Volatility Risk and Monotonic Pricing Kernels*

Steven Heston
University of Maryland

Kris Jacobs
University of Houston

Hyung Joo Kim
Federal Reserve Board

October 16, 2025

Abstract

The option valuation literature often specifies the pricing kernel indirectly by defining the price of risk associated with the state variables. Some specifications amount to pricing kernels that are inconsistent with economic theory or result in identification problems. The resulting time series of the pricing kernel is excessively volatile in good times, and the marginal kernel has multiple inflection points. Explicitly specifying the pricing kernel avoids these problems. A parsimonious exponential-affine kernel monotonic in the state variables resolves anomalies. It achieves good option fit, provides reasonable estimates of equity and variance risk premiums, and improved estimates of expected option returns.

*Heston: sheston@umd.edu; Jacobs: kjacobs@bauer.uh.edu; Kim: hyungjoo.kim@frb.gov. An earlier version of this paper was circulated under the title "The Pricing Kernel in Options". We would like to thank Caio Almeida, David Bates, Ian Dew-Becker, Hitesh Doshi, Bjørn Eraker, Xiaohui Gao, Stefano Giglio, Massimo Guidolin, Alex Kostakis, Paola Pederzoli, Jean-Paul Renne, seminar participants at the 2023 AFA Conference, the 2022 Finance Down Under Conference, the 2022 SoFiE Conference, the 2022 FMA Conference on Derivatives and Volatility, K.U. Leuven, Syracuse University, the Universities of Houston and Liverpool, and especially our discussants Gurdip Bakshi, Mikhail Chernov, and Jeroen Dalderop for helpful comments. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by the Federal Reserve Board or other members of its staff.

1 Introduction

This paper addresses the specification of the pricing kernel in dynamic option pricing models. Derivatives pricing often specifies reduced form risk premia to implicitly define the pricing kernel as the ratio of the risk-neutral and physical probabilities. In the Black-Scholes model, this approach merely requires the specification of the price per unit of equity risk, but in more complex state-of-the-art models with multiple state variables, it requires the specification of a price of risk for each state variable. In principle, the risk premium approach is equivalent to the pricing kernel approach. But we show that seemingly benign risk premium specifications can result in pricing kernels that are inconsistent with economic theory and/or have undesirable economic and statistical properties. Specifically, these pricing kernels can be non-monotonic as a function of the variance; alternatively, their dependence on variance may be subject to identification problems. Moreover, the resulting time path of the pricing kernels can be excessively volatile, even when times are good, and the projection of marginal utility on stock market returns, which the literature refers to as the marginal pricing kernel, may be characterized by multiple inflection points.

Explicitly specifying the pricing kernel avoids these problems. We also advocate the use of a specific pricing kernel that corresponds to a parsimonious specification of the price of volatility risk. To illustrate the properties of well-behaved pricing kernels and to illustrate potential problems with the specification of the price of risk, we study a class of pricing kernels that extend the power kernels in Rubinstein (1976) and Brennan (1979) to be functions of the paths of latent variance $v(t)$ and the index level $S(t)$. These kernels are consistent with the conventional assumption of affine variance dynamics under the physical and risk-neutral measure in the square root stochastic volatility model (Heston, 1993), and they explicitly incorporate variance risk aversion. To study the relation between these pricing kernels and existing specifications of the price of risk, we first analyze the often-used specification of an equity (market) risk premium $\mu v(t)$ and a variance risk premium $\lambda v(t)$, which are directly proportional to the variance. We characterize a pricing kernel that is consistent with these risk premia, and derive the parameter restrictions consistent with the martingale conditions and absence of arbitrage. We contrast this risk premium with a more general specification with an equity risk premium $\mu_0 + \mu_1 v(t)$ and/or a variance risk premium $\lambda_0 + \lambda_1 v(t)$. We refer to this latter specification as the affine price of risk specification. We refer to the former specification, which is of course a special case, as the affine price of risk specification with zero intercepts.

Our empirical investigation illustrates how these specification problems affect estimation results for option pricing models with stochastic volatility (SV) and the properties of the implied pricing kernels. This empirical analysis relies on a rich option data set, using data for the January 1996 to June 2019 period. To highlight the implications of these modeling choices for the pricing kernel, it is of course critical to use a joint likelihood based on options and the underlying index returns.[1] While the potential problems with the specification of the price of risk apply to all SV models, any empirical illustration is of course model-dependent. For clarity, we conduct most of our analysis using evidence based on the simplest possible model, a one-factor stochastic volatility model, but we also present robustness results for several models with additional factors. To highlight where potential mis-specification in the model and the resulting pricing kernel might surface, we compare the various estimated pricing kernels with the properties of a pricing kernel that imposes the stochastic volatility model structure but does not put restrictions on the mapping between the risk-neutral and physical measures. We refer to this kernel as "unrestricted". It is directly defined by the ratio of independently estimated P-probabilities (estimated from returns) and Q-probabilities (estimated from options). The properties of this "unrestricted" kernel are a useful benchmark because mis-specified and over-parameterized kernels will presumably try to capture them.

Our first contribution to the literature is a characterization of the potential problems associated with the most general affine price of risk used in SV option pricing models. We start from the observation that the more general affine price of risk provides significantly better fit than the nested affine specification with zero intercepts. However, we find that this improved fit due to the intercepts in the affine specification comes at the cost of counterintuitive properties of the implied

[1] 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), Andersen, Fusari, and Todorov (2017), Bardgett, Gourier, and Leippold (2019), and Aït-Sahalia, Karaman, and Mancini (2020)

pricing kernel. Specifically, it is straightforward to show that for some parameter combinations, the resulting pricing kernel is non-monotonic as a function of the variance, which results in a mis-specified model. For other parameter combinations, the pricing kernel is monotonic in the variance, but estimating its dependence on variance may be subject to identification problems because identification relies on functional form.

Next we provide intuition for the empirical problems resulting from mis-specification by documenting the properties of the resulting pricing kernel from two perspectives. First, we inspect the time series of the pricing kernel that results from inserting the realized time paths of the relevant state variables. We find that the time path of the pricing kernels can be excessively volatile, even in good times, and consequently exhibits large positive outliers when times are good. Second, we study the properties of the projection of marginal utility on wealth, proxied by the stock market return. The literature refers to this as the marginal pricing kernel, and it takes a central place in this literature. It is by now well understood that this marginal kernel is not necessarily monotonic in wealth for several possible reasons. However, we find that the model-implied marginal kernel implied by the most general affine price of risk is characterized by multiple inflection points. This is inconsistent with the mechanics and intuition underlying the stochastic volatility model, which should result in at most one inflection point. Finally, we show that the properties of the implied kernel are due to the model's attempt to mimick the properties of the unrestricted kernel, which does not impose restrictions on the relation between the two measures. These findings stand in stark contrast to the stylized facts we document for the exponential-affine kernel, which results in a marginal kernel and a time series of the kernel that closely align with economic intuition. We then show that these models perform very differently for the purpose of matching (expected) option returns, indicating that these improved economic properties have important implications when the models are evaluated using criteria that are not explicitly taken into account when optimizing (in-sample) fit.

In summary, our first set of results concludes that the general affine specification allows for implausible economic assumptions and results in over-parameterized models that may give rise to identification problems and unrealistic state prices. Seemingly innocuous generalization of the price of risk specifications used in the literature correspond to different pricing kernels with radically different economic implications. Consequently, we recommend the use of the exponential-affine kernel, which corresponds to the affine price of risk specification with zero intercepts.

Our second set of findings imposes additional restrictions on the exponential-affine kernel to study its economic properties. Because this kernel includes separate components for volatility and index return risk, it allows us to examine the distinct origins of the equity and variance risk premiums. We find that these additional restrictions have a minimal effect on how the model fits options and the underlying returns, and we do not reject the null hypothesis that either variance or market return shocks alone account for the risk premiums in index returns and options.[2] In other words, we are unable to statistically identify the origins of these risk premiums because these risks largely span each other. Because the innovations to market returns and market variance are highly (negatively) correlated, when we restrict variance aversion to zero, the index level risk aversion parameter absorbs most of that effect, and vice versa. However, these restrictions result in large differences in the state prices embodied in the marginal pricing kernel and the realized time series path of the kernel. We show that these different economic properties once again result in very different predictions for (expected) option returns. Variance risk aversion is critical to match the patterns in expected option returns.

A final set of empirical results investigates other restrictions on the exponential-affine kernel to provide insights into two other related literatures. First, our approach clarifies the path-dependent nature of the plausible exponential-affine pricing kernel. Imposing the path-independence restriction from recovery theory (Ross, 2015) significantly impacts the time series of the pricing kernel. Second, because the kernel explicitly accounts for volatility risk, we can verify if volatility risk aversion results in a U-shaped marginal pricing kernel, as argued for example by Christoffersen, Heston, and Jacobs (2013), and Song and Xiu (2016). We find that whereas the marginal kernel is indeed convex in log aggregate wealth, which confirms the importance of volatility risk, in our

[2]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.

sample the estimate of the variance risk aversion parameter is too small to generate a U-shaped marginal pricing kernel, unless index level risk aversion is very small.

Our results are most closely related to a large and well-established literature that estimates reduced form dynamic option pricing models.[3] Most of these studies that use reduced form risk premia to define the relationship between physical and risk-neutral probabilities use the affine price of risk specification with zero intercepts, which is isomorphic to the exponential-affine kernel. A non-exhaustive list of studies that use alternative specifications is Chernov and Ghysels (2000) and Song and Xiu (2016), who use an affine price of equity and/or variance risk, and Eraker (2004), who specifies a constant return drift. Singleton (2006, p. 392) points out potential violations of no-arbitrage in SV option pricing models due to the general affine specification of the price of risk. Cheridito, Filipović, and Kimmel (2007) and Heston, Loewenstein, and Willard (2006) discuss necessary conditions for no-arbitrage in affine term structure models for this specification. This paper instead emphasizes the economic implications of this specification of the price of risk, which include pricing kernels that are nonmonotonic in volatility and potential identification problems. We empirically characterize the economic properties of the resulting pricing kernels and find that the time series of the pricing kernel is excessively volatilie in good times and that the marginal kernel has multiple inflection points.

More broadly, our paper is related to several other strands of literature. An important literature studies the properties of the pricing kernel by specifying its relation to aggregate consumption. This literature is too extensive to cite in full here.[4] A more closely related literature uses general equilibrium models to analyze how preferences and pricing kernels impact index option prices.[5]

[3]See for instance Andersen, Benzoni, and Lund (2002), Bates (1996, 2000, 2006, 2012, 2019), Bakshi, Cao, and Chen (1997), Chernov and Ghysels (2000), Duffie, Pan, and Singleton (2000), Pan (2002), Eraker, Johannes, and Polson (2003), Eraker (2004), Christoffersen, Heston, and Jacobs (2009), Christoffersen, Jacobs, and Mimouni (2010), Aït-Sahalia, Cacho-Diaz, and Laeven (2015), Andersen, Fusari, and Todorov (2015a,b, 2017), Hurn, Lindsay, and McClelland (2015), Bardgett, Gourier, and Leippold (2019), Du and Luo (2019), Aït-Sahalia, Karaman, and Mancini (2020), Aït-Sahalia, Li, and Li (2021), Li and Zinna (2018), and Gruber, Tebaldi, and Trojani (2021).

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

[5]See, for instance, Garcia, Luger, and Renault (2003), Drechsler (2013), Shaliastovich (2015), Eraker and Yang (2019), and Seo and Wachter (2019). Eraker and Shaliastovich (2008) study the specification of affine models. Liu,

The pricing kernels we study are instead extensions of the power utility of Rubinstein (1976) and provide a direct link with existing empirical implementations of (reduced-form) parametric dynamic option pricing models. They are straightforward to implement using option data, which allows us to explore the impact of stock index volatility on the pricing kernel.[6] The role of stochastic volatility in pricing kernels for these models has not been sufficiently emphasized in the literature, which has proceeded by specifying prices of risk instead, with some exceptions. The stochastic volatility setup in Bates (2000) relies on logarithmic utility. Chabi-Yo (2012), Christoffersen, Heston, and Jacobs (2013), and Song and Xiu (2016) discuss volatility-dependent kernels. Liu, Brennan, and Xia (2006) specify and estimate pricing kernels with multiple state variables.

Our results also have important implications for the option pricing literature which studies the properties of the marginal pricing kernel using index returns and index option prices (Jackwerth, 2000). Much of this literature is non-parametric, building on the insights of Breeden and Litzenberger (1978).[7] A substantial part of this literature addresses the puzzling non-monotonic (U-shaped) estimates of the pricing kernel and provides potential explanations.[8] Christoffersen, Heston, and Jacobs (2013), Song and Xiu (2016), and Sichert (2023) emphasize the role of volatility risk in explaining U-shaped pricing kernels. The time series path of the pricing kernel has received less attention. A notable early exception is Chernov (2003), who reverse engineers the pricing kernel based on options on various securities and studies the time path of the realized pricing kernel to learn about state variables and the relation between the pricing kernel and economic conditions. Ghosh, Julliard, and Taylor (2017) also explore the relation between the pricing kernel and business cycle fluctuations, but do not use options to estimate the kernel.

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.

[6] While our approach is partial equilibrium, it is nevertheless useful to point out pitfalls with existing empirical applications, and it can be used as a building block for future general equilibrium approaches.

[7] See for instance Aït-Sahalia and Lo (1998), Aït-Sahalia and Lo (2000), Jackwerth and Rubinstein (1996), Jackwerth (2000) and Rosenberg and Engle (2002).

[8] 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 marginal pricing kernel is monotonic.

Finally, our findings are related to a rich parametric and nonparametric literature that uses index option and return data to characterize the moments of index returns and the corresponding risk premia (Duan and Zhang, 2014; Martin, 2017; Martin and Wagner, 2019; Chabi-Yo and Loudis, 2020, 2023). Our results confirm existing findings that it is difficult to precisely estimate time-variation in the equity and variance risk premiums (Merton, 1980). We emphasize that it is challenging to identify the source of these risk premiums due to the correlation between return and variance shocks. This is related to Jackwerth and Vilkov (2019), who emphasize the time-variation in this correlation and use SPX and VIX options to improve identification. Our estimates of the pricing kernel can be used to interpret empirical findings on risk premiums such as Dew-Becker and Giglio (2022), who find that the variance risk premium has declined recently and that it can be entirely attributed to a negative stock beta.

The paper proceeds as follows. Section 2 introduces the data and the Heston (1993) stochastic volatility model. Section 3 specifies the class of pricing kernels that connect the risk-neutral and physical dynamics. Section 4 discusses our estimation approach based on returns and options data. Sections 5 and 6 present and discuss the empirical results. Section 7 presents the results of robustness exercises and extensions, and Section 8 concludes.

2 Data and Model

We discuss the data and briefly summarize the stylized affine Heston (1993) stochastic volatility model.

2.1 Data

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

1. Discard options with implied volatility smaller than 5% or greater than 150%.

2. Discard options with volume or open interest less than ten contracts.
3. Discard options with mid price less than $0.50 or bid price less than $0.375.
4. Discard options with data errors – where bid price exceeds offer price, or a negative price is implied through put-call parity.
5. Discard options with moneyness less than 0.75 or greater than 1.25.

Following 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. We follow Dufays, Jacobs, Liu, and Rombouts (2023) who construct an option sample that is balanced over time by selecting representative options that span the implied volatility surface. We use six moneyness bins and six maturity bins and select the most liquid option contract in each moneyness-maturity bin.[9] This procedure results in a dataset with 33,535 option contracts. Table 1 presents descriptive statistics. Note that each bin does not contain exactly the same number of option contracts, because on a given day options may not be available for every moneyness or maturity bin.

We obtain S&P 500 index returns for 1996-2019 from CRSP and VIX data for 1996-2019 from the Federal Reserve Bank of St. Louis Economic Database. The time series for the risk-free rate is proxied by the one-month Treasury Bill rate obtained from CRSP. Following existing work, options are valued using a maturity-specific risk-free rate. We apply a cubic spline interpolation to the data obtained from OptionMetrics.

2.2 The Model

Our baseline analysis focuses on the simplest possible stochastic volatility model with a single diffusive volatility factor. 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, and we study some of these extensions in Sections 7.3 and 7.4. We deliberately

[9] In the moneyness (defined as the underlying price divided by the exercise price) dimension, we define bins with moneyness levels spaced every 0.04 between 0.94 and 1.10, and assign two additional bins for options with moneyness below 0.94 or above 1.10. In the maturity dimension, we define bins with maturities of 30 days or less, 31-60 days, 61-90 days, 91-120 days, 121-180 days, and more than 180 days.

focus on the simplest possible model in our baseline analysis because it suffices to illustrate our main argument and we want to avoid comparisons between models and factors.

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

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

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

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

where $\mu(v(t))$ denotes the equity premium as a function of $v(t)$, and dz_1 and dz_2 are Wiener processes under the physical measure. Note that 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 A Taxonomy of Pricing Kernels

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

3.1 Affine Prices of Risk

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

Heston (1993) restricts the variance risk premium to be proportional to the variance, i.e., equal to $\lambda v(t)$. Because Heston (1993) does not perform (joint) estimation of the model parameters, he does not explicitly address the structure of the equity premium. However, several seminal estimation exercises (see, e.g., Bates, 2000; Pan, 2002; Aït-Sahalia and Kimmel, 2007) impose a similar proportional structure on the equity premium, i.e., equal to $\mu v(t)$. We therefore start our analysis with this specification of the equity and variance risk premium.[10] Note that in this case, the relation between the risk-neutral and physical parameters in equations (1) and (2) are given by $\kappa = \kappa^* - \lambda$ and $\theta = \kappa^* \theta^* / \kappa$.

