test with (11)	0.0016	0.0036			
Panel B. Economic Implications					
Instantaneous Equity Premium	0.0833	0.0832	0.0833	0.0831	0.0791
Instantaneous Variance Premium	-0.0606	-0.0456	-0.0607	-0.0594	-0.0567
Daily Sharpe Ratio	0.0236	0.0255	0.0329	0.0536	0.0364

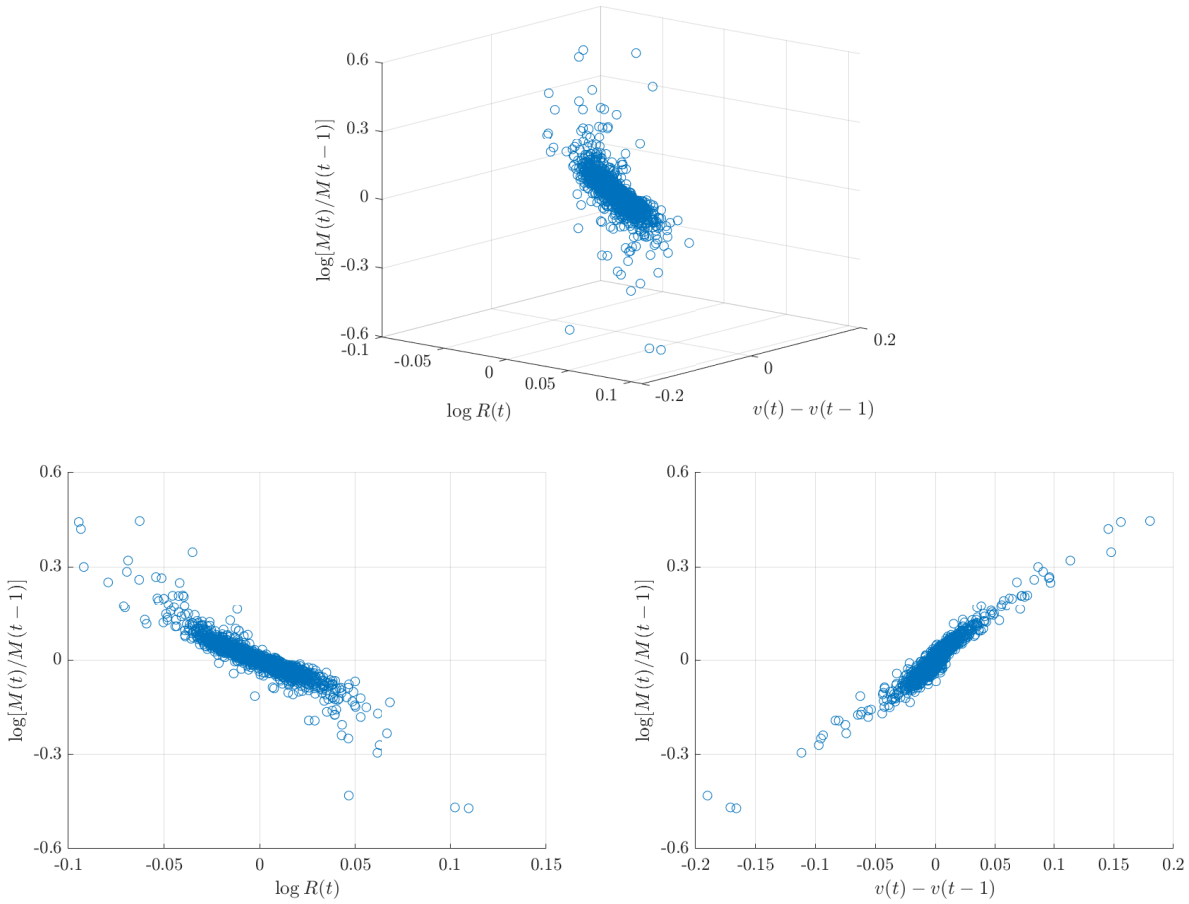
Notes: Panel A reports the results of joint maximum likelihood estimation. The risk preference parameters γ, ξ, and α as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions. Robust standard errors are in parentheses. We report on the more general pricing kernel (PK) from equation (23) in column (9), the more general PK subject to $\xi = 0$ in column (10), and the affine price of risk (POR) specification in column (11). Columns (12) and (13) report on the specifications in (9) and (11) with the Feller condition imposed. For columns (9) and (10), we report the P-values for the LR test against the specification in column (11). Panel B presents the sample means for the implied instantaneous equity and variance risk premia and the daily 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 $v(t) = \eta_0 + \eta_1 \mathrm{VIX}^2(t)$, the option-based variance is computed as $v^*(t) = \eta_0^* + \eta_1^* \mathrm{VIX}^2(t)$, and the difference is computed as $v(t) - v^*(t)$, where $\eta_0,\ \eta_1,\ \eta_0^*$, and η_1^* are from Table 2.