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

[10] The literature sometimes refers to this specification, with equity and variance risk premiums that are proportional to the stochastic variance, as *completely affine* risk premiums (Singleton, 2006).

risk premiums with zero intercepts.

$$M(t) = M(0)\exp\left(-rt - \frac{1}{2}\int_0^t Var\left[\frac{dM(s)}{M(s)}\right] - \int_0^t \pi_1\sqrt{v(s)}dz_1(s) - \int_0^t \pi_2\sqrt{v(s)}dz_2(s)\right), \quad (3)$$

where $Var\left[\frac{dM(s)}{M(s)}\right] = \left(\pi_1^2 + \pi_2^2 + 2\rho\pi_1\pi_2\right)v(s)ds$, and π_1 and π_2 are price of risk parameters, which are related to the risk premium parameters as follows:

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

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

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

$$\begin{aligned} M(t) = M(0)\exp\Bigg(-rt - \frac{1}{2}\int_0^t Var\left[\frac{dM(s)}{M(s)}\right] &- \int_0^t \left(\frac{\pi_{1,0}}{\sqrt{v(s)}} + \pi_{1,1}\sqrt{v(s)}\right)dz_1(s) \\ &- \int_0^t \left(\frac{\pi_{2,0}}{\sqrt{v(s)}} + \pi_{2,1}\sqrt{v(s)}\right)dz_2(s)\Bigg), \quad (5) \end{aligned}$$

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

$$\mu_0 = \pi_{1,0} + \rho\pi_{2,0}, \quad \mu_1 = \pi_{1,1} + \rho\pi_{2,1}, \quad \lambda_0 = \rho\sigma\pi_{1,0} + \sigma\pi_{2,0}, \quad \lambda_1 = \rho\sigma\pi_{1,1} + \sigma\pi_{2,1}, \quad (6)$$

and

$$Var\left[\frac{dM(s)}{M(s)}\right] = \Bigg[\left(\pi_{1,0}^2 + \pi_{2,0}^2 + 2\rho\pi_{1,0}\pi_{2,0}\right)\frac{1}{v(s)} + \left(\pi_{1,1}^2 + \pi_{2,1}^2 + 2\rho\pi_{1,1}\pi_{2,1}\right)v(s)$$
$$+\ 2\left(\pi_{1,0}\pi_{1,1} + \pi_{2,0}\pi_{2,1} + \rho\pi_{1,0}\pi_{2,1} + \rho\pi_{2,0}\pi_{1,1}\right)\Bigg]ds. \tag{7}$$

One critical difference between the pricing kernels in equations (3) and (5) is that (5) may allow for arbitrage when the variance reaches zero (Singleton, 2006). This can be addressed by imposing the Feller (1951) condition(s). We discuss this in more detail below. In the absence of this condition, our prior is that the specification in equation (5) will lead to problems at times when the variance is low.

3.2 A General Pricing Kernel

The pricing kernels in equations (3) and (5) contain complicated, opaque stochastic integrals that lack a clear economic interpretation. The purpose of this section is to simplify and interpret these kernels using (generalizations of) economically motivated pricing kernels from the finance literature. These include the seminal power utility kernel of Rubinstein (1976) and Brennan (1979) and the transition-independent kernel of Ross (2015).

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

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

$$M(t) = M(0) \left(\frac{S(t)}{S(0)} \right)^{-\gamma} \left(\frac{v(t)}{v(0)} \right)^{\alpha} \exp \left(\beta t + \int_0^t \eta(v(s))ds + \xi(v(t) - v(0)) \right). \tag{8}$$

Note that this is not the most general possible kernel with state variables $S(t)$ and $v(t)$. We start our analysis with this kernel because it provides insight into the affine equity and variance risk premium specification, or equivalently into the pricing kernel in equation (5). Specifically, Online Appendix B shows that the mapping between the preference parameters in equation (8) and the reduced-form parameters characterizing the affine risk premiums $\mu_0 + \mu_1 v(t)$ and $\lambda_0 + \lambda_1 v(t)$ is as follows:

$$\mu_0 = -\alpha\sigma\rho, \quad \mu_1 = \gamma - \rho\sigma\xi, \quad \lambda_0 = -\alpha\sigma^2, \quad \lambda_1 = \gamma\rho\sigma - \xi\sigma^2. \tag{9}$$

From an economic perspective, we note two important features of the kernel in equation (8). First, while it contains an integral that depends on the history of the variance, it does not depend on the history of the stock price. This represents a substantial simplification from the general stochastic integrals in (3) and (5). Second, the kernel can be seen as extending the power utility of Rubinstein (1976) and Brennan (1979) to include the pricing of variance risk. However, the kernel contains two parameters (α and ξ) that govern the pricing of variance, and this is a potential source of problems. Any economically plausible pricing kernel should produce a monotonic variance premium. Because α affects the intercepts of the equity and variance risk premiums and ξ affects the slopes (see equation (9)), the pricing kernel can be a nonmonotonic function of variance, which directly affects the cross-section of option prices. Similar nonmonotonicities occur when the equity premium has a nonzero intercept but the variance risk premium does not, or vice versa.

Finally, a brief discussion of the parameter count in these kernels is in order. The most general affine risk premium specification contains four parameters $(\mu_0, \mu_1, \lambda_0, \lambda_1)$. The kernel (8) characterizes these four parameters in terms of the parameters in the physical dynamic and the three preference parameters γ, α and ξ. We therefore refer to (8) as a three-parameter kernel.

The kernel (8) contains two additional preference parameters, β and η. But using the martingale restriction on the kernel itself, Online Appendix B shows that the time-preference parameter β and the path-dependence function $\eta(v(t))$ are given by:

$$\begin{aligned} \beta &= -(1-\gamma)r + \gamma\mu_0 - \xi\kappa\theta + \alpha\kappa + \gamma\alpha\sigma\rho - \xi\alpha\sigma^2, && (10) \\ \eta(v(t)) &= \left[\left(\mu_1 - \frac{1}{2}\right)\gamma + \xi\kappa - \frac{1}{2}\left(\gamma^2 - 2\gamma\xi\sigma\rho + \xi^2\sigma^2\right)\right] v(t) - \left[\kappa\theta - \frac{1}{2}\sigma^2 + \frac{1}{2}\alpha\sigma^2\right]\frac{\alpha}{v(t)}. && (11) \end{aligned}$$

Restrictions (10)-(11) are required to match the risk-free rate, which is an affine function of the variance. Restriction (10) constrains the intercept in this affine function by constraining the β parameter, and restriction (11) constrains the slope. Equation (11) stipulates that given the physical parameters and the three preference parameters γ, α and ξ, there is a (time-varying) path-dependence parameter $\eta(v(t))$ that imposes no-arbitrage. The implications of (not) imposing the restrictions (10) and (11) are outside the scope of our study.[11] We merely point out that for the purpose of the mapping between the kernel (8) and the affine risk premiums, the kernel effectively uses three preference parameters to characterize the four parameters μ_0, μ_1, λ_0, and λ_1.

3.3 A Monotonic Kernel

To prevent nonmonotonicities, we restrict α in equation (8) to be zero, which gives the following simpler two-parameter specification:

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

This specification is easier to interpret economically, because one parameter γ determines equity risk-aversion, and another parameter ξ governs variance-aversion. Moreover, Online Appendix C shows that given the risk-neutral dynamics (1), the physical dynamics (2), no-arbitrage, and affine equity and variance risk premiums with zero intercepts $\mu v(t)$ and $\lambda v(t)$, the two pricing kernel

[11]The restrictions (10) and (11) should in principle be imposed in empirical work *in addition* to the affine risk premiums $\mu_0 + \mu_1 v(t)$ and $\lambda_0 + \lambda_1 v(t)$, but usually these restrictions are ignored. This amounts to stating that one has no prior on the magnitude of the path-dependence parameter, and that one is willing to accept the risk-free rate implied by the values of $\mu_0, \mu_1, \lambda_0, \lambda_1$, regardless of the resulting value of the risk-free rate.

parameters γ and ξ determine the equity premium μ and variance premium λ parameters, as follows:

$$\text{Equity risk premium:} \quad \mu v(t) = (\gamma - \xi\sigma\rho)v(t), \tag{13}$$

$$\text{Variance risk premium:} \quad \lambda v(t) = (\rho\sigma\gamma - \sigma^2\xi)v(t) = (\rho\sigma\mu - (1-\rho^2)\sigma^2\xi)v(t). \tag{14}$$

Equations (13) and (14) indicate that due to the correlation ρ between stock returns and variance, both the equity premium μ and variance premium λ depend on both the preference parameters γ and ξ. The variance premium λ may be negative because the pricing kernel parameter ξ is positive, or because variance is negatively correlated with equity returns ($\rho < 0$) and the equity premium is positive (i.e., variance has a negative beta). We can test these hypotheses separately, by restricting ξ or λ to be zero. Similarly, we can test the hypothesis that the equity risk-aversion γ is zero separately from the hypothesis that the equity premium μ itself is zero. Finally, note that for this pricing kernel the no-arbitrage restrictions (10) and (11) reduce to:

$$\beta = -(1-\gamma)r - \xi\kappa\theta, \tag{15}$$

$$\eta = \left(\mu - \frac{1}{2}\right)\gamma + \xi\kappa - \frac{1}{2}\left(\gamma^2 - 2\gamma\xi\sigma\rho + \xi^2\sigma^2\right). \tag{16}$$

Henceforth we refer to the kernel in equation (12) as the exponential-affine pricing kernel. Next, we discuss two important special cases of this kernel.

3.4 The Power Utility Pricing Kernel

A first economically important special case emanates from setting ξ equal to zero in equation (12). This imposes absence of variance preference, and gives:

$$M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^{-\gamma} \exp\left(\beta t + \int_0^t \eta v(s)ds\right). \tag{17}$$

The local fluctuations in the pricing kernel (17) now result exclusively from changes in the spot price $S(t)$, and there is no additional premium due to variance aversion. The martingale condition once again implies that the expected growth of the pricing kernel must equal the opposite of the interest rate. This restricts β and η, effectively leaving only a single index level preference parameter γ to price risk.

For convenience, we will refer to the kernel in (17) as the power utility pricing kernel. However, the kernel (17) differs from the power utility pricing kernels of Rubinstein (1976) and Brennan (1979) because it contains a path-dependent component $\eta \int_0^t v(s)ds$, even in the absence of variance preference.

3.5 A Path Independent Pricing Kernel

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

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

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

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

This special case effectively sets $\eta = 0$ in equation (16), which is a joint restriction on the equity risk aversion parameter γ and variance risk aversion ξ, and therefore also reduces the pricing kernel to a single parameter. We test this restriction in our empirical application.

4 Return-Based and Option-Based Estimation

We estimate the stylized affine Heston (1993) under the physical measure, exclusively based on returns, and under the risk-neutral measure, exclusively based on options. Then we compare the resulting estimates and we explain how to perform joint estimation.

4.1 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.[12] We use the fact that the stochastic variance $v(t)$ can be represented as a linear function of $\text{VIX}^2(t)$. This directly follows from the model specification: When $v(t)$ follows a CIR process, $\text{VIX}^2(t)$ is a linear function of $v(t)$. Specifically, the model-implied $\text{VIX}^2(t)$ is given by:

$$\begin{aligned}\text{VIX}^2(t) &= \frac{1}{\Delta_{1m}} E_t^* \left[\int_t^{t+\Delta_{1m}} v(u) du \right] \\ &= \theta^* + \frac{e^{-\kappa^* \Delta_{1m}} - 1}{-\kappa^* \Delta_{1m}} (v(t) - \theta^*),\end{aligned} \tag{18}$$

where $\Delta_{1m} \approx 30/365$. Rearranging equation (18) yields

$$v(t) = \frac{\text{VIX}^2(t) - \theta^*(1-w)}{w}, \tag{19}$$

where $w = (1 - \exp(-\kappa^* \Delta_{1m}))/(\kappa^* \Delta_{1m})$. In implementation, we can add a measurement error because equations (18) and (19) use the model-implied $\text{VIX}^2(t)$. Equation (18) in conjunction

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

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 (19) and omit the measurement error. Specifically, we assume:

$$v(t) = \eta_0 + \eta_1 \mathrm{VIX}^2(t). \tag{20}$$

We then use equation (20) in the valuation formula for all options in the sample. As a result, options are a function not only of the stochastic $v(t)$, but also of the observable VIX. This implementation follows Aït-Sahalia and Kimmel (2007), who use it in a sample which contains a single short-maturity at-the-money option at each time t. We next discuss the details of this estimation approach when using returns and when using options. Our use of the VIX as a proxy for the stochastic variance has implications for both estimation exercises.

4.2 Return-Based and Option-Based Estimation

4.2.1 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 Ito's lemma and the Euler discretization to equation (2), which results in:

$$
\begin{aligned}
\log R(t+\Delta) &= \left[r+\mu(v(t))-\frac{1}{2}v(t)\right]\Delta+\epsilon_R(t+\Delta), \\
v(t+\Delta)-v(t) &= \kappa(\theta-v(t))\Delta+\epsilon_v(t+\Delta),
\end{aligned}
\tag{21}
$$

where $R(t+\Delta)=S(t+\Delta)/S(t)$ represents the gross return and $\Delta=1/252$.[13] The errors $\epsilon(t+\Delta)=(\epsilon_R(t+\Delta),\ \epsilon_v(t+\Delta))'$ follow a joint normal distribution, and their mean and variance-covariance matrix are given by

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

The joint log-likelihood function is given by:

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

where $f(\log R(t+\Delta),v(t+\Delta)|v(t))$ is the conditional density of the discretized $\log R(t+\Delta)$ and $v(t+\Delta)$, $\mathcal{J}(t+\Delta)$ is the Jacobian between $\mathrm{VIX}^2(t+\Delta)$ and $v(t+\Delta)$, which is given by η_1 from equation (20), 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:

$$
\max_{\Theta}\log L^R. \tag{22}
$$

[13]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.

4.2.2 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

$$C(S(t), v(t), t) = \frac{e^{-k\chi}}{\pi} \int_0^{\infty} Re\left[e^{-iuk}\psi(u)\right] du, \tag{23}$$

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 $e^{k\chi}$ for $\chi > 0$. We found that $\chi = 4$ works well. The function $\psi(u)$ is calculated as follows:

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

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).[14] The price of a put option with the same strike price and maturity can be obtained through put-call parity. Note that the option pricing formula in equation (23) 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.

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

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

the characteristic function solution is given by

$$f_{\tau}^{CH}(\phi|S(t), v(t)) = e^{C + Dv(t) + i\phi \log S(t)}, \tag{24}$$

where

$$\begin{aligned} &C = r\phi i\tau + \tfrac{a}{\sigma^2}\left\{(b - \rho\sigma\phi i + d)\tau - 2\log\left[\tfrac{1-ge^{d\tau}}{1-g}\right]\right\}, \quad D = \tfrac{b-\rho\sigma\phi i+d}{\sigma^2}\left[\tfrac{1-e^{d\tau}}{1-ge^{d\tau}}\right], \\ &g = \tfrac{b-\rho\sigma\phi i+d}{b-\rho\sigma\phi i-d}, \quad \text{and} \quad d = \sqrt{(\rho\sigma\phi i - b)^2 - \sigma^2(2u\phi i - \phi^2)}. \end{aligned} \tag{25}$$

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

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

where ν_i^{Mkt} is the Black-Scholes vega of option i.[15]

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.* $\epsilon_{o,i} \sim \mathcal{N}(0, s_o^2)$, 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:

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

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

$$\log L^O = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log s_o^2 - \frac{1}{2s_o^2}\sum_{i=1}^{N}\epsilon_{o,i}^2. \tag{26}$$

4.2.3 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 have signs that are consistent with economic theory. Consistent

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

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 1996 to June 2019 sample period, close to the sample average of 8.22%. 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 18.8% (32.9%).

4.3 Joint Maximum Likelihood Estimation with No-Arbitrage Restrictions

Rather than separately estimating the physical and risk-neutral dynamics as in Table 2, we now discuss how to 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 (12) for example, or we directly impose a structure on the parameters that characterize the risk premia such as μ and λ in the affine specification without intercepts.

For kernels nested in the exponential-affine specification in equation (12), 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 (13) and (14). Likewise, for the more general kernel in equation (8), 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 (9). We also investigate specifications that restrict the risk premium parameters instead. For the affine price of risk specification, the parameter vector is $\Theta^{PR} = \{\mu_0,\ \mu_1,\ \lambda_0,\ \lambda_1,\ \kappa,\ \theta,\ \sigma,\ \rho,\ \eta_0,\ \eta_1\}$. Regardless of the parameterization, estimation is based on the weighted sum of the return and option log-likelihoods, following Christoffersen, Jacobs, and Ornthanalai (2012). That is, we solve the following optimization problem:

$$\max_{\Theta} \frac{N_R + N_O}{2} \left(\frac{\log L^R}{N_R} + \frac{\log L^O}{N_O} \right),$$

where N_R and N_O represent the number of returns and options data points, and Θ can refer to Θ^{PK1}, Θ^{PK2}, or Θ^{PR}.

Table 3 presents estimation results for the exponential-affine kernel (12) in column (1), the affine specification of the price of risk in column (2), the more general kernel in equation (8) in column (5), and the special case of this kernel with $\xi = 0$ in column (6). Table 4 reports on the exponential-affine kernel (12) and six restricted cases. In columns (7) and (8), we have two restricted estimation exercises based on risk premium restrictions: $\mu = 0$ and $\lambda = 0$ respectively. Next, we have several cases based on preference parameter restrictions. In column (9) we impose $\gamma = 0$, and in column (10) we impose $\xi = 0$. The restriction $\xi = 0$ corresponds to the power utility pricing kernel in equation (17). We also test the joint restrictions $\gamma = \xi = 0$ in column (11).[16] Finally, we impose $\eta = 0$ in column (12). This case provides insight into the literature on recovery theory (see, e.g., Ross, 2015; Borovička, Hansen, and Scheinkman, 2016; Qin and Linetsky, 2016). Online Appendix D summarizes the implementation of the restrictions used in these tables.

5 Affine Prices of Risk: Empirical Results

We estimate the stochastic volatility model subject either to the affine risk premium specification $\mu_0 + \mu_1 v(t)$ and $\lambda_0 + \lambda_1 v(t)$ or the nested specification without intercepts $\mu v(t)$ and $\lambda v(t)$. We

[16] From equations (13) and (14), the joint restriction $\mu = \lambda = 0$ is equivalent to the joint restriction $\gamma = \xi = 0$.

compare the resulting model estimates and option fit with results that impose the pricing kernels in equations (8) or (12). Our results are based on the joint likelihood composed of returns and options. We discuss how to estimate the marginal pricing kernel that plots the pricing kernel as a function of log index returns and we present estimates of these marginal pricing kernels for different specifications of the price of risk and the pricing kernel. We also study the time series paths of the pricing kernels for these different specifications. Finally, we show how well the various pricing kernels and price-of-risk specifications fit expected option returns.

5.1 Parameter Estimates and Option Fit

Table 3 reports the results of the joint MLE estimation. We report robust standard errors for all parameters, except for those that are implied by other parameter estimates (see Online Appendix D). Column (1) presents the estimates for the MLE estimation of the exponential-affine pricing kernel in equation (12).[17] We can compare the risk-neutral estimates in column (1) with the option-based estimates in Table 2. The risk-neutral mean-reversion $\kappa^* = \kappa + \lambda$ from joint estimation in Table 3 is 1.3976, somewhat higher than the 0.8401 estimate in Table 2. The risk-neutral long-run variance $\theta^* = \kappa\theta/\kappa^*$ in Table 3 is 0.0747, somewhat lower than the 0.1085 estimate in Table 2. This finding is not surprising, because the option-implied variance typically exceeds the average variance implied by returns.

We can also compare the physical parameters in column (1) of Table 3 with the return-based estimates in Table 2. The physical mean reversion κ in Table 3 is 2.816, substantially smaller than the 3.837 estimate in Table 2. The physical long-run variance θ in Table 3 is 0.037, similar to the 0.035 estimate in Table 2. The most important difference is that the return-based estimates of ρ and σ in Table 2 are smaller (in absolute value) compared to Table 3. Recall that in joint estimation 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.

[17] Recall that this kernel is isomorphic to the affine price of risk specification without intercepts, i.e. with equity premium $\mu v(t)$ and variance risk premium $\lambda v(t)$.

The estimate of the index level preference parameter ("risk aversion") γ in the exponential-affine kernel in column (1) is 1.309 and the estimate of the variance preference parameter ξ is 1.872. These signs are consistent with economic intuition. The implied 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 consistent with economic priors about variance aversion. A negative λ implies $\kappa > \kappa^*$ and $\theta < \theta^*$. It also implies that the ordinal relations between the physical and risk-neutral parameters are the same as in Table 2, which reports estimates from returns only and options only.

Column (5) of Table 3 presents results for the more general kernel in equation (8). The risk-neutral mean reversion is similar to that for the exponential-affine kernel in column (1), while the physical mean-reversion is higher. The other parameters and option RMSE are very similar. Column (2) in Table 3 reports on the affine specification of the price of risk with an equity premium $\mu_0 + \mu_1 v(t)$ and a variance risk premium $\lambda_0 + \lambda_1 v(t)$. The risk-neutral parameter estimates θ^*, σ, and ρ are again very similar to the estimates in column (1), but 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 most important conclusion is that the joint log-likelihood in column (2) exceeds the log-likelihood in column (1), as well as the one in column (5). The likelihood ratio tests in Table 5 show that these differences are statistically significant at the 1% level. It seems that including an (unrestricted) intercept in the risk premium specification provides a better fit to the data. This improvement in fit is due to improved fit of the return data. Note that in column (5), the more general kernel in equation (8) also contains an intercept in the risk premiums, but this intercept is subject to the restrictions in equation (9), and this specification does not result in such a big improvement.

The affine risk premium specification in column (2) clearly provides superior model fit.[18] Panel B of Table 5 presents descriptive statistics for the equity and variance risk premia. The annualized

[18]This finding is related to findings in the bond pricing literature, where the affine specification with zero intercepts is outperformed by more general specifications.

equity premium for the exponential-affine kernel in column (1) is 8.28%, close to the sample estimate of 8.22%, and the equity risk premiums in columns (2) and (5) are similar. The variance risk premium for the exponential-affine pricing kernel in column (1) is equal to -0.0524, which amounts to -22.89% in annual standard deviation terms.[19] The results on the affine price of risk specification in column (2) and the more general kernel in column (5) again yield similar estimates. The third row in Panel B of Table 5 reports the time-series averages of the daily Sharpe ratios for the various kernel specifications based on the model-implied conditional means and standard deviations. Figure A.1 plots the time series of these Sharpe ratios. The exponential-affine pricing kernel in column (1) results in an annualized Sharpe ratio of 0.386, compared to the sample average of 0.431. The more general kernel in column (5) yields a similar Sharpe ratio, but the Sharpe ratio for the affine price of risk specification in column (2) is higher. Panel B of Figure A.1 shows that this Sharpe ratio also displays very limited variation over time, unlike the Sharpe ratios in Panels A and C.

We conclude that the affine risk premium produces mixed results. It produces average risk premiums that are similar to the exponential-affine kernel in column (1), but the signs of the (implied) risk premium parameters in Table 3 imply that the signs of the equity and variance risk premiums may not be economically reasonable on each day t, and the time series of the Sharpe ratio displays limited variation. We now further explore the economic implications of these different kernels and price-of-risk specifications.

5.2 Estimating Marginal Pricing Kernels

Following the literature, we refer to the univariate pricing kernel in the aggregate wealth dimension, i.e., as a function of the log index return, as the marginal distribution of the pricing kernel. The traditional view of this marginal pricing kernel is guided by the power utility pricing kernel in equation (17), which predicts a linear downward sloping log pricing kernel as a function of the log index return. It is by now well understood that variance-dependent pricing kernels can lead to

[19] For the 1996 to 2019 period, the sample mean of the estimated volatility is 18.77%. The implied variance risk premium is therefore slightly larger than the index variance itself. This finding is similar to Bollerslev, Tauchen, and Zhou (2009).

nonmonotonicities in the pricing kernel (Christoffersen, Heston, and Jacobs, 2013; Song and Xiu, 2016). Figure 1 plots the estimated exponential-affine kernel in equation (12) as a function of the log stock return $\log R(t)$ and the daily change in variance $v(t) - v(t-1)$, based on the MLE estimates in column (1) of Table 3. The two bottom pictures depict the univariate relations. Figure 1 shows that the log return and the change in variance are negatively and positively correlated respectively with the log pricing kernel, due to the positive estimates of the index return preference γ and 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 positive 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.

We now discuss how to estimate the marginal pricing kernel. In general, it is well-known that it is not straightforward to reliably estimate pricing kernels. We need to characterize both the physical and risk-neutral distribution. Characterizing the risk-neutral distribution is relatively straightforward because of the abundance of option data on a given day, but characterizing the physical distribution is more challenging. The literature contains several non-parametric and parametric approaches. We proceed parametrically based on the parameter estimates in Tables 2, 3, and 4. Given the parametric stock return dynamics in equations (1) and (2), we can generate the probability density of returns under both the physical and risk-neutral measures. For simplicity, we let the initial stock price $S(0) = 1$ or equivalently $\log S(0) = 0$. Following Heston (1993), the cumulative distribution function of log returns can be calculated as follows:

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

where the characteristic function f_τ^{CH} follows equation (24). 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:

$$\begin{aligned} Pr(\log R(\tau) = x) &= \frac{dPr(\log R(\tau) \leq x)}{dx} \\ &= -\frac{1}{\pi} \int_0^{\infty} Re\left[-e^{-i\phi x + C + Dv}\right] d\phi, \end{aligned} \tag{27}$$

where C and D are given in equation (25).

We calculate the numerical integral in equation (27) 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 $\bar{r} = 0.0210$ (2.10%), the sample mean for our 1996-2019 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:

$$\log M(\tau) - \log M(0) = -\bar{r}\tau + \log Q(\log R(\tau)|v(0) = \theta) - \log P(\log R(\tau)|v(0) = \theta). \tag{28}$$

By changing the value of $\log R(\tau)$, we can therefore generate the log pricing kernel as a function of the log index return.

5.3 The Unrestricted Marginal Pricing Kernel

We study the $\tau-$maturity log pricing kernel in equation (28) for various physical and risk-neutral parameter vectors, using the estimates resulting from various restrictions on the pricing kernel in Tables 3 and 4. We first consider the marginal pricing kernel implied by the estimates in Table 2, i.e., using physical estimates obtained from returns and risk-neutral estimates obtained (separately) from options. These estimates do not impose restrictions on the pricing kernel, but they do of course impose a parametric structure on the probability distributions.

Panel A of Figure 2 presents results for the one-month pricing kernel. For convenience, we express the x-axis in standardized log returns (the log return in deviation from its mean and divided by its standard deviation).[20] The pricing kernel has a W-shape. This is consistent with some existing findings (Cuesdeanu and Jackwerth, 2018), although a U-shape is more common. Aït-Sahalia and Lo (2000) report an S-shape for the unconditional pricing kernel. See also Sichert (2023) on S-shaped and U-shaped pricing kernels.[21] Panel A of Figure A.2 reports on the 6-month pricing kernel. Results are qualitatively similar across maturities.

5.4 Marginal Pricing Kernels from Joint Estimation

Panels B-D of Figure 2 investigate how the shape of the pricing kernel as a function of the log index return depends on the specification of the price of risk and/or the parameterization of the pricing kernel. A central question is if the implied marginal pricing kernels can capture the patterns in Panel A. Panel B is based on the estimates in column (1) of Table 3, i.e., the exponential-affine pricing kernel. Recall that the affine price of risk specification with zero intercepts $\mu v(t)$ and $\lambda v(t)$ is isomorphic to the exponentially-affine pricing kernel. While the kernel in Panel B is downward sloping as a function of the log index return, it is not linear. Because the estimate of ξ is positive (and ρ is negative) in column (1) of Table 3, the function is convex.

Panel C reports on the affine price of risk estimates in column (2) of Table 3, which imposes an equity premium of $\mu_0 + \mu_1 v(t)$ and a variance risk premium of $\lambda_0 + \lambda_1 v(t)$. Panel D reports on the general pricing kernel in equation (8), corresponding to the estimates in column (5) of Table 3. The marginal kernel in Panel D is similar to the one in Panel B. The non-monotonicity in variance in the more general kernel (8) is theoretically problematic because it can lead to option prices that are non-monotonic in the variance. While this feature does not seem to affect the shape of

[20] Since we already have the physical probability density, we can calculate the expectation and variance of the log return as:

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

[21] The estimation of the pricing kernel is much more precise for low returns. For high returns, confidence intervals are typically much larger due to more limited option information. We do not report confidence intervals in Figure 2, and moreover this issue is also impacted by imposing a parametric structure.

the marginal kernel much, when the non-monotonicity is captured through the more general risk premiums $\mu_0 + \mu_1 v(t)$ and $\lambda_0 + \lambda_1 v(t)$, as in Panel C, the resulting marginal kernel is very different. The additional flexibility of the resulting kernel helps capture the patterns in the data, but this results in a marginal kernel with features similar to the unrestricted pricing kernel in Panel A, which does not impose any restrictions across measures.

Finally, Figure A.3 plots the same four marginal kernels conditional on the level of the variance. The left panels plot the kernels conditional on a low variance level, while the right panels are conditonal on high variance. The figure clearly confirms our prior that the problems with the general affine price of risk specification surface at low volatility levels. This shortcoming of the specification is mechanical and due to the functional form of the price of risk.

5.5 Pricing Kernel Dynamics

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. We insert 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. Generating this time path is straightforward when the pricing kernel is explicitly specified.[22] See Chernov (2003) for a related exercise.

Figure 3 presents the results. In general, the time series of the pricing kernels illustrate the differences between the different specifications even more dramatically than the marginal kernels in Figure 2. First compare the time series of the unrestricted kernel in Panel A of Figure 3 with that of the exponential-affine kernel in Panel B. Both time series are most variable during the financial crisis. However, the differences between the financial crisis and the rest of the sample are much less pronounced in Panel A. More importantly, the outliers in Panel A are of an entirely different order of magnitude compared to the exponential-affine pricing kernel in Panel B.[23] Explaining the return

[22]For the affine price of risk specification and when the physical and risk-neutral parameters are independently estimated, this computation is more complex. Online Appendix E discusses these cases.

[23]A value of -0.02 for the log pricing kernel on the vertical axis corresponds to a daily discount factor of 0.98. In

and option data without linking the underlying P and Q estimates results in an implied pricing kernel in Panel A that is often very volatile in calm times.

The marginal pricing kernel for the affine price of risk in Panel C of Figure 2 captures some of the patterns of the unrestricted kernel in Panel A. Consistent with this, the time series of the pricing kernels in Panels A and C of Figure 3 both contain many more outliers compared to the exponential-affine kernel in Panel B. Similar to the exponential-affine kernel in Panel B, some of the largest outliers in Panel C occur during the financial crisis. However, Panel C contains many more outliers, for instance during 2004-2007. As a result, the standard deviation of the path of the daily pricing kernel in Panel C is 0.0601, compared to 0.0345 in Panel B. This results in an annualized ratio of the standard deviation to the mean of the pricing kernel of 0.9554 in Panel C, as compared to 0.5483 in Panel B. Using the insights of Hansen and Jagannathan (1991), these ratios can be compared to the sample Sharpe ratio, which is 0.431.[24] Finally, the time series in Panels B and D are rather similar. This also confirms the conclusions from the marginal pricing kernels in Figure 2.

We conclude that the specification of no-arbitrage conditions and pricing kernels is far from straightforward. The implied unrestricted kernel in Panel A of Figures 2 and 3 indicates that matching the return and option data imposes stringent demands on the pricing kernel. While the affine price of risk specification in Panel C provides the best fit to the option data and the resulting marginal kernel better captures the patterns in the unrestricted kernel in Panel A of Figure 2, the resulting pricing kernel is excessively variable, even in stable economic times. The specification of the exponential-affine pricing kernel is consistent with economic theory, and this is evident in Panel B of Figures 2 and 3. The specification of the more general kernel in equation (8) is also subject to nonmonotonicity problems but the empirical implications (in Panel D of Figures 2 and 3) are

Panel B, the outliers in the financial crisis imply that the discount factor fluctuates between approximately 0.7 and 1.3. But in Panel A, the outliers during the financial crisis and throughout the entire sample yield extremely large and small discount factors (note that the y-axis in Panel A is scaled differently).

[24]Rather than illustrating the problems emanating from the affine price of risk through the time series of the implied pricing kernel, we can also inspect the time series of the prices of risk itself. Figure A.4 plots these time series. The affine price of equity risk in Panel C frequently spikes up when the variance becomes very small. This mechanism also underlies the problems with the implied pricing kernel, as evidenced by equations (5) and (7).

relatively limited.[25]

5.6 Expected Option Returns

Section 5.1 concludes that the joint-log likelihood of the various pricing kernels and price of risk specifications differs, and that some of these differences are statistically significant. Sections 5.4 and 5.5 show how the implied pricing kernels differ. However, Table 3 indicates that these differences do not result in differences in the vega-weighted option price RMSE. This reflects a well-known problem in the option pricing literature. The estimation of mis-specified models results in parameter estimates that optimize option fit, potentially at the cost of other model aspects. This implies that it is preferable to asses model differences along different dimensions. We now illustrate model differences using a criterion that embodies the models' economic implications.

The pricing kernel links risk-neutral and physical dynamics, and (expected) option returns incorporate information from both measures. The current option price is determined by the models' risk-neutral dynamics. In contrast, an index option's payoff is determined by the difference between the spot price at maturity and the strike price, and is therefore determined by the physical dynamics. We compare the hold-to-expiration realized option returns with the model-implied expected option returns. Since this comparison involves future option returns, the models' ability to fit expected option returns can be regarded as out-of-sample, as opposed to the in-sample fitting of current prices.

We compute the ex-post realized option return as:

$$Ret^{call}_{t,T} = \frac{(S(T)-K)^+}{E^*_t\left[e^{-r(T-t)}(S(T)-K)^+\right]}, \quad \text{and} \quad Ret^{put}_{t,T} = \frac{(K-S(T))^+}{E^*_t\left[e^{-r(T-t)}(K-S(T))^+\right]}, \tag{29}$$

and the expected option return as:

$$ERet^{call}_{t,T} = \frac{E_t\left[(S(T)-K)^+\right]}{E^*_t\left[e^{-r(T-t)}(S(T)-K)^+\right]}, \quad \text{and} \quad ERet^{put}_{t,T} = \frac{E_t\left[(K-S(T))^+\right]}{E^*_t\left[e^{-r(T-t)}(K-S(T))^+\right]}. \tag{30}$$

[25] We rely on one-month marginal pricing kernels and daily time series of pricing kernel innovations. Figure A.5 shows that we obtain similar differences based on monthly time series of pricing kernel innovations.

To ensure a consistent comparison over our sample period, we use the Volatility Surface data from OptionMetrics, which provides interpolated option-implied volatilities on a uniform grid of Black-Scholes Deltas and fixed maturities on each day. We further linearly interpolate implied volatilities on a uniform grid of standardized log moneyness at each maturity, defined as $(\log(K/S(t)))/\sqrt{v \times (T-t)}$.[26] We then investigate 1-, 3-, and 6-month returns. Consistent with our main sample, we use Wednesday option data. We include both in-the-money and out-of-the-money options. We compute both realized and expected option returns and compare their time-series averages.