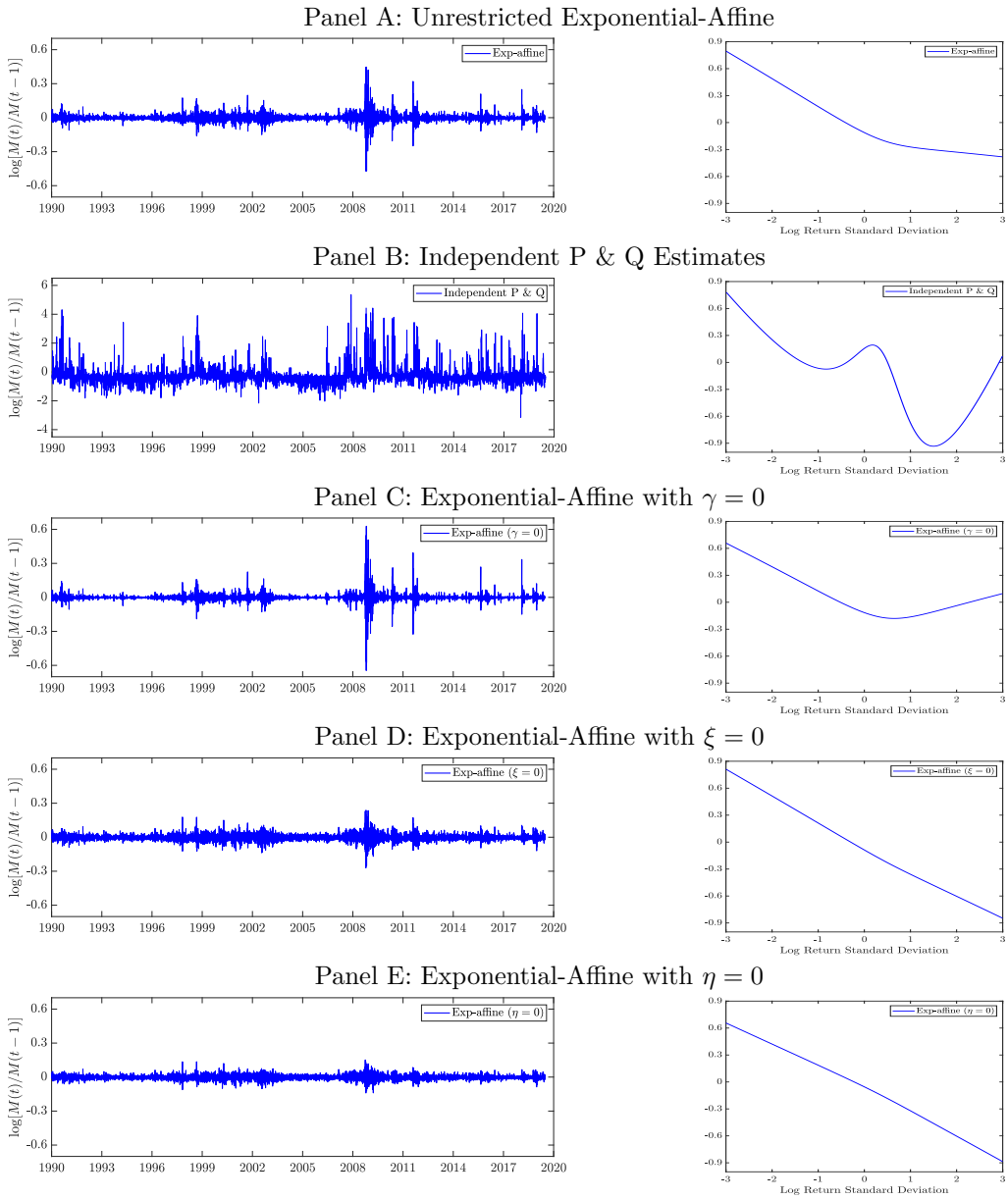
Figure 2: The Log Pricing Kernel as a Function of Log Return and Variance



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 Pricing Kernels

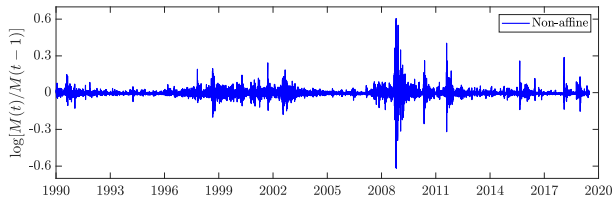


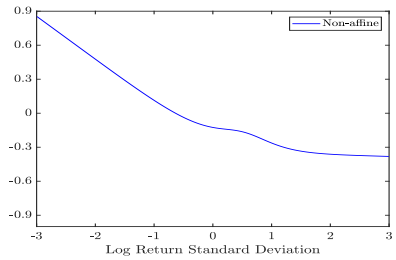
Notes: We plot the time series of the daily log pricing kernels and the implied 6-month log pricing kernels for the following five specifications: unrestricted exponential-affine (Panel A), independent estimation of the P- and Q-parameters (Panel B), the exponential-affine specifications with a restriction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

Figure 4: Other Log Pricing Kernels

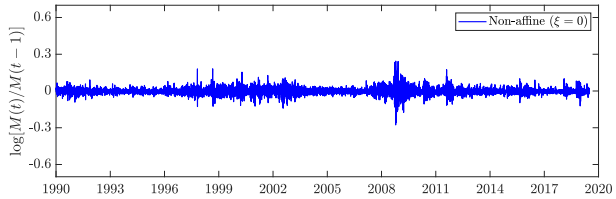
Daily (Realized) Log Pricing Kernels 1990-2019 | Implied Six-Month Log Pricing Kernels

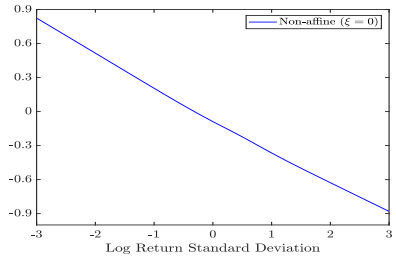
Panel A: Unrestricted PK from Eq. (23)



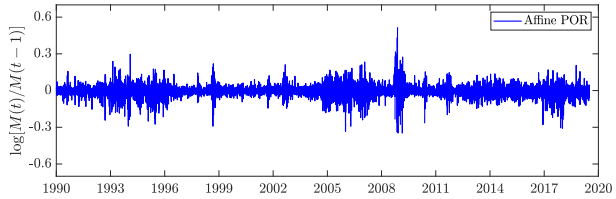


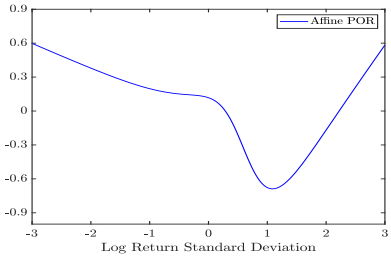
Panel B: PK from Eq. (23) with $\xi = 0$