Panels A through F of Figure 4 display the resulting expected option returns based on the exponential-affine pricing kernel (blue solid line), the affine price of risk specification (red dashed line), and the realized option returns (black dotted line). Regardless of option type (call/put), maturity, or moneyness, the exponential-affine kernel significantly outperforms the affine price of risk specification for the purpose of fitting the option returns. Panel G of Figure 4 indicates that that the exponential-affine kernel also outperforms the affine price of risk model for the purpose of fitting at-the-money (ATM) straddle returns. The time series averages of the annualized Sharpe ratios associated with the straddle returns in Panel C of Table 5 confirm these results. The differences between the average straddle returns in Panel C are larger than the differences between the risk premia in in Panel B.

We conclude that these findings highlight the importance of a well-specified pricing kernel to capture adequate physical and risk-neutral dynamics. We jointly estimate the physical and risk-neutral dynamics. While the affine price of risk specification slightly outperforms the exponential-affine pricing kernel when judged by the likelihood, the results for option returns suggest that this improved in-sample fit comes at the cost of overfitting, which surfaces when we inspect model implications that are not explicitly taken into account in the estimation, especially when these model implications depend on the model's performance under the P measure, as is the case for expected returns.

[26]To calculate this standardized log moneyness, we use the time series of spot variances (v) from the return-only estimation in Panel A of Table 2.

6 Restricting the Exponential-Affine Pricing Kernel

We study the implications of restricting the exponential-affine pricing kernel (12). Column (1) in Table 4 repeats the estimates for the exponential-affine kernel, while columns (7)-(12) report on estimation subject to various restrictions on this kernel. First, to understand the relative importance of index and variance risk aversion, we study the implications of restricting the preference parameters γ and ξ and we compare this with restrictions on the risk premium parameters μ and λ. We then impose a path-independence restriction, which provides insights into the recovery (Ross, 2015) literature. We discuss the implications of our findings for the literature on pricing kernel puzzles and the shape of the pricing kernel, we discuss the power utility pricing kernel, and we document the relative size of and time-variation in the components of the pricing kernel.

6.1 Restrictions on Preferences

Columns (9) and (10) in Table 4 restrict risk preferences, and set either γ or ξ equal to zero. The resulting estimates of respectively ξ and γ have the same sign as in column (1), but are larger in absolute value. The other parameters are very similar to column (1) and the overall likelihood is little affected. Panel A of Table 5 shows that setting one of the preference parameters (γ or ξ) equal to zero in columns (9) or (10) is not statistically rejected. The intuition for these findings is the robustly high (in absolute value) correlation between the innovations to aggregate variance and equity risk. When we set the representative agent's equity risk aversion or variance aversion equal to zero, the other parameter simply takes over the task of the restricted parameter. This is confirmed by the parameter estimates in columns (9) and (10) of Table 4, where respectively ξ and γ increase when the other parameter is set to zero. This finding is consistent with the recent findings of Dew-Becker and Giglio (2022) that the entire index variance risk premium can be entirely attributed to a negative stock beta.

Not surprisingly, column (9) in Table 5 shows that the $\gamma = 0$ restriction results in a substantially lower equity premium compared to column (1). Similarly, the $\xi = 0$ restriction in column (10) results in a smaller (in absolute value) variance risk premium. Panel B of Table 5 also shows that when setting $\xi = 0$, the Sharpe ratio remains very similar to the one in column (1), but that it is lowered significantly when imposing $\gamma = 0$. Panels A and B in Figure 5 plot the resulting marginal kernels and the time series of the kernels. The $\xi = 0$ has a much greater impact on the time series of the pricing kernel. Not surprisingly, the $\gamma = 0$ restrictions results in a more convex marginal kernel, due to the resulting higher estimate of variance risk aversion ξ.

Model misspecification once again manifests itself when exploring (out-of-sample) model features that are sensitive to the models' implications for the pricing kernel. Figure 6 illustrates this by comparing expected option returns based on the exponential-affine pricing kernel (the blue solid line) with the restricted specifications (the red dashed line for $\gamma = 0$ and the green dashed-dotted line for $\xi = 0$). The black dotted line plots the average return data. A first important conclusion is that the $\xi = 0$ specification robustly underperforms the unrestricted exponential-affine kernel. This finding highlights the critical role of variance aversion in linking the P- and the Q-dynamics and capturing the cross-section of option returns. The second observation is that while the $\gamma = 0$ specification has difficulty matching average put returns, it substantially outperforms the unrestricted exponential-affine kernel for call returns regardless of maturity and moneyness, and the outperformance increases as moneyness increases. This confirms our intuition: OTM call returns are highly negative, and call option payoffs critically depend on the probability of high positive underlying returns, which is captured by the right tail of the marginal pricing kernel. Panel A of Figure 5 shows that the higher positive ξ estimates when $\gamma = 0$ result in a U-shaped marginal pricing kernel which helps match these return patterns. A U-shaped kernel is consistent with a lower P-density and/or a higher Q-density in the right tail. This decreases the expected physical option payoff and/or increases option prices, which results in lower expected OTM call returns. Panel C of Table 5 reports descriptive statistics for the ATM straddle returns. The straddle returns illustrate the implications of these restrictions more dramatically than the risk premia in in Panel B.

Finally, column (11) in Table 4 imposes the joint restriction $\gamma = \xi = 0$, which is equivalent to setting the risk premium parameters equal to zero ($\mu = \lambda = 0$). These restrictions impact the estimates of the drift parameters κ and θ, but not the parameters ρ and σ, which determine skewness and kurtosis. In all three columns, the risk-neutral parameter estimates κ^*, θ^*, σ, and ρ are similar to column (1), and consequently once again the vega-weighted option RMSEs are also very similar. However, the economic implications are considerable. Panel B of Table 5 shows that by design the resulting risk premia are zero. More interestingly, Panel C reports that the resulting fit for ATM straddle returns considerably worsens.

We conclude that option pricing models can imply very different equity and variance risk premia. These risk premia cannot be easily distinguished statistically even when using returns and options in estimation. These findings are related to the observation of Merton (1980) that it is difficult to precisely estimate the equity premium. Our findings are also related to Bates (2003), who argues 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. Our results show that even when using a joint likelihood based on returns and options, it is difficult to statistically distinguish the option fit of models with very different economic implications.[27]

6.2 Restrictions on Risk Premiums

When we instead impose restrictions on the risk premium parameters μ and λ, the resulting estimates of respectively γ and ξ have the wrong sign. The implied P-parameters also substantially differ from the estimates in column (1). While the risk-neutral parameter estimates κ^*, θ^*, σ, and ρ remain remarkably similar and the vega-weighted option RMSEs are nearly identical, there are small differences between the log-likelihoods in columns (1), (7) and (8). Panel A of Table 5 shows that the restriction $\mu = 0$ is rejected at the 1% level against the exponential-affine kernel in column (1), and the same holds for the restriction $\lambda = 0$. Panel B of Table 5 also indicates that restricting one of the risk premium parameters to zero greatly reduces the size of the other risk premium. This is consistent with the smaller (in absolute value) estimates of the other risk premium parameter in Table 4. Finally, Panel C of Table 5 reports on the model fit of ATM straddle returns. Restricting

[27]Bardgett, Gourier, and Leippold (2019) and Jackwerth and Vilkov (2019) propose to better identify equity and variance risk premia by including VIX options in estimation together with SPX options.

one of the two risk premium parameters (μ or λ) to zero worsens the fit to option returns much more than setting one of the preference parameters (γ or ξ) to zero. In the latter case, as discussed above, the other parameter can take over the task of the restricted parameters due to the high correlation between the innovations to aggregate variance and equity risk.

6.3 Recovery Theory

Ross (2015) proposes to retrieve the physical distribution of the underlying asset using recovery theory, which is accomplished by imposing path-independence.[28] In columns (1) and (7)-(11) of Table 4, the martingale property implies the value of η, and the resulting estimates are different from zero. The exponential affine kernel contains a permanent component proportional to a power of the stock price and a transient variance-dependent component. Alvarez and Jermann (2005) show that the permanent component of the pricing kernel is quite variable. Bakshi, Chabi-Yo, and Gao (2017) reject the restriction from Ross recovery theory that the permanent component is constant.

Column (12) in Table 4 imposes the path-independence restriction, which amounts to $\eta = 0$, on the exponential-affine pricing kernel. This restriction lowers the likelihood ratio compared to the exponential-affine kernel in column (1). The $\eta = 0$ restriction results in a negative estimate of the variance preference parameter ξ, contrary to economic intuition. Table 5 shows that it also results in a smaller (in absolute value) variance risk premium, and Panel G of Figure 5 indicates that the restriction greatly affects the time series of the pricing kernel. We therefore confirm existing findings (Jackwerth and Menner, 2020) that the path-independence restriction is not supported by the data, and advocate the use of the exponential-affine kernel in equation (12), which does not impose this restriction.

[28] For other contributions to recovery theory, see Borovička, Hansen, and Scheinkman (2016), Jensen, Lando, and Pedersen (2019), Martin and Ross (2019), and Qin and Linetsky (2016).

6.4 The Pricing Kernel Puzzle

We previously discussed how variance-dependent pricing kernels can lead to nonmonotonicities in the pricing kernel. Several papers have argued that this mechanism may provide an explanation of the U-shaped pricing kernel puzzle of Jackwerth (2000). Christoffersen, Heston, and Jacobs (2013) propose this argument using a GARCH model. Song and Xiu (2016) make a similar point in a continuous-time setup using stochastic volatility. We now briefly discuss the implications of our estimates for this debate.

First, note that the marginal kernel in Panel A of Figure 2, which does not impose restrictions across the P- and Q-measures, is not U-shaped but rather W-shaped. As mentioned above, this is consistent with some existing findings (Cuesdeanu and Jackwerth, 2018); however, many studies document a U-shape. While the affine price of risk specification in Panel C of Figure 2 can capture the W-shaped pattern, we discussed above that resulting time series of the kernel in Panel C of Figure 3 is too variable.[29]

Panel B of Figure 2 shows that while the marginal kernel corresponding to the exponential-affine kernel is not linear in log returns due to the ξ parameter, the estimate of ξ is not sufficiently high to generate a U-shaped pricing kernel. In Section 7.1 below, we discuss how this stylized facts is impacted by the Feller conditions. Finally, Panel A of Figure 5 shows that we also obtain a U-shaped marginal kernel under the $\gamma = 0$ restriction. While the resulting time series of the kernel corresponds more closely to our priors, a pricing kernel with only variance aversion is not economically appealing. We conclude that while variance aversion can in principle generate a U-shaped pricing kernel, the estimates of our simple one-factor model do not readily produce one.

6.5 The Power Utility Pricing Kernel

We offer two observations on the power utility pricing kernel. First, the kernel in (17) differs from the traditional definition of the power utility pricing kernel. The seminal work by Rubinstein

[29]The W-shaped pattern resulting from the affine price of risk specification is more pronounced at longer horizons (see Figure A.2) and at low volatility levels (see Figure A.3).

(1976) and Brennan (1979) specifies economically motivated pricing kernels for economies with constant variance. The kernel in (17) highlights that we cannot simply adopt those static kernels in the presence of stochastic volatility. We need to include a path dependent component $\eta \int_0^t v(s)ds$ even in the absence of variance preference. As we showed in Section 6.3, it is possible to impose additional restrictions that eliminate this component, but these restrictions are not supported by the data.

A second remark is that risk-neutralization with affine prices of risk and zero intercepts cannot be supported by a power utility pricing kernel. If the parameters characterizing the equity and variance risk premium are left unconstrained, this argument is incorrect, as can be seen from equations (13) and (14). Indeed, with $\xi = 0$, we have $\mu = \gamma$ and $\lambda = \rho\sigma\gamma$, hence the two risk premiums are perfectly (negatively) correlated. In other words, the affine price of risk specification with zero intercepts relies on the existence of variance aversion and a nonzero ξ in the exponential-affine pricing kernel.

6.6 The Magnitudes of the Pricing Kernel Components

We decompose the pricing kernel into its different components: the one related to the stock return ($\gamma \log R(t)$) and to the change in variance ($\xi[v(t) - v(t-1)]$). Figure 8 presents results for the exponential-affine specification in column (1) of Table 3. We omit the path-dependent component for convenience. It is small and very persistent. Figure 8 shows how the different components of the pricing kernel account for the overall variation of the 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 8 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 Robustness and Extensions

We first show how imposing the Feller conditions impacts the estimation results and the properties of the pricing kernels. Then we report on two additional restricted versions of the affine risk premium specification. We also investigate whether our results depend on the specification of the underlying dynamics. To this end, we consider two additional models: a stochastic volatility model with jumps in the return process and a stochastic volatility model with two volatility factors.

7.1 Imposing the Feller Condition(s)

Feller (1951) provides conditions for the inaccessibility of the origin for a square root process. Following Cox, Ingersoll, and Ross (1985), these conditions have often been discussed in relation to no-arbitrage conditions in term structure models and option pricing models with stochastic volatility (Heston, Loewenstein, and Willard, 2006; Wong and Heyde, 2006; Cheridito, Filipović, and Kimmel, 2007; Singleton, 2006). Heston, Loewenstein, and Willard (2006) articulate necessary conditions for absence of arbitrage: i) the Feller (1951) condition makes the origin inaccessible under both measures, so $2\kappa\theta > \sigma^2$ and $2\kappa^*\theta^* > \sigma^2$; or ii) if the origin is accessible then $\kappa\theta = \kappa^*\theta^*$. The first condition prevents positive prices for contingent claims that pay out when volatility hits zero if this is impossible under the physical measure.[30] The second condition says that if volatility can become zero, then the intercepts in the affine price of risk specification must also be zero.

The second condition always holds when specifying the exponential-affine pricing kernel in equation (17), which is equivalent to the affine price of risk specification with zero intercept. When the price of risk contains an intercept, or for the more general kernel in equation (8), absence of arbitrage requires imposing the restrictions under both measures articulated in condition i). Columns (2) and (3) in Table 6 repeat the estimation in columns (2) and (5) in Table 3 but now impose these no-arbitrage conditions. While these conditions are not required for no-arbitrage in the case of the exponential-affine kernel, for completeness we also report on this case in column

[30]This condition also prohibits the converse situation in which contingent claims that pay out when volatility vanishes have zero prices despite having a positive probability of payment.

(1). A first conclusion from Table 6 is that the resulting estimates of σ are smaller. Second, option fit worsens compared to Table 3. Third, the intercept λ_0 is driven to zero in columns (2) and (3). Finally, compared to Table 3 the estimate of equity risk aversion γ is much smaller and the estimate of ξ is much larger. This confirms our earlier findings that it is very difficult to identify the sources of the equity and variance risk premia using these data.

Panels C and D of Figure 5 are based on the parameter estimates in columns (2) and (3) of Table 6. The marginal kernels are similar but the time series of the kernels, which capture conditional information, differ. A comparison with Panels C and D of Figure 2 indicates that we now obtain U-shaped pricing kernels, which is due to the larger estimates of ξ. In summary, our results show that the Feller restrictions have important implications for the economic properties of these models.[31]

7.2 Restrictions on Affine Risk Premiums

Section 5.4 documents the counter-intuitive properties of the marginal pricing kernel implied by the most flexible affine price of risk specification. We now investigate two restricted versions of the affine risk premium specification: $\mu_0 = 0$ and $\lambda_0 = 0$. This allows us to determine whether the counter-intuitive patterns in the kernel are due to the affine price of equity risk or the affine price of variance risk, which includes a component inversely proportional to $v(t)$.[32]

Columns (3) and (4) of Table 3 present the parameter estimates, and Panels E and F of Figure 5 depict the time series of the realized log pricing kernels and the implied marginal pricing kernels. The parameter estimates, log-likelihoods, and option price RMSEs are generally similar to those of the unrestricted affine price of risk specification in column (2) of Table 3. The pricing kernel

[31] Existing studies do not always discuss whether they impose the Feller condition in estimation. For option-based and joint (based on options and returns) parameter estimates in which the Feller condition is satisfied, see for example Eraker (2004), Pan (2002), Aït-Sahalia and Kimmel (2007), and Christoffersen, Jacobs, and Mimouni (2010). See for instance Bakshi, Cao, and Chen (1997), Bates (2000), and Hurn, Lindsay, and McClelland (2015) for parameter estimates where the Feller condition is not satisfied.

[32] Online Appendix F shows that $\mu_0 = 0$ corresponds to the case where the equity risk premium, independent of variance risk, is solely proportional to $\sqrt{v(t)}$ and not to $1/\sqrt{v(t)}$. Similarly, it shows that $\lambda_0 = 0$ corresponds to the case where the variance risk premium, independent of equity risk, is solely proportional to $\sqrt{v(t)}$.

patterns for the restricted models are therefore also similar to those for the unrestricted affine specification in Panel C of Figures 2 and 3. For instance, the daily realized log pricing kernels are not only volatile in the financial crisis, but also during some relatively calm periods. We conclude that the problems with the affine price of risk specification are not exclusively due to equity return risk or variance risk.[33]

Another interesting case nested by the affine price of risk specification is a minimum variance level (variance floor). Consider the following risk-neutral dynamic:

$$\begin{aligned} dS(t)/S(t) &= rdt + \sqrt{v_0 + v(t)}dz_1^*(t), \\ dv(t) &= \kappa^*(\theta^* - v(t))dt + \sigma\sqrt{v(t)}dz_2^*(t), \end{aligned}$$

where v_0 represents the minimum variance level, which we set to 0.0006 in our implementation. This is admittedly ad-hoc. It is based on the minimum value of the annualized 5-minute realized variance in our sample, which is 0.000625 (or 0.025 in volatility terms). Online Appendix G provides a more detailed description of the dynamics, as well as the resulting option valuation formula.

Column (1) of Table A.1 reports the resulting parameter estimates, and Panel A of Figure A.6 depicts the implied marginal pricing kernel. While the weighted log likelihood slightly exceeds the likelihood in the main specification in Table 3, the vega-weighted option price RMSE is indistinguishable. Note that under this alternative specification, the equity premium contains an intercept: it is equal to $\mu v_0 + \mu v(t)$. However, unlike Panel F of Figure 5, which is also based on an equity premium with an intercept, the pricing kernel in Panel A of Figure A.6 is monotonic in the log return. Intuitively, this result is due to the fact that the minimum-variance specification does not require a price of risk inversely proportional to $v(t)$.

These findings once again confirm that the essence of the problem is the specification of the pricing kernel and/or the prices of risk, and not the existence of an intercept in the risk premiums

[33] Note that the more general kernel in equation (8) also produces risk prices proportional to $1/\sqrt{v}$ or risk premiums with constant terms. However, because these are constrained by the structure of the kernel, this does not translate into similar patterns in the pricing kernel.

in itself.

7.3 A Stochastic Volatility Model with Jumps in Returns

We augment our baseline stochastic volatility model with jumps in returns. The risk-neutral dynamics are as follows:

$$
\begin{aligned}
dS(t)/S(t) &= (r - \nu^* l^*(t))\, dt + \sqrt{v(t)} dz_1^*(t) + \left[e^{J^*(t)} - 1\right] dN^*(t), \\
dv(t) &= \kappa^*(\theta^* - v(t))dt + \sigma\sqrt{v(t)} dz_2^*(t),
\end{aligned}
\tag{31}
$$

where the jump occurrence $N^*(t)$ follows a Poisson distribution with intensity $l^*(t) = l_0^* + l_1^* v(t)$. The jump compensation is denoted by $\nu^* l^*(t)$ with associated parameter $\nu^* = E^*\left[e^{J^*(t)} - 1\right]$, and the jump size $J^*(t)$ follows a normal distribution of mean μ_J^* and standard deviation σ_J. These assumptions imply that ν^* is governed by the moment generating function of a normally distributed random variable, which gives $\nu^* = e^{\mu_J^* - \frac{1}{2}\sigma_J^2} - 1$. We focus on two specifications: one with time-varying jump intensity ($l_0^* = 0$) and one with constant jump intensity ($l_1^* = 0$).

To obtain the specification of the physical dynamics, suppose the jump component of the pricing kernel dynamic has the form $\left[e^{-\zeta J(t)} - 1\right] dN(t)$, where $J(t)$ and $N(t)$ correspond to the jump size and the jump occurrence under the physical measure, and ζ represents the price of risk for jumps. Note that the pricing kernel in equation (8) or (12) implies $\zeta = \gamma$. We treat ζ as a free parameter under the affine POR specification. Girsanov's theorem then implies the following physical return dynamic:

$$
dS(t)/S(t) = [r + \mu(v(t)) - \nu l(t)]\, dt + \sqrt{v(t)} dz_1(t) + \left[e^{J(t)} - 1\right] dN(t),
\tag{32}
$$

where $J(t) \sim \mathcal{N}(\mu_J, \sigma_J)$ with $\mu_J = \mu_J^* + \zeta\sigma_J^2$, $\nu = E\left[e^{J(t)} - 1\right]$, and $l(t) = l^*(t)/E\left[e^{-\zeta J(t)}\right]$. We re-express the jump intensity as $l(t) = l_0 + l_1 v(t)$, where $l_i = l_i^*/E\left[e^{-\zeta J(t)}\right]$ for $i = 1, 2$.

Due to the presence of the jump component, we need to adopt a slightly different estimation approach. To facilitate the calculation of the return-based likelihood, we assume that at most one jump occurs within a single day.[34] Bayes' rule then implies the following log likelihood function:

$$\begin{aligned}
\log L^R &= \sum_{t=1}^{T-1} \log\left[f(\log R(t+\Delta), v(t+\Delta)|v(t)) \times \mathcal{J}(t+\Delta)\right] \\
&\approx \sum_{t=1}^{T-1} \log\left[(f(\log R(t+\Delta), v(t+\Delta)|v(t), \Delta N(t)=0)P(\Delta N(t)=0)\right. \\
&\qquad \left. +\ f(\log R(t+\Delta), v(t+\Delta)|v(t), \Delta N(t)=1)P(\Delta N(t)=1)) \times \mathcal{J}(t+\Delta)\right],
\end{aligned}$$

where the Poisson probability $P(\Delta N(t) = k) = \frac{(l(t)\Delta)^k e^{-(l(t)\Delta)}}{k!}$. Moreover, the conditional distribution of returns and variances given one jump is given by:

$$\begin{aligned}
& f(\log R(t+\Delta), v(t+\Delta)|v(t), \Delta N(t)=1) \\
&\qquad = \int_{-\infty}^{\infty} f(\log R(t+\Delta), v(t+\Delta)|v(t), \Delta N(t)=1, J(t)=j)P(J(t)=j)dj.
\end{aligned}$$

We price options following Duffie, Pan, and Singleton (2000) and calculate the option-based likelihood as in equation (26). We perform joint maximum likelihood estimation as described in Section 4.3 using both the return-based and option-based likelihoods. Table 7 reports the resulting parameter estimates. Both the estimated signs and magnitudes of the risk preference parameters are very similar to those in Table 3. A notable difference concerns the variance dynamics: for all specifications, the estimate of σ is smaller. This is to be expected, because the jump component now accounts for part of the return kurtosis.

In both the stochastic and constant jump intensity cases, the annualized jump intensity parameter (l_0 or l_1) and the average jump size parameter (μ_J) under the affine jump-diffusion price of risk specification (columns (2) and (5)) are lower compared to the cases where we directly specify the pricing kernel (columns (1), (3), (4), and (6)). Incorporating the jump process noticeably improves model fit, especially for option prices, under both stochastic and constant intensity models. Panels

[34] This is a reasonable assumption because the likelihood that more than one jump occurs is negligible. For instance, $P(\Delta N(t) \leq 1) = 100\%$ and 99.92% when $l(t) = 1$ and 10, respectively. Moreover, our estimation results imply that $l(t)$ does not exceed 2 in our sample, i.e., on average, we have no more than two jumps per year.

A through F of Figure 7 display the implied log marginal pricing kernels. Even when including jumps, the marginal pricing kernel for the affine price of risk specification is very different and displays nonmonotonicities.[35]

7.4 A Two-Factor Stochastic Volatility Model

We next investigate a two-factor stochastic volatility, which assumes the following risk-neutral underlying dynamics.

$$\begin{aligned} dS(t)/S(t) &= rdt + \sqrt{v_1(t)}dz_{11}^*(t) + \sqrt{v_2(t)}dz_{12}^*(t), \\ dv_1(t) &= \kappa_1^*(\theta_1^* - v_1(t))dt + \sigma_1\sqrt{v_1(t)}dz_{21}^*(t), \\ dv_2(t) &= \kappa_2^*(\theta_2^* - v_2(t))dt + \sigma_2\sqrt{v_2(t)}dz_{22}^*(t), \end{aligned} \tag{33}$$

where z_{11}^* and z_{21}^* are Wiener processes with correlation coefficient ρ_1, and z_{12}^* and z_{22}^* are Wiener processes with correlation coefficient ρ_2. All other correlations among these Wiener processes are zero. The total variance is $v(t) = v_1(t) + v_2(t)$.

We assume that both risk factors v_1 and v_2 are priced. To maintain consistency with the pricing kernel specifications used in our main analysis, we further assume that the two factors share common risk preference parameters, ξ and/or α.[36] For the affine price of risk (or affine risk premia) specification, we assume that the equity premium is given by $\mu_0 + \mu_1 v_1(t) + \mu_2 v_2(t)$, and the risk premium for the i-th factor is given by $\lambda_{0i} + \lambda_i v_i(t)$. Under these assumptions, we obtain

[35] The implied 6-month marginal pricing kernels (not reported) exhibit even more pronounced differences between the exponential-affine pricing kernel and the kernel implied by the affine jump-diffusion price of risk.

[36] That is, $M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^{-\gamma}\left(\frac{v_1(t)}{v_1(0)}\right)^{\alpha}\left(\frac{v_2(t)}{v_2(0)}\right)^{\alpha}\exp\left(\beta t + \eta_1\int_0^t v_1(s)ds + \eta_2\int_0^t v_2(s)ds + \xi(v(t) - v(0))\right)$ for the more general kernel.

the following physical dynamics for the return and the two factors:

$$\begin{aligned} dS(t)/S(t) &= (r + \mu_0 + \mu_1 v_1(t) + \mu_2 v_2(t))\, dt + \sqrt{v_1(t)} dz_{11}(t) + \sqrt{v_2(t)} dz_{12}(t), \\ dv_1(t) &= \kappa_1(\theta_1 - v_1(t))dt + \sigma_1 \sqrt{v_1(t)} dz_{21}(t), \\ dv_2(t) &= \kappa_2(\theta_2 - v_2(t))dt + \sigma_2 \sqrt{v_2(t)} dz_{22}(t), \end{aligned} \tag{34}$$

where $\kappa_i = \kappa_i^* - \lambda_i$ and $\theta_i = (\kappa_i^* \theta_i^* + \lambda_{0i})/\kappa_i$ for $i = 1$ and 2.

Once again, we need to adopt a slightly different estimation approach for this richer specification. We can no longer use the relationship between $v(t)$ and $\text{VIX}^2(t)$ in equation (20), because $v_1(t)$ and $v_2(t)$ cannot be separately identified through this relation. Instead, we use the exact formula for $\text{VIX}^2(t)$ similar to equation (18) as well as a measurement error to filter $v_1(t)$ and $v_2(t)$ from the time series of $\text{VIX}^2(t)$.[37] We use the linear Kalman filter, following Duan and Simonato (1999) and Chen and Scott (2003) without parameter optimization during the filtering step. Once we filter out the two factors, we compute the return-based likelihood and the option-based likelihood to optimize the parameters through joint maximum likelihood estimation.[38] We use variance targeting by requiring $\theta_1 + \theta_2$ to match the sample return variance. Option prices are computed following Christoffersen, Heston, and Jacobs (2009).

Table 8 presents the resulting parameter estimates. The estimated risk preference parameters γ and ξ remain consistent with economic theory in columns (1) and (3). Panels G, F, and I of Figure 7 depict the implied log marginal pricing kernels, which confirm our baseline results. Overall, the results confirm our findings from the baseline model. However, we now obtain substantial differences in fit across pricing kernel or price of risk specifications.[39] In the two-factor setting, the presence of an additional volatility factor allows the additional price of risk parameters to play a more

[37] In the two factor model, $\text{VIX}^2(t)$ is given by $\text{VIX}^2(t) = \theta_1^* + \frac{e^{-\kappa_1^* \Delta_{1m}} - 1}{-\kappa_1^* \Delta_{1m}}(v_1(t) - \theta_1^*) + \theta_2^* + \frac{e^{-\kappa_2^* \Delta_{1m}} - 1}{-\kappa_2^* \Delta_{1m}}(v_2(t) - \theta_2^*)$.

[38] For completeness, we also estimate our baseline model specifications using the linear Kalman filter. Table A.1 and Figure A.6 in the Appendix show that our main results do not depend on the method used to extract the time-varying variance.

[39] This is not due to the estimation approach. We do not observe such differences in columns (2)–(4) of Table A.1, which report estimation results for the one-factor model using the same filtering method.

significant role. Model fit improves with the more general kernel (column (3)) and the affine price of risk specification (column (2)), and the parameter estimates vary across different specifications.

8 Conclusion

The pricing kernel is a critical concept in asset pricing, and options are instrumental to estimate the risk-neutral probabilities that are required to identify the pricing kernel. However, in the option pricing literature, the mapping between the risk-neutral and physical probabilities is often defined in terms of risk premia, which implicitly define the pricing kernel as the ratio of these probabilities. We show that seemingly benign affine risk premium specifications can result in pricing kernels and state prices that are inconsistent with economic theory or that exhibit identification problems.

Explicitly specifying volatility-dependent pricing kernels can avoid these problems. A parsimonious exponential-affine kernel that is monotonic in the variance and generalizes the economically appealing kernels in Rubinstein (1976) and Brennan (1979) is consistent with economic theory. We use these kernels to study the economic implications of affine price of risk specifications. The exponential-affine kernel is equivalent to zero intercepts in the affine price of risk specification, and results in estimates that are consistent with economic intuition. Nonzero intercepts in the price of risk specification can result in state prices are non-monotonic functions of volatility, or their dependence of volatility is difficult to identify. The resulting marginal kernels display multiple inflection points, and the time series of the pricing kernel is excessively volatile in calm times. We find that while the in-sample fit to option prices of this price of risk specification is good, the resulting models fails to match expected option returns.

We use the parsimonious exponential-affine kernel to examine the role of variance aversion and index return risk aversion in determining the equity and variance risk premiums. Because the innovations to market returns and market variance are highly negatively correlated, it is difficult to statistically identify the sources of these risk premiums, and this results in a very similar fit to option prices. However, the resulting state prices are very different, which results in very different predictions for option returns. Variance aversion is critical to match the patterns in expected option returns. Finally, the proposed kernels are path-dependent and nest path-independent kernels (Ross, 2015) as special cases, We find that path-independence is not supported by the data. We also investigate the implications of our pricing kernels for the pricing kernel puzzle of Jackwerth (2000).

Our implementation uses plain-vanilla index options. Bardgett, Gourier, and Leippold (2019) and Jackwerth and Vilkov (2019) propose to improve the identification of equity and variance risk premia by including VIX options in estimation together with SPX options. In future work, we plan to study the role of these additional data in identifying risk premia and pricing kernels within our framework. We also plan to investigate the predictive power of the various pricing kernels for future returns and macroeconomic variables.

References

Aït-Sahalia, Yacine, Julio Cacho-Diaz, and Roger J. A. Laeven, 2015, Modeling financial contagion using mutually exciting jump processes, *Journal of Financial Economics* 117, 585–606.

Aït-Sahalia, Yacine, Mustafa Karaman, and Loriano Mancini, 2020, The Term Structure of Equity and Variance Risk Premia, *Journal of Econometrics* 219, 204–230.

Aït-Sahalia, Yacine, and Robert Kimmel, 2007, Maximum Likelihood Estimation of Stochastic Volatility Models, *Journal of Financial Economics* 83, 413–452.

Aït-Sahalia, Yacine, Chenxu Li, and Chen Xu Li, 2021, Implied Stochastic Volatility Models, *Review of Financial Studies* 34, 394–450.

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.

Alvarez, Fernando, and Urban J. Jermann, 2005, Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth, *Econometrica* 73, 1977–2016.

Andersen, Torben G., Luca Benzoni, and Jesper Lund, 2002, An Empirical Investigation of Continuous-Time Models for Equity Returns, *Journal of Finance* 57, 1239–1284.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015a, Parametric Inference and Dynamic State Recovery from Option Panels, *Econometrica* 83, 1081–1145.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015b, The Risk Premia Embedded in Index Options, *Journal of Financial Economics* 117, 558–584.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2017, Short-Term Market Risks Implied by Weekly Options, *Journal of Finance* 72, 1335–1386.

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

Bakshi, Gurdip, Fousseni Chabi-Yo, and Xiaohui Gao, 2017, A Recovery that We Can Trust? Deducing and Testing the Restrictions of the Recovery Theorem, *The Review of Financial Studies* 31, 532–555.

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.

Bardgett, Chris, Elise Gourier, and Markus Leippold, 2019, Inferring Volatility Dynamics and Risk Premia from the S&P 500 and VIX Markets, *Journal of Financial Economics* 131, 593–618.

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., 1996, Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options, *Review of Financial Studies* 9, 69–107.

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.

Bates, David S., 2012, U.S. Stock Market Crash Risk, 1926-2010, *Journal of Financial Economics* 105, 229–259.

Bates, David S., 2019, How Crashes Develop: Intradaily Volatility and Crash Evolution, *Journal of Finance* 74, 193–238.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22, 4463–4492.

Borovička, Jaroslav, Lars Peter Hansen, and José A. Scheinkman, 2016, Misspecified Recovery, *Journal of Finance* 71, 2493–2544.

Breeden, Douglas T., and Robert H. Litzenberger, 1978, Prices of State-Contingent Claims Implicit in Option Prices, *Journal of Business* 51, 621–651.

Brennan, Michael J., 1979, The Pricing of Contingent Claims in Discrete Time Models, *Journal of Finance* 34, 53–68.

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.

Chabi-Yo, Fousseni, and Johnathan Loudis, 2020, The Conditional Expected Market Return, *Journal of Financial Economics* 137, 752–786.

Chabi-Yo, Fousseni, and Johnathan A. Loudis, 2023, A Decomposition of Conditional Risk Premia and Implications for Representative Agent Models, *Management Science*.

Chen, Ren-Raw, and Louis Scott, 2003, Multi-factor Cox-Ingersoll-Ross models of the term structure: Estimates and tests from a Kalman filter model, *The Journal of Real Estate Finance and Economics* 27, 143–172.

Cheridito, Patrick, Damir Filipović, and Robert L Kimmel, 2007, Market Price of Risk Specifications for Affine Models: Theory and Evidence, *Journal of Financial Economics* 83, 123–170.

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, 2009, The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well, *Management Science* 55, 1914–1932.

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.

Christoffersen, Peter, Kris Jacobs, and Chayawat Ornthanalai, 2012, Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options, *Journal of Financial Economics* 106, 447–472.

Cox, John C., Jonathan C. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.

Cuesdeanu, Horatio, and Jens Carsten Jackwerth, 2018, The Pricing Kernel Puzzle: Survey and Outlook, *Annals of Finance* 14, 289–329.

Dew-Becker, Ian, and Stefano Giglio, 2022, Risk Preferences Implied by Synthetic Options, *Working Paper, Northwestern University and Yale University.*

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

Du, Du, and Dan Luo, 2019, The Pricing of Jump Propagation: Evidence from Spot and Options Markets, *Management Science* 65, 2360–2387.