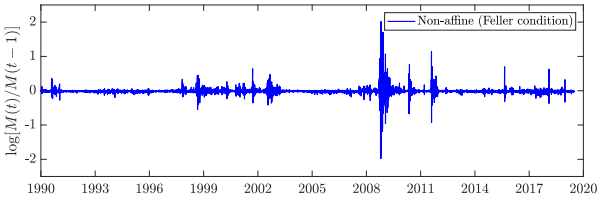


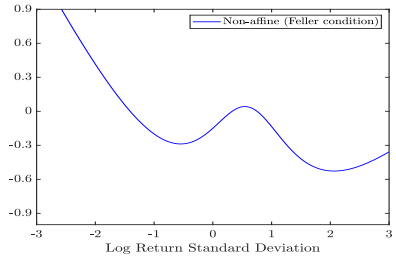
Panel C: Affine Price of Risk





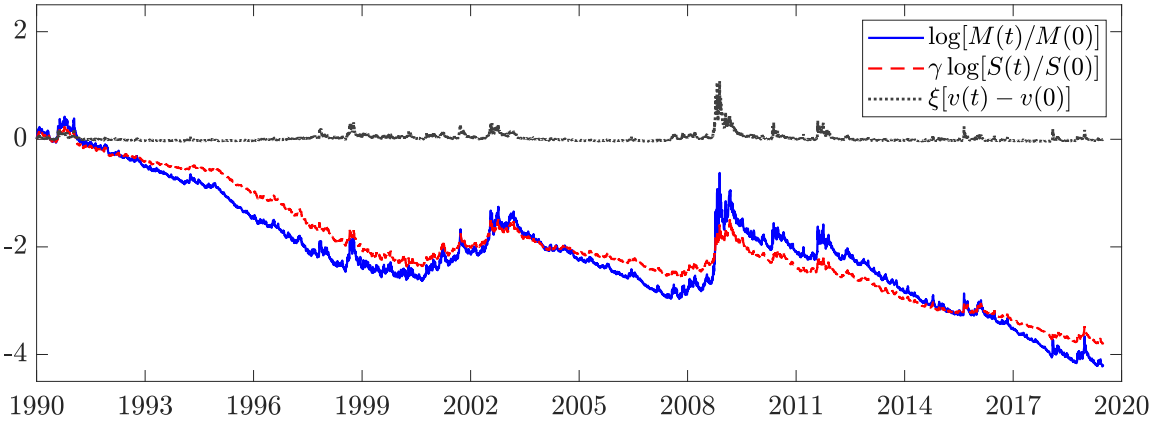
Panel D: PK from Eq. (23) with the Feller No-Arbitrage Condition





Notes: We plot the time series of the daily log pricing kernels and the implied 6-month log pricing kernels for the following four specifications: the unrestricted pricing kernel from equation (23) (Panel A), the kernel from equation (23) with restriction $\xi = 0$ (Panel B), the affine price of risk specification (Panel C), and the pricing kernel from equation (23) with the Feller condition imposed (Panel D). Parameter values are from Table 4. For the implied 6-month kernels, the x-axis represents the standard deviations of log return from the expected 6-month log return.

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.

Appendix

A No-Arbitrage Restrictions for the Exponential-Affine Pricing Kernel

Applying Ito's lemma to equation (11) 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

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

In order to find the restrictions on β and η, we use the fact that $e^{rt}M(t)$ is a martingale – the drift term of $d\left(e^{rt}M(t)\right)$ must be zero. That is, from equation (A.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,

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

Since equation (A.2) must hold for any values of v, it implies that (A) and (B) must be zero. Therefore, we have the following restrictions on β and η:

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

Furthermore, to find the expression for the equity risk premium parameter μ as a function of the preference parameters, we employ the fact that $S(t)M(t)$ is a martingale and calculate the equity risk premium. Since $\frac{d(SM)}{SM} = \frac{dS}{S} + \frac{dM}{M} + \frac{dS}{S}\frac{dM}{M}$ and $E\left[\frac{d(SM)}{SM}\right]$ must be zero, the equity risk premium follows

$$E\left[\frac{dS}{S}\right] - rdt = -E\left[\frac{dS}{S}\frac{dM}{M}\right].$$

Note that we have used the fact that the drift of M is equal to $-rMdt$. Then, we finally obtain

$$E\left[\frac{dS}{S}\right] - rdt = (-\gamma - \rho\sigma\xi)v.$$

Since the equity risk premium should be equal to μv under the physical process of equation (2), we have

$$\mu = -\gamma - \rho\sigma\xi.$$

Likewise, we find the expression for the variance risk premium as $-E\left[dv\frac{dM}{M}\right] = (-\gamma\rho\sigma - \xi\sigma^2)v$. When the variance risk premium is an affine function of v, say λv, we have

$$\lambda = -\gamma\rho\sigma - \xi\sigma^2.$$

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

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

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

$$\begin{aligned} \kappa &= \kappa^* - \lambda, \\ \theta &= \frac{\kappa^*\theta^*}{\kappa}. \end{aligned}$$

B No-Arbitrage Restrictions for the More General Pricing Kernel

We follow the same step as in Appendix A to find the no-arbitrage restrictions between the risk-neutral dynamics in equation (1) and physical dynamics in equation (26) under the more general pricing kernel. Applying Ito's lemma to equation (23) 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 $d\log v = \frac{1}{v}\left(\kappa(\theta - v) - \frac{1}{2}\sigma^2\right)dt + \frac{\sigma}{\sqrt{v}}dz_2$. Again by Ito's lemma, we have