Duan, Jin-Chuan, and Jean-Guy Simonato, 1999, Estimating and testing exponential-affine term structure models by Kalman filter, *Review of Quantitative Finance and Accounting* 13, 111–135.

Duan, Jin-Chuan, and Weiqi Zhang, 2014, Forward-Looking Market Risk Premium, *Management Science* 60, 521–538.

Dufays, Arnaud, Kris Jacobs, Yuguo Liu, and Jeroen Rombouts, 2023, Fast Filtering with Large Option Panels: Implications for Asset Pricing, *Journal of Financial and Quantitative Analysis* pp. 1–56.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* 68, 1343–1376.

Eraker, Bjørn, 2004, Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, *Journal of Finance* 59, 1367–1404.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The Impact of Jumps in Volatility and Returns, *Journal of Finance* 58, 1269–1300.

Eraker, Bjørn, and Ivan Shaliastovich, 2008, An Equilibrium Guide to Designing Affine Pricing Models, *Mathematical Finance* 18, 519–543.

Eraker, Bjørn, and Yue Wu, 2017, Explaining the Negative Returns to VIX Futures and ETNs: An Equilibrium Approach, *Journal of Financial Economics* 125, 72–98.

Eraker, Bjørn, and Aoxiang Yang, 2019, The Price of Higher Order Catastrophe Insurance: The Case of VIX Options, *Working Paper, University of Wisconsin-Madison.*

Feller, William, 1951, Two Singular Diffusion Problems, *The Annals of Mathematics* 54, 173–182.

Gabaix, Xavier, 2012, An Exactly Solved Framework for Ten Puzzles in Macro-Finance, *Quarterly Journal of Economics* 127, 645–700.

Garcia, Rene, Richard Luger, and Eric Renault, 2003, Empirical Assessment of an Intertemporal Option Pricing Model with Latent Variables, *Journal of Econometrics* 116, 49–83.

Ghosh, Anisha, Christian Julliard, and Alex P. Taylor, 2017, What is the Consumption-CAPM Missing? An Information-Theoretic Framework for the Analysis of Asset Pricing Models, *Review of Financial Studies* 30, 442–504.

Gruber, Peter H., Claudio Tebaldi, and Fabio Trojani, 2021, The Price of the Smile and Variance Risk Premia, *Management Science* 67, 4056–4074.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of Security Market Data for Models of Dynamic Economies, *Journal of Political Economy* 99, 225–262.

Hansen, Lars Peter, and Ken Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.

Heston, Steven L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 327–343.

Heston, Steven L., Mark Loewenstein, and Gregory A. Willard, 2006, Options and Bubbles, *Review of Financial Studies* 20, 359–390.

Heston, Steven L., and Saikat Nandi, 2000, A Closed-Form GARCH Option Valuation Model, *Review of Financial Studies* 13, 585–625.

Hurn, A. Stan, Kenneth A. Lindsay, and Andrew J. McClelland, 2015, Estimating the Parameters of Stochastic Volatility Models Using Option Price Data, *Journal of Business & Economic Statistics* 33, 579–594.

Jackwerth, Jens, and Grigory Vilkov, 2019, Asymmetric Volatility Risk: Evidence from Option Markets, *Review of Finance* 23, 777–799.

Jackwerth, Jens Carsten, 2000, Recovering Risk Aversion from Option Prices and Realized Returns, *Review of Financial Studies* 13, 433–451.

Jackwerth, Jens Carsten, and Marco Menner, 2020, Does the Ross recovery theorem work empirically?, *Journal of Financial Economics* 137, 723–739.

Jackwerth, Jens Carsten, and Mark Rubinstein, 1996, Recovering Probability Distributions from Option Prices, *Journal of Finance* 51, 1611–1631.

Jensen, Christian Skov, David Lando, and Lasse Heje Pedersen, 2019, Generalized recovery, *Journal of Financial Economics* 133, 154–174.

Jones, Christopher S., 2003, The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, *Journal of Econometrics* 116, 181–224.

Li, Junye, and Gabriele Zinna, 2018, The Variance Risk Premium: Components, Term Structures, and Stock Return Predictability, *Journal of Business & Economic Statistics* 36, 411–425.

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.

Liu, Xiaoquan, Michael J. Brennan, and Yihong Xia, 2006, Option Pricing Kernels and the ICAPM, *Working Paper, University of California, Los Angeles.*

Martin, Ian, 2017, What Is the Expected Return on the Market?, *Quarterly Journal of Economics* 132, 367–433.

Martin, Ian W. R., and Stephen A. Ross, 2019, Notes on the Yield Curve, *Journal of Financial Economics* 134, 689–702.

Martin, Ian W. R., and Christian Wagner, 2019, What Is the Expected Return on a Stock?, *Journal of Finance* 74, 1887–1929.

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.

Sichert, Tobias, 2023, The Pricing Kernel is U-Shaped, *Working Paper, Stockholm School of Economics*.

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.

Varin, Cristiano, Nancy Reid, and David Firth, 2011, An Overview of Composite Likelihood Methods, *Statistica Sinica* pp. 5–42.

Wachter, Jessica A., 2013, Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?, *Journal of Finance* 68, 987–1035.

Wong, Bernard, and C.C. Heyde, 2006, On Changes of Measure in Stochastic Volatility Models, *Journal of Applied Mathematics and Stochastic Analysis* 1.

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	4,934	5,600	5,781	5,767	5,622	5,831	33,535
Average IV (%)	15.59	15.94	18.15	20.28	22.50	26.09	19.88
Average price	9.68	24.20	45.90	34.20	23.76	14.74	25.81
Average spread	1.06	1.46	1.84	1.63	1.41	1.19	1.44
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	4,144	7,032	6,346	4,284	5,225	6,504	33,535
Average IV (%)	20.85	20.35	19.70	19.39	19.32	19.71	19.88
Average price	6.67	12.22	19.70	26.71	33.96	51.49	25.81
Average spread	0.61	1.01	1.38	1.55	1.71	2.19	1.44
Panel C: Log Stock Returns (annualized)							
Mean (%)	8.56						
Standard deviation (%)	18.75						
Skewness	-0.26						
Kurtosis	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, 1996 to June 30, 2019 period. Mean and standard deviation are annualized.

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

Panel A: Return-Based Physical Parameters							
μ	κ	θ	σ	ρ	η_0	η_1	$\log L^R$
2.3252	3.8367	0.0353	0.5671	-0.7188	-0.0069	0.8993	43,332
(1.1272)	(1.3507)	(0.0096)	(0.0283)	(0.0123)	(0.0003)	(0.0242)	
Panel B: Option-Based Risk-Neutral Parameters							
μ^*	κ^*	θ^*	σ	ρ	η_0	η_1	$\log L^O$
0	0.8401	0.1085	0.7422	-0.7459	-0.0030	0.8526	84,468
	(0.0894)	(0.0082)	(0.0072)	(0.0025)	(0.0003)	(0.0094)	

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. The sample period is January 10, 1996 to June 26, 2019.

Table 3: Joint MLE Estimation. Various Pricing Kernels

	Exp-Affine	Affine POR			PK in Eq. (8)	
Restrictions	None	None	$\mu_0 = 0$	$\lambda_0 = 0$	None	$\xi = 0$
	(1)	(2)	(3)	(4)	(5)	(6)
Risk Preference Parameters						
γ	1.3088				1.3964	2.4662
	(0.2516)				(0.0550)	(0.0369)
ξ	1.8720				2.8557	0
	(0.5093)				(0.0833)	
α					-0.0399	-0.0168
					(0.0011)	(0.0010)
Implied Risk Premia Parameters						
μ_0		0.0418	0	0.0627	-0.0198	-0.0084
		(0.0020)		(0.0231)		
μ or μ_1	2.2402	1.1084	2.2400	0.5444	2.8170	2.4662
		(0.1643)	(0.2673)	(0.0985)		
λ_0		0.0173	0.0381	0	0.0163	0.0069
		(0.0013)	(0.0010)			
λ or λ_1	-1.4182	-1.8830	-2.4486	-1.4176	-1.8639	-1.2273
		(0.1076)	(0.0970)	(0.0702)		
P-Parameters						
κ	2.8157	3.2726	3.8365	2.8098	3.2587	2.6263
	(0.1722)	(0.1423)	(0.0998)	(0.0681)	(0.1086)	(0.0736)
θ	0.0371	0.0371	0.0370	0.0371	0.0370	0.0424
	(0.0022)	(0.0013)	(0.0007)	(0.0011)	(0.0004)	(0.0006)
σ	0.6401	0.6397	0.6395	0.6399	0.6399	0.6401
	(0.0208)	(0.0203)	(0.0211)	(0.0193)	(0.0214)	(0.0206)
ρ	-0.7774	-0.7774	-0.7775	-0.7774	-0.7774	-0.7774
	(0.0096)	(0.0093)	(0.0094)	(0.0095)	(0.0096)	(0.0095)
η_0	-0.0063	-0.0063	-0.0063	-0.0063	-0.0063	-0.0063
	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0004)	(0.0003)
η_1	0.9235	0.9235	0.9232	0.9236	0.9233	0.9237
	(0.0076)	(0.0064)	(0.0087)	(0.0055)	(0.0099)	(0.0085)
Q-Parameters						
κ^*	1.3976	1.3896	1.3879	1.3922	1.3948	1.3990
θ^*	0.0747	0.0749	0.0749	0.0748	0.0748	0.0746
Model Fit						
$\log L^R$	43,207.41	43,212.70	43,211.47	43,212.52	43,207.52	43,207.21
$\log L^O$	83,397.42	83,393.44	83,396.19	83,392.69	83,398.54	83,397.58
Weighted $\log L$	193,178.18	193,193.48	193,190.98	193,192.43	193,179.20	193,177.58
Vega-Weighted Option Price RMSE	0.0201	0.0201	0.0201	0.0201	0.0201	0.0201

Notes: This table reports the results of joint maximum likelihood estimation using the 1996-2019 sample. The risk preference parameters γ, ξ, and α (or the risk premia parameters μ_0, μ_1, λ_0, and λ_1 for columns (2)-(4)) as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions. Robust standard errors are in parentheses. Column (1) reports on the exponential-affine pricing kernel (PK) specification, column (2) on the affine price of risk (POR) specification, columns (3)-(4) on the affine POR subject to $\mu_0 = 0$ and $\lambda_0 = 0$, respectively, column (5) on the more general PK from equation (8), and column (6) on the more general PK with $\xi = 0$. We report the weighted log likelihood and the vega-weighted option price RMSE for each specification.

Table 4: Joint MLE Estimation. Restrictions on the Exponential-Affine Pricing Kernel

	Exponential-Affine						
Restrictions	None	$\mu = 0$	$\lambda = 0$	$\gamma = 0$	$\xi = 0$	$\gamma = \xi = 0$	$\eta = 0$
	(1)	(7)	(8)	(9)	(10)	(11)	(12)
Risk Preference Parameters							
γ	1.3088	-0.9309	1.3202	0	2.2399	0	2.1578
	(0.2516)				(0.4090)		
ξ	1.8720	1.8709	-1.6029	3.4666	0	0	-0.5279
	(0.5093)	(1.1679)	(0.0747)	(0.2792)			(0.0594)
Implied Risk Premia Parameters							
μ	2.2402	0	0.5224	1.7247	2.2399	0	1.8951
λ	-1.4182	-0.3033	0	-1.4199	-1.1147	0	-0.8576
P-Parameters							
κ	2.8157	1.7010	1.4013	2.8158	2.5142	1.3997	2.2580
	(0.1722)	(0.1982)	(0.0515)	(0.0650)	(0.0974)	(0.0688)	(0.0738)
θ	0.0371	0.0614	0.0745	0.0371	0.0415	0.0746	0.0463
	(0.0022)	(0.0070)	(0.0018)	(0.0013)	(0.0028)	(0.0026)	(0.0009)
σ	0.6401	0.6401	0.6403	0.6400	0.6402	0.6402	0.6402
	(0.0208)	(0.0209)	(0.0214)	(0.0201)	(0.0216)	(0.0208)	(0.0208)
ρ	-0.7774	-0.7774	-0.7774	-0.7774	-0.7774	-0.7774	-0.7774
	(0.0096)	(0.0096)	(0.0095)	(0.0095)	(0.0099)	(0.0096)	(0.0096)
η_0	-0.0063	-0.0063	-0.0063	-0.0063	-0.0063	-0.0063	-0.0063
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0003)
η_1	0.9235	0.9235	0.9239	0.9233	0.9237	0.9237	0.9238
	(0.0076)	(0.0076)	(0.0067)	(0.0076)	(0.0122)	(0.0076)	(0.0076)
Q-Parameters							
κ^*	1.3976	1.3977	1.4013	1.3959	1.3995	1.3997	1.4004
θ^*	0.0747	0.0747	0.0745	0.0748	0.0746	0.0746	0.0746
Model Fit							
$\log L^R$	43,207.41	43,205.24	43,205.37	43,207.07	43,207.19	43,205.01	43,207.00
$\log L^O$	83,397.42	83,397.42	83,397.00	83,397.71	83,397.32	83,397.32	83,397.19
Weighted $\log L$	193,178.18	193,170.92	193,171.10	193,177.20	193,177.36	193,170.10	193,176.66
Vega-Weighted Option Price RMSE	0.0201	0.0201	0.0201	0.0201	0.0201	0.0201	0.0201

Notes: This table reports the results of joint maximum likelihood estimation of the exponential-affine pricing kernel with various restrictions, using the 1996-2019 sample. 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) and (7) to (12). Robust standard errors are in parentheses. Column (1) reports on the exponential-affine kernel (column (1) in Table 3). The other columns report on the model subject to various restrictions. We report the weighted log likelihood and the vega-weighted option price RMSE for each specification.

Table 5: Economic Implications

		Exp-Affine	Affine POR	PK in Eq. (8)		Exponential-Affine					
Restrictions	Data	None	None	None	$\xi = 0$	$\mu = 0$	$\lambda = 0$	$\gamma = 0$	$\xi = 0$	$\gamma = \xi = 0$	$\eta = 0$
		(1)	(2)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Likelihood Ratio Test P-Values											
Comparison with (1)						0.0002	0.0003	0.1775	0.2174	0.0006	0.0931
Comparison with (2)		0.0002		0.0001	0.0002	1.8E-05	1.9E-05	0.0005	0.0005	3.7E-05	0.0004
Panel B: Risk Premiums and Sharpe Ratios											
Instantaneous ERP	0.0822	0.0828	0.0827	0.0843	0.0828	0	0.0193	0.0637	0.0828	0	0.0701
Instantaneous VRP		-0.0524	-0.0523	-0.0526	-0.0385	-0.0112	0	-0.0525	-0.0412	0	-0.0317
Annualized Sharpe Ratio	0.4310	0.3860	0.4917	0.3612	0.3661	0	0.0900	0.2971	0.3860	0	0.3266
Panel C: ATM Option Straddle Sharpe Ratios											
1M maturity	-0.1993	-0.0712	-0.0220	-0.0756	-0.0636	-0.0315	-0.0287	-0.0699	-0.0632	-0.0243	-0.0562
3M maturity	-0.3034	-0.1506	-0.0484	-0.1638	-0.1304	-0.0578	-0.0494	-0.1512	-0.1296	-0.0396	-0.1138
6M maturity	-0.2583	-0.1784	-0.0339	-0.1995	-0.1417	-0.0516	-0.0335	-0.1908	-0.1414	-0.0212	-0.1219
12M maturity	0.0008	-0.0734	0.2536	-0.1367	-0.0597	0.1048	0.0687	-0.1293	-0.0710	0.1495	-0.0611

Notes: Panel A reports the P-values for the composite Likelihood Ratio tests of Varin, Reid, and Firth (2011) against the specifications in columns (1) and (2). Panel B reports the sample means for the implied instantaneous equity and variance risk premia, and for the annualized Sharpe ratio. Panel C presents the Sharpe ratio of at-the-money option straddle position with maturities ranging from one month to one year.

Table 6: Joint MLE Estimation. Feller Conditions Imposed

Panel A: Parameter Estimates			
	Feller Condition Imposed		
	Exp-Affine	Affine POR	PK in Eq. (8)
	(1)	(2)	(3)
Risk Preference Parameters			
γ	0.2844		0.2844
	(0.0236)		(0.0116)
ξ	4.9723		4.9723
	(0.3061)		(0.5006)
α			-5.8E-12
			(2.9E-11)
Implied Risk Premia Parameters			
μ_0		0.0115	-2.4E-12
		(0.0008)	
μ or μ_1	2.3171	1.9920	2.3171
		(0.1313)	
λ_0		1.1E-08	1.6E-12
		(1.7E-10)	
λ or λ_1	-1.5155	-1.5155	-1.5155
		(0.0742)	
P-Parameters			
κ	3.9658	3.9642	3.9657
	(0.1028)	(0.0360)	(0.0505)
θ	0.0355	0.0355	0.0355
	(0.0007)	(0.0006)	(0.0012)
σ	0.5305	0.5303	0.5305
	(0.0150)	(0.0163)	(0.0146)
ρ	-0.7707	-0.7706	-0.7707
	(0.0089)	(0.0089)	(0.0091)
η_0	-0.0067	-0.0067	-0.0067
	(0.0002)	(0.0003)	(0.0003)
η_1	0.8977	0.8977	0.8977
	(0.0070)	(0.0074)	(0.0079)
Q-Parameters			
κ^*	2.4503	2.4487	2.4503
θ^*	0.0574	0.0574	0.0574
Model Fit			
$\log L^R$	43,187.33	43,187.25	43,187.33
$\log L^O$	78,685.94	78,687.37	78,685.95
Weighted $\log L$	190,340.09	190,340.63	190,340.09
Vega-Weighted Option Price RMSE	0.0232	0.0232	0.0232
Panel B: Economic Implications			
Instantaneous ERP	0.0820	0.0820	0.0820
Instantaneous VRP	-0.0536	-0.0536	-0.0536
Annualized Sharpe Ratio	0.3891	0.4209	0.3890

Notes: This table reports the results of joint maximum likelihood estimation using the 1996-2019 sample. The risk preference parameters γ, ξ, and α (or the risk premia parameters μ_0, μ_1, λ_0, and λ_1 for column (2)) as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions. Robust standard errors are in parentheses. Column (1) reports on the exponential-affine PK specification, column (2) on the affine price of risk (POR) specification, and column (3) on the general pricing kernel (PK) from equation (8). The Feller condition is imposed under both the P- and Q-measures.

Table 7: Joint MLE Estimation. Stochastic Volatility Model with Return Jumps

	Stochastic jump intensity			Constant jump intensity		
	Exp-Affine	AJD POR	PK in Eq. (8)	Exp-Affine	AJD POR	PK in Eq. (8)
	(1)	(2)	(3)	(4)	(5)	(6)
Risk Preference Parameters						
γ	1.8020		1.8068	1.6794		1.6883
	(0.3229)		(0.1264)	(0.0992)		(0.0951)
ξ	1.3983		4.4471	1.3770		3.9029
	(0.3995)		(0.5005)	(0.1663)		(0.1765)
α			-0.0931			-0.0934
			(0.0068)			(0.0111)
Implied Risk Premia Parameters						
μ_0		0.0340	-0.0334	0.0061	0.0466	-0.0338
		(0.0668)			(0.0125)	
μ or μ_1	2.7092	1.2996	3.8080	2.2688	1.0120	3.3567
		(0.3810)			(0.2188)	
λ_0		0.0228	0.0228		0.0342	0.0305
		(0.0009)			(0.0020)	
λ or λ_1	-0.9896	-1.7111	-1.7362	-1.1694	-2.1664	-1.9957
		(0.1508)			(0.2398)	
ζ		8.9559			4.0945	
		(0.2345)			(0.1047)	
P-Parameters						
κ	2.6084	3.3112	3.3493	2.2455	3.2269	3.0623
	(0.2704)	(0.0113)	(0.0630)	(0.0536)	(0.0266)	(0.0601)
θ	0.0305	0.0306	0.0305	0.0340	0.0341	0.0348
	(0.0042)	(0.0002)	(0.0004)	(0.0010)	(0.0002)	(0.0009)
σ	0.4952	0.4930	0.4947	0.5720	0.5712	0.5713
	(0.0200)	(0.0234)	(0.0040)	(0.0275)	(0.0301)	(0.0075)
ρ	-0.7247	-0.7245	-0.7247	-0.7483	-0.7484	-0.7482
	(0.0110)	(0.0133)	(0.0183)	(0.0089)	(0.0103)	(0.0148)
l_0	0	0	0	0.0653	0.0590	0.0665
				(0.0013)	(0.0013)	(0.0011)
l_1	3.6138	1.0205	3.6314	0	0	0
	(0.2396)	(0.0583)	(0.0714)			
μ_J	-0.2201	-0.1627	-0.2195	-0.1091	-0.0185	-0.1077
	(0.0176)	(0.0035)	(0.0070)	(0.0016)	(0.0011)	(0.0035)
σ_J	0.0809	0.0831	0.0810	0.1935	0.1906	0.1924
	(0.0069)	(0.0028)	(0.0026)	(0.0016)	(0.0008)	(0.0038)
η_0	-0.0042	-0.0041	-0.0042	-0.0062	-0.0062	-0.0061
	(0.0005)	(0.0004)	(0.0006)	(0.0003)	(0.0003)	(0.0003)
η_1	0.7370	0.7334	0.7364	0.8557	0.8551	0.8549
	(0.0205)	(0.0154)	(0.0249)	(0.0162)	(0.0170)	(0.0213)
Q-Parameters						
κ^*	1.6189	1.6001	1.6130	1.0761	1.0604	1.0666
θ^*	0.0491	0.0491	0.0491	0.0710	0.0714	0.0713
Model Fit						
$\log L^R$	43,232.33	43,236.27	43,232.11	43,287.64	43,297.19	43,289.32
$\log L^O$	86,110.71	86,131.70	86,117.21	84,762.62	84,760.94	84,763.11
Weighted $\log L$	194,857.14	194,882.62	194,860.23	194,248.73	194,279.63	194,254.64
Vega-Weighted Option Price RMSE	0.0186	0.0185	0.0186	0.0193	0.0193	0.0193

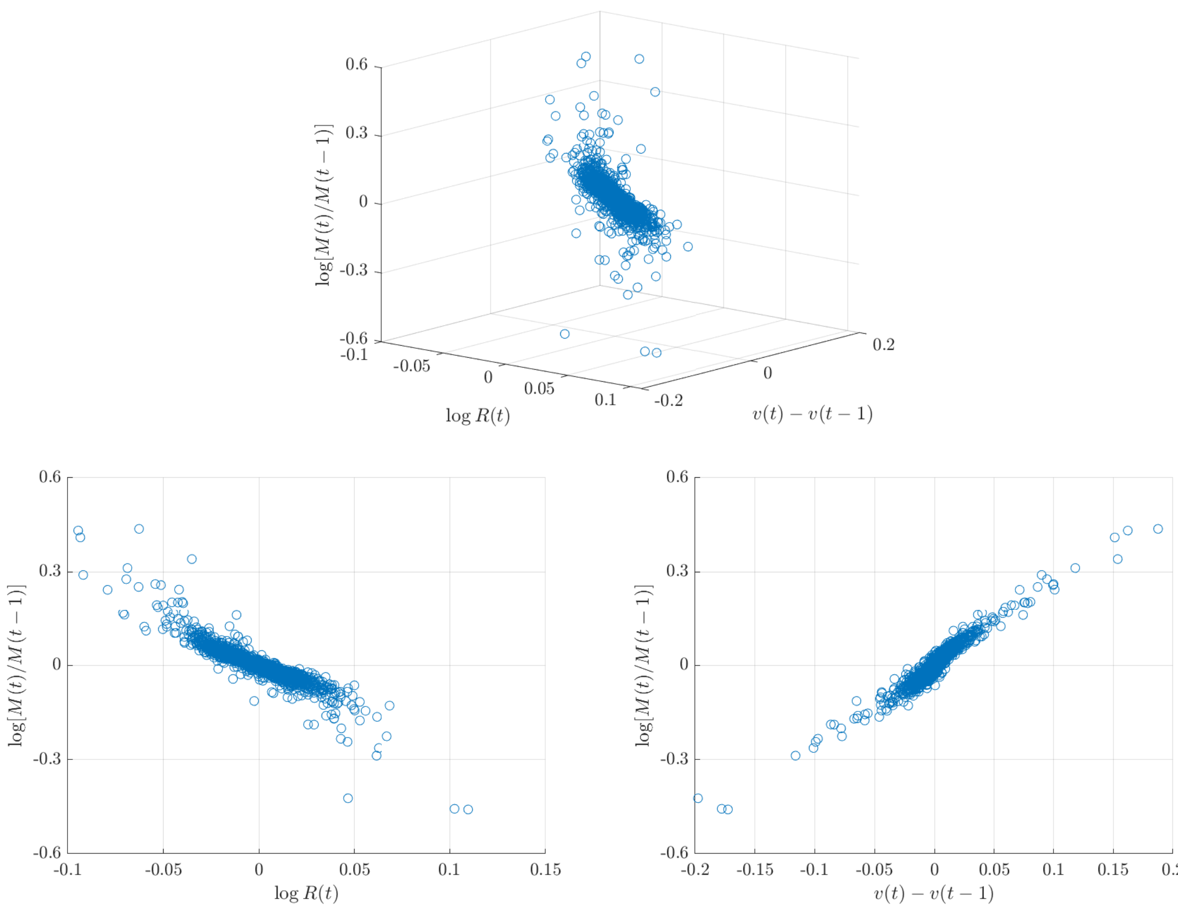
Notes: This table reports the results of joint maximum likelihood estimation of a model with one stochastic volatility factor and a return jump. AJD stands for Affine Jump Diffusion. Jump occurrence follows a Poisson distribution with jump intensity $l(t) = l_0 + l_1 v(t)$. The jump size $J(t)$ follows a normal distribution $N(\mu_J, \sigma_J^2)$. In columns (1), (3), (4), and (6), jumps in the pricing kernel are given by $\left(e^{-\gamma J(t)} - 1\right) dN(t)$. In the POR specification (columns (2) and (5)), the jumps in the pricing kernel follow $\left(e^{-\zeta J(t)} - 1\right) dN(t)$, with an additional jump price of risk parameter ζ.

Table 8: Joint MLE Estimation. Two-Factor Stochastic Volatility Model

	Two-factor		
	Exp-Affine	Affine POR	PK in Eq. (8)
	(1)	(2)	(3)
Risk Preference Parameters			
γ	0.9771		1.4835
	(0.1246)		(0.1079)
ξ	0.3056		0.7206
	(0.0543)		(0.0375)
α			-0.0141
			(0.0006)
Implied Risk Premia Parameters			
μ_0		0.0002	-0.0248
		(2.9E-05)	
μ_1	1.4062	1.0021	2.5007
		(0.1198)	
μ_2	1.0882	1.0622	1.7303
		(0.1011)	
λ_{01}		0.1412	0.0436
		(0.0020)	
λ_{02}		-0.0286	0.0023
		(0.0028)	
λ_1	-2.3439	-3.2205	-4.3195
		(5.9E-09)	
λ_2	-0.4126	-1.4009	-0.6273
		(0.1074)	
P-Parameters			
κ_1	8.2689	8.3490	9.7545
	(0.1994)	(0.0024)	(1.3E-05)
θ_1	0.0182	0.0183	0.0175
	(0.0002)	(0.0001)	(0.0001)
σ_1	1.7835	1.8163	1.7574
	(0.0338)	(3.2E-10)	(0.0240)
ρ_1	-0.7872	-0.9011	-0.8033
	(2.8E-07)	(0.0003)	(8.8E-07)
κ_2	1.0824	1.5886	1.2508
	(0.0273)	(0.0510)	(0.0222)
θ_2	0.0169	0.0169	0.0176
σ_2	0.4324	0.4744	0.4066
	(0.0066)	(0.0042)	(0.0050)
ρ_2	-0.8413	-0.8068	-0.8425
	(0.0072)	(0.0057)	(0.0068)
Q-Parameters			
κ_1^*	5.9250	5.1285	5.4350
θ_1^*	0.0255	0.0022	0.0234
κ_2^*	0.6699	0.1877	0.6235
θ_2^*	0.0273	0.2953	0.0316
Model Fit			
$\log L^R$	21,507.74	21,485.27	21,528.56
$\log L^O$	85,621.13	88,537.42	86,209.29
Weighted $\log L$	122,102.44	123,742.76	122,517.84
Vega-Weighted Option Price RMSE	0.0188	0.0173	0.0185

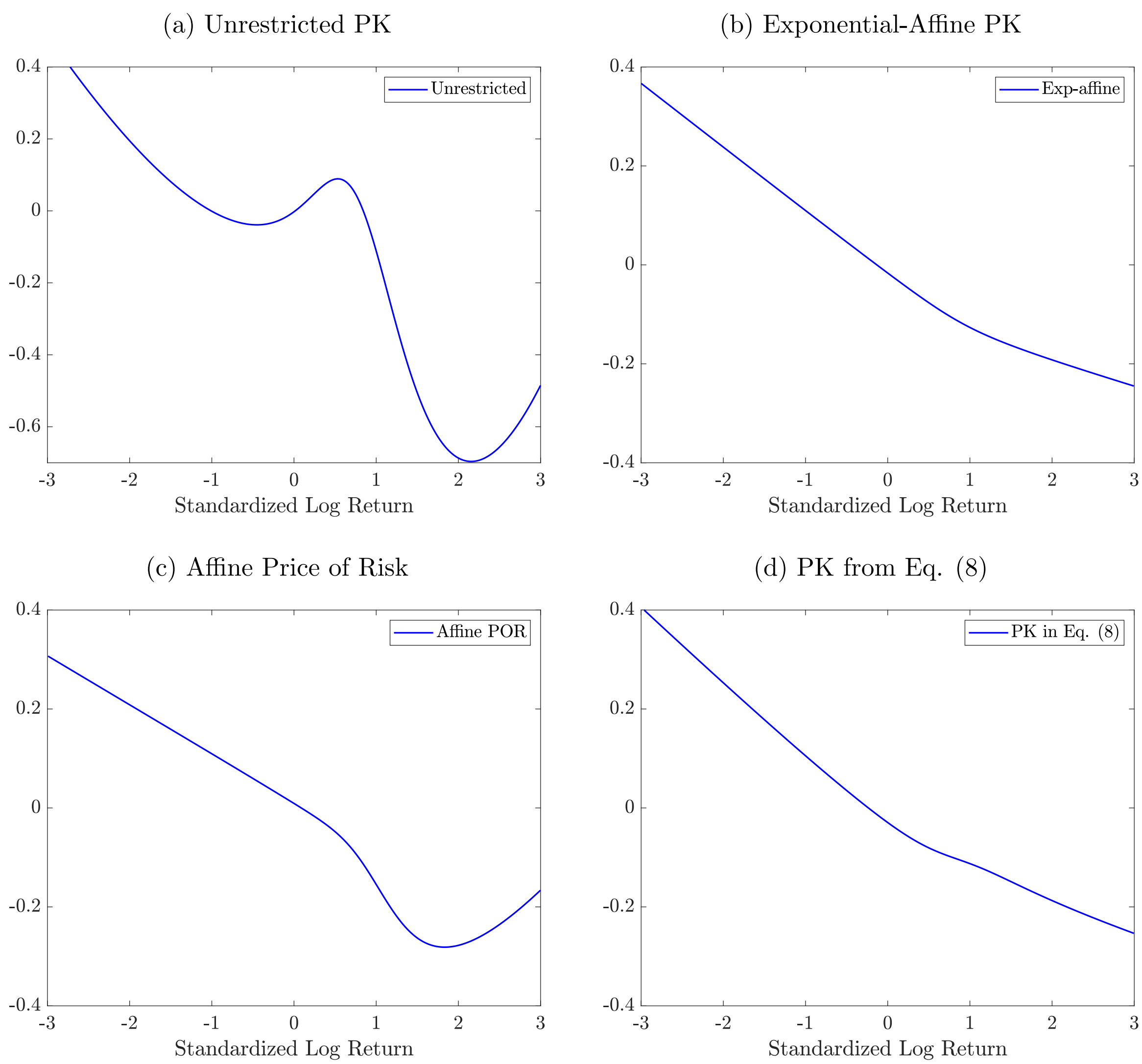
Notes: This table reports the results of joint maximum likelihood estimation of a two-factor stochastic volatility model using the 1996-2019 sample. The risk preference parameters γ, ξ, and α (or the risk premia parameters μ_0, μ_1, λ_0, and λ_1 for columns (2)-(4)) as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions. Robust standard errors are in parentheses. Column (1) reports on the exponential-affine pricing kernel (PK) specification, column (2) on the affine price of risk (POR) specification, and column (3) reports on the more general PK from equation (8). We report the weighted log likelihood and the vega-weighted option price RMSE for each specification.

Figure 1: 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.

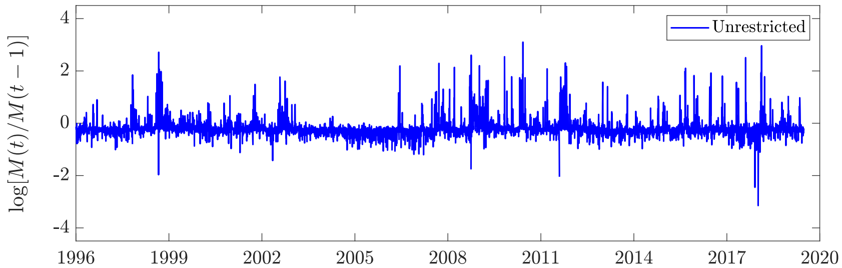
Figure 2: Implied Log Marginal Pricing Kernels



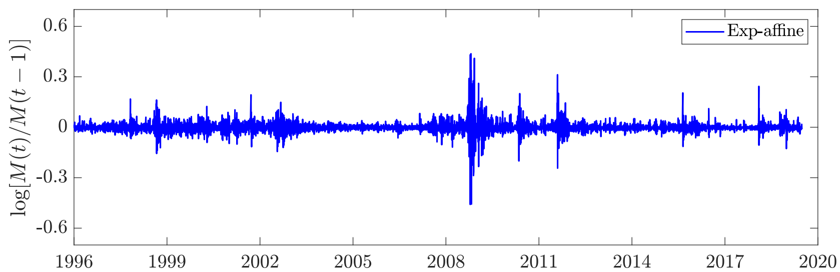
Notes: We plot the implied 1-month log marginal pricing kernels for the following four specifications: independent estimation of the P- and Q-parameters (Panel A), the unrestricted exponential-affine pricing kernel (Panel B), the affine price of risk specification (Panel C), and the unrestricted pricing kernel from equation (8) (Panel D). Parameter values for Panel A are from Table 2, and those for Panels B, C, and D are from Table 3. The x-axis represents the log return in deviation from its mean and divided by its standard deviation.