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

The drift of M should be equal to $-rMdt$. By rearranging the drift term in equation (B.1),

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

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

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

Moreover, by calculating $-E\left[\frac{dS}{S}\frac{dM}{M}\right]$ and $-E\left[dv\frac{dM}{M}\right]$ for the equity and variance risk premia, respectively, we obtain the risk premia expressed as $\mu_0 + \mu_1 v(t)$ and $\lambda_0 + \lambda_1 v(t)$, where

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

The physical dynamics in equation (26) are restricted according to

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

C A Closed-Form Expression for the Joint Probability Distribution

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

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

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

$$g_0^{CH}(\phi_x, \phi_v | x, v) = e^{i\phi_x x + i\phi_v v}. \tag{C.2}$$

Suppose g has the following functional form:

$$g_\tau^{CH}(\phi_x, \phi_v | x, v) = e^{G(\tau)+H(\tau)v+i\phi_x x}. \tag{C.3}$$

By substituting equation (C.3) into equation (C.1), we get the following ordinary differential equations (ODEs) for $G(\tau)$ and $H(\tau)$:

$$\begin{aligned} G'(\tau) &= r\phi_x i + aH(\tau), \\ H'(\tau) &= -\frac{1}{2}\phi_x^2 + \rho\sigma\phi_x i H(\tau) + \frac{1}{2}\sigma^2 H(\tau)^2 + u\phi_x i - bH(\tau), \end{aligned} \tag{C.4}$$

and the terminal conditions of the ODEs are $G(0) = 0$ and $H(0) = i\phi_v$ inferred from equation (C.2). This system of ODEs expressed in equation (C.4) has the following closed-form solution:

$$\begin{aligned} H(\tau) &= \frac{A_2(D_m X - iA_2 Y)\phi_v - 2iA_1(-D_m - 2A_3\phi_v + (Y-1)(D_m - 2A_3\phi_v))}{iA_2 D_m X - A_2^2 Y + 4A_1A_3Y - 2A_3 D_m X\phi_v}, \\ G(\tau) &= r\phi_x i\tau \\ &\quad + \frac{1}{4A_3}\Bigg[-2aA_2\tau - \frac{2ia\tau(A_2^2 - 4A_1A_2)}{D_m} + 2ia\arctan\left(\frac{A_2A_3X^2\phi_v}{A_2^2(Y-1) - A_1A_3Y^2 + A_3^2X^2 + \phi_v^2}\right) \\ &\quad + \frac{4aD_p}{D_m}\arctan\left(\frac{iA_2^2 + 2A_2A_3X\phi_v - 2iA_3(A_1Y - A_3X\phi_v^2)}{D_p(A_2 + 2iA_3\phi_v)}\right) \\ &\quad - \frac{4iaD_p}{D_m}\operatorname{arctanh}\left(\frac{D_p}{A_2 + 2iA_3\phi_v}\right) + a\log(D_p) \\ &\quad - a\log\left(A_2^4(Z-1) + A_1^2A_3^2Y^4 + A_2^2A_3^2X^2Z\phi_v^2 + A_3^4X^4\phi_v^4 - 2A_1A_3Y^2(A_2^2(Y-1) + A_3^2X^2\phi_v^2)\right)\Bigg], \end{aligned}$$

where

$$A_1 = -\tfrac{1}{2}\phi_x^2 + u\phi_x i, \quad A_2 = \rho\sigma\phi_x i - b, \quad A_3 = \tfrac{1}{2}\sigma^2$$

$$D_m = \sqrt{-A_2^2 + 4A_1A_3}, \quad D_p = \sqrt{A_2^2 - 4A_1A_3}$$

$$X = -1 + e^{iD_m\tau}, \quad Y = 1 + e^{iD_m\tau}, \quad Z = 1 + e^{2iD_m\tau}.$$

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,

$$Pr(x, v|x(t), v(t)) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\phi_x x - i\phi_v v} g_\tau^{CH}(\phi_x, \phi_v|x(t), v(t)) d\phi_x d\phi_v \tag{C.5}$$

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 τ-horizon log return. Then, equation (C.5) can be written as

$$Pr(x, v|v(t)) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\phi_x x - i\phi_v v + G(\tau|\phi_x,\phi_v) + H(\tau|\phi_x,\phi_v)v(t)} d\phi_x d\phi_v. \tag{C.6}$$