Figure 3: Daily (Realized) Log Pricing Kernels

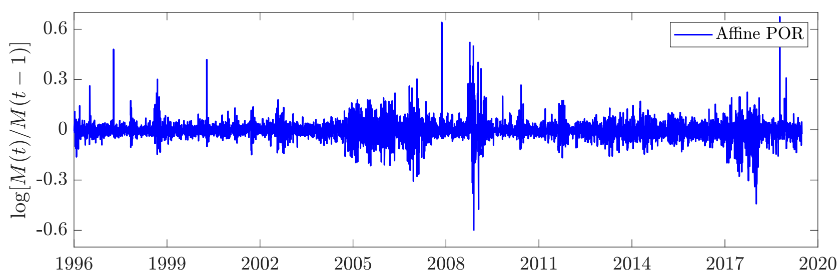
(a) Unrestricted PK



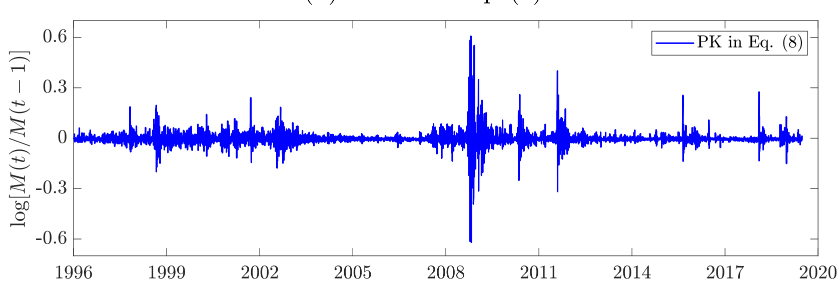
(b) Exponential-Affine PK



(c) Affine Price of Risk

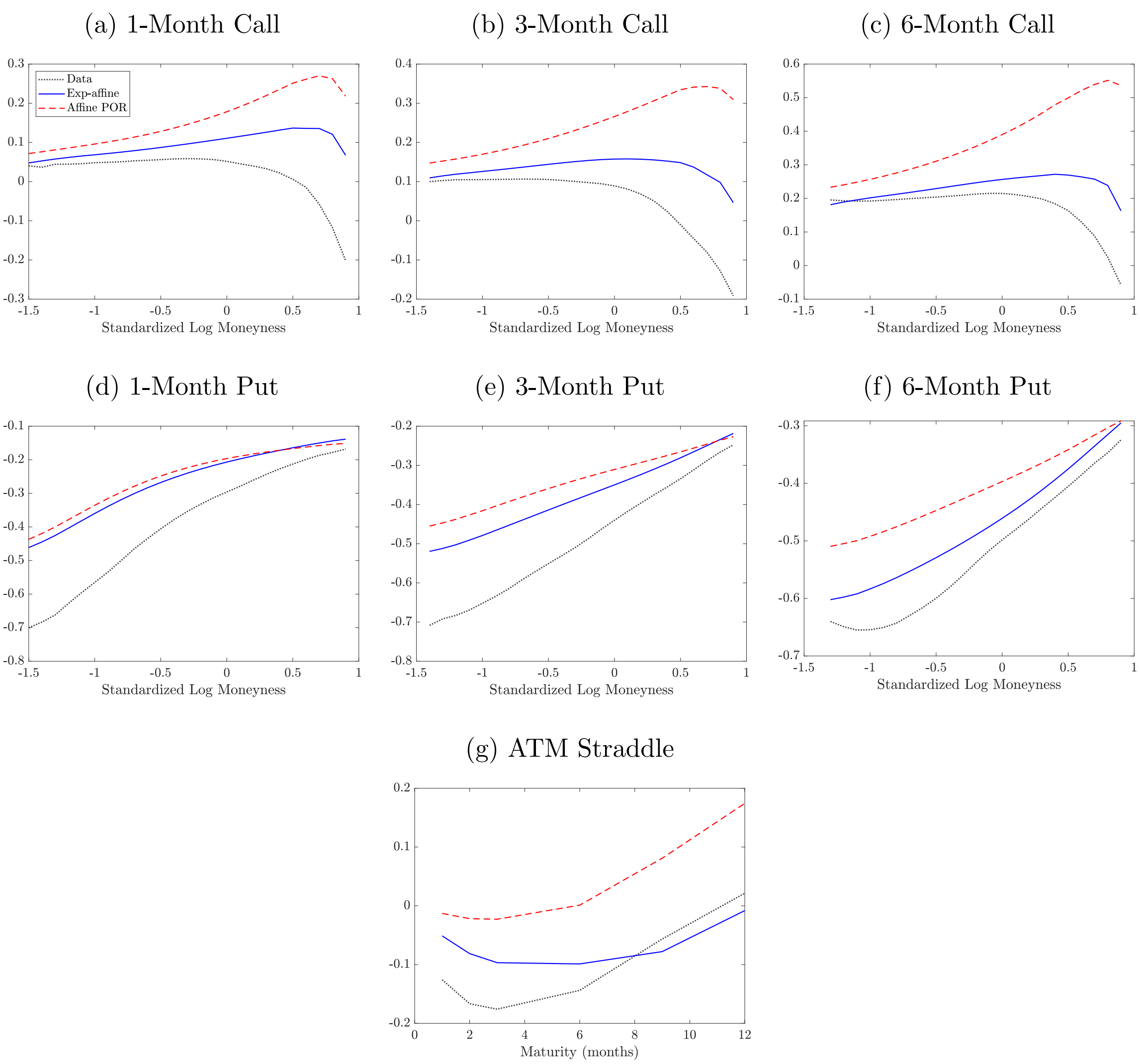


(d) PK from Eq. (8)



Notes: We plot the time series of the daily log pricing kernels for the following four specifications: independent estimation of the P- and Q-parameters (Panel A), the unrestricted exponential-affine (Panel B), the affine price of risk specification (Panel C), and the unrestricted pricing kernel from equation (8) (Panel D). Parameter values for Panel A are from Table 2, and those for Panels B, C, and D are from Table 3.

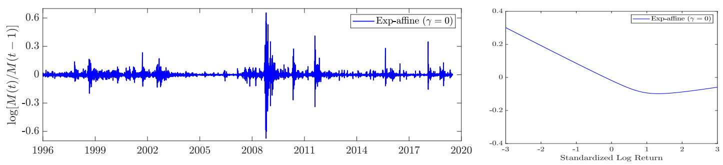
Figure 4: Expected Option Returns



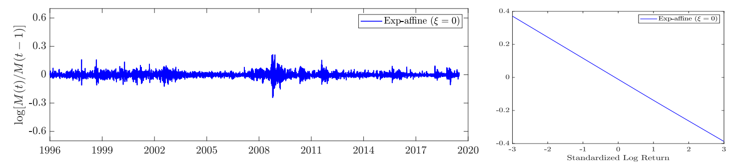
Notes: We plot model-implied expected option returns for the exponential-affine pricing kernel and the affine price of risk specification, and compare them to the data. Panels A through F plot time series averages of realized option returns and expected returns as a function of option moneyness. Panel G plots time series averages of realized ATM straddle returns and expected returns as a function of maturity.

Figure 5: Other Log Pricing Kernels

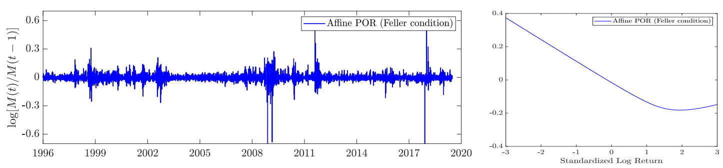
(a) Exponential-Affine with $\gamma = 0$



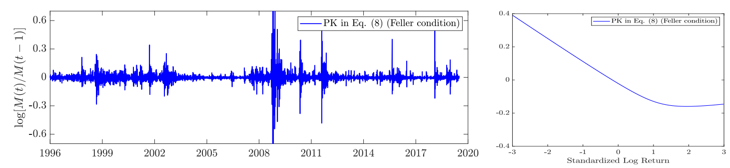
(b) Exponential-Affine with $\xi = 0$ (no independent variance premium)



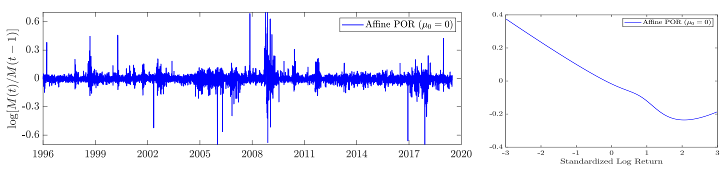
(c) Affine POR with the Feller No-Arbitrage Conditions Imposed



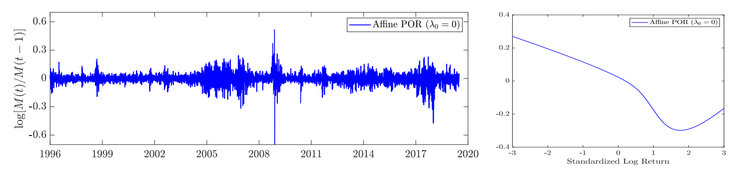
(d) PK from Eq. (8) with the Feller No-Arbitrage Conditions Imposed



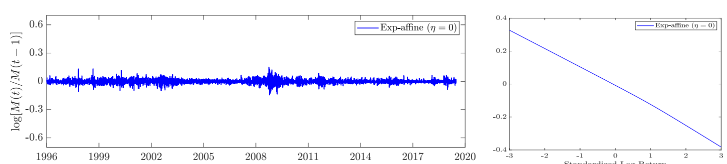
(e) Affine Price of Risk with $\mu_0 = 0$ (no intercept in equity premium)



(f) Affine Price of Risk with $\lambda_0 = 0$ (no intercept in variance premium)

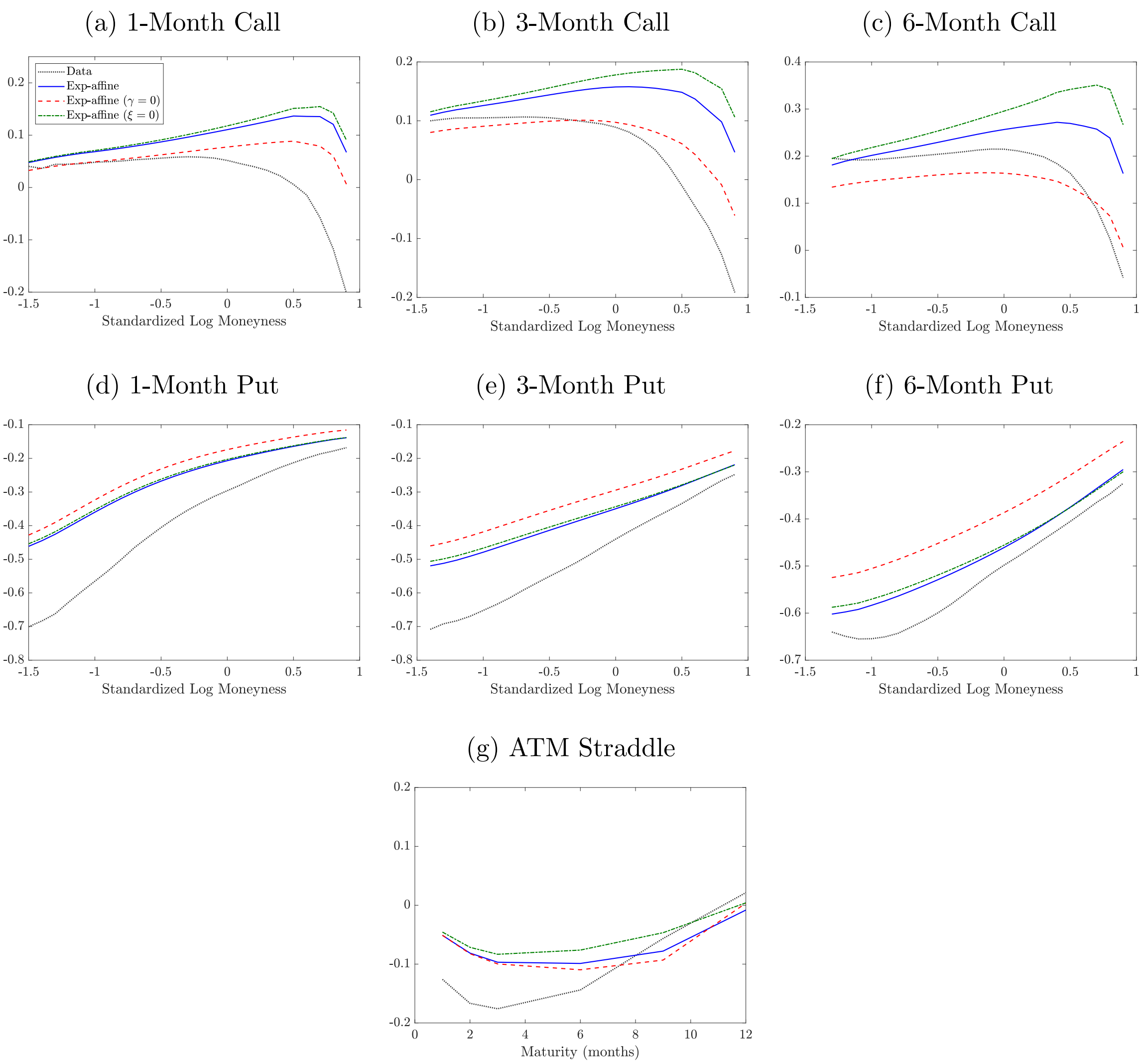


(g) Exponential-Affine with $\eta = 0$



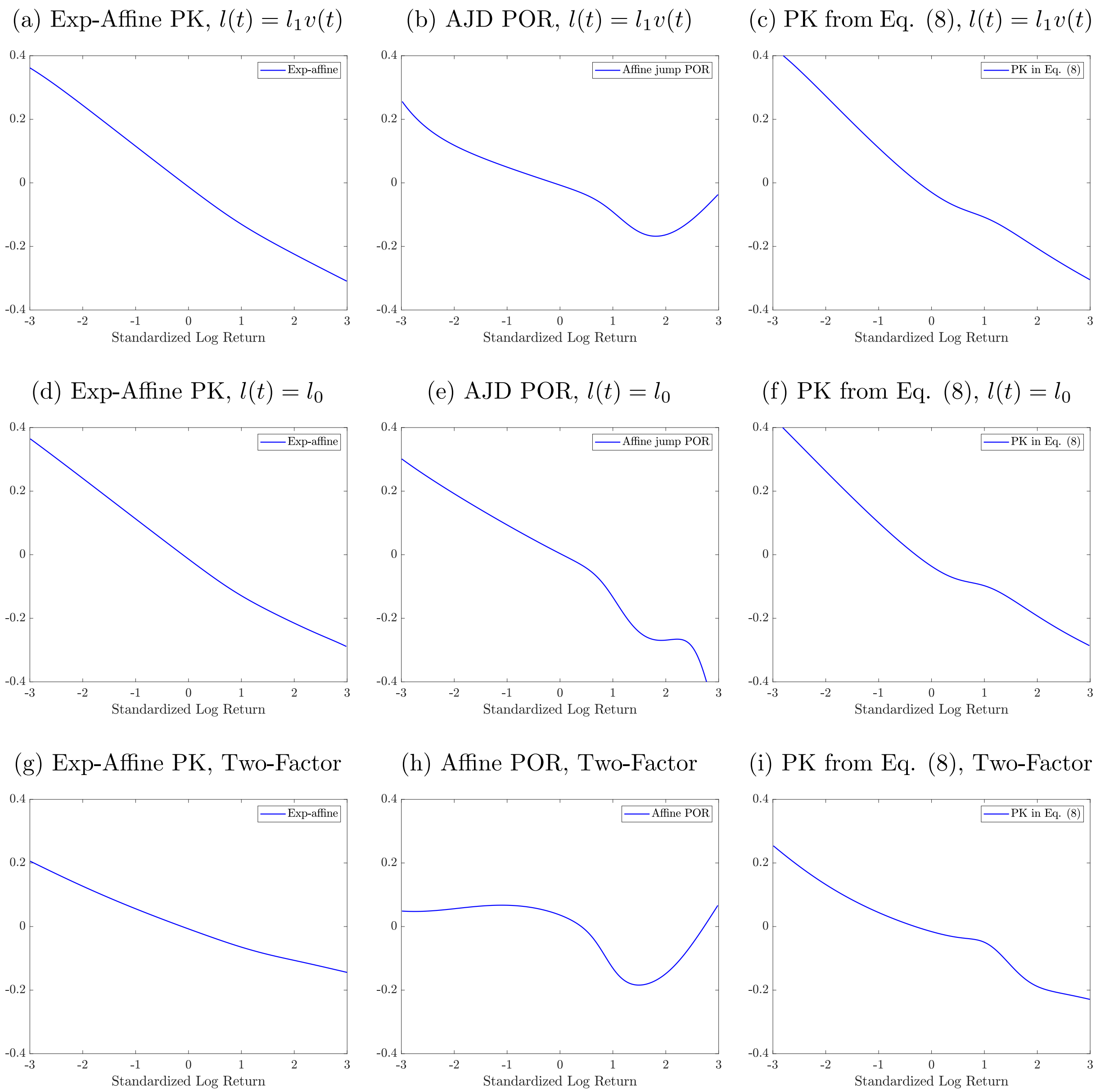
Notes: We plot the time series of the daily log pricing kernels and the implied 1-month log marginal pricing kernels for the following five specifications: the exponential-affine specifications with restriction $\gamma = 0$ (Panel A), $\xi = 0$ (Panel B), and $\eta = 0$ (Panel G), the affine price of risk with intercept with the Feller condition imposed (Panel C), the pricing kernel from equation (8) with the Feller condition imposed (Panel D), the affine price of risk specification with restriction $\mu_0 = 0$ (Panel E), and the affine price of risk specification with restriction $\lambda_0 = 0$ (Panel F). Parameter values for Panel A, B, E, F, and G are from Table 4, and those for Panels C and D are from Table 6. For the implied 1-month marginal kernels, the x-axis represents the log return in deviation from its mean and divided by its standard deviation.

Figure 6: Expected Option Returns. Restrictions on Preferences



We plot model-implied expected option returns for the exponential-affine pricing kernel and two special cases ($\gamma = 0$ and $\xi = 0$), and compare them to the data. Panels A through F plot time series averages of realized option returns and expected returns as a function of option moneyness. Panel G plots time series averages of realized ATM straddle returns and expected returns as a function of maturity.

Figure 7: Implied Log Marginal Pricing Kernels. Additional Models



(a) Exp-Affine