Numerical integration of equation (C.6) is not tractable because the exponential function to be integrated decays very slowly, especially over the ϕ_v dimension. To expedite its calculation, we use the following weight function for the j-th grid point of ϕ_v, w_j:

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

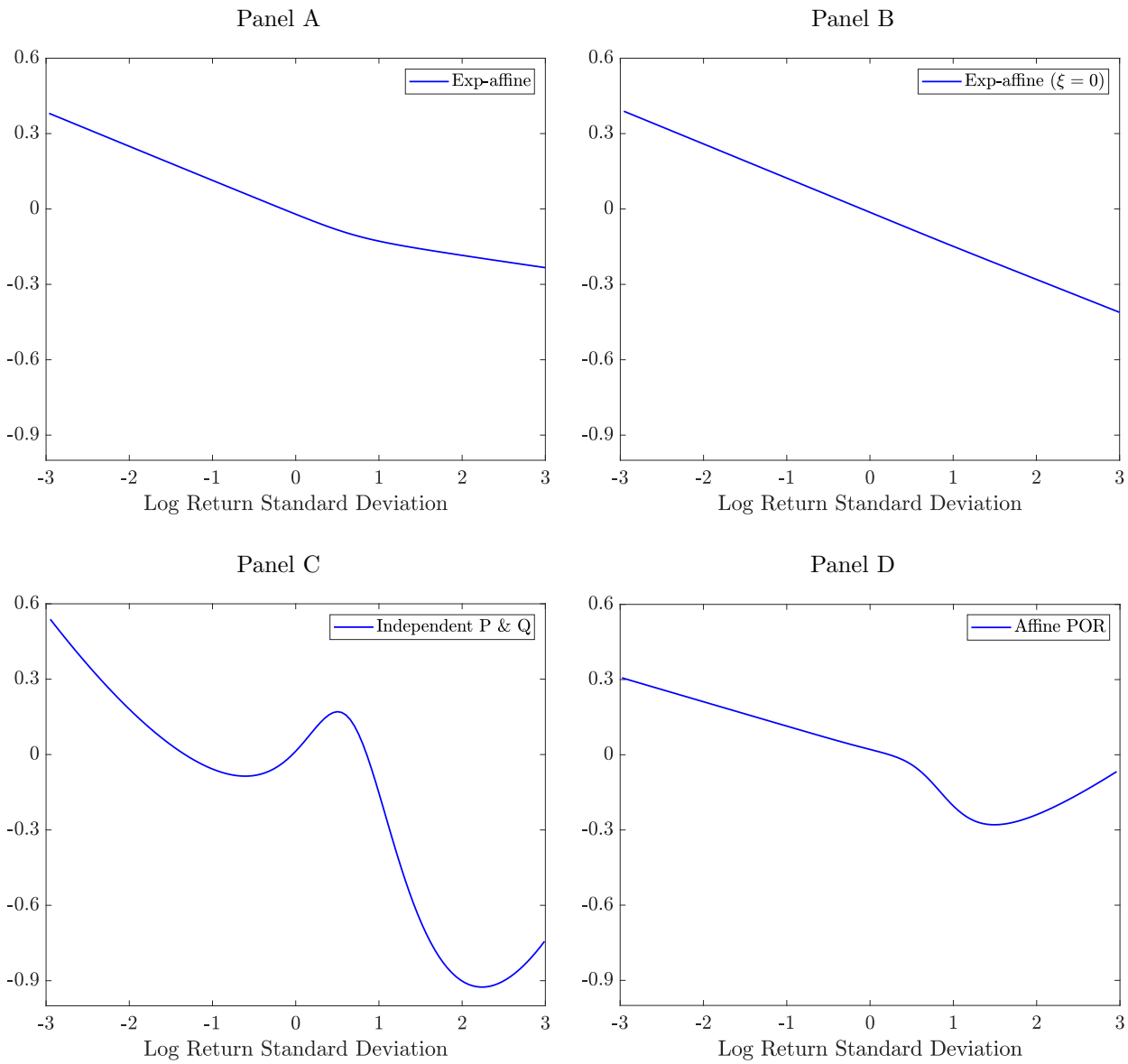
where $2N_v$ is the number of grid points of ϕ_v. This weight function forces the exponential function to decay much faster (see, for example, Ooura, 2001).[24] Hence, we numerically compute equation (C.6) as

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

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

[24] 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 standard deviations of log return from the expected 1-month log return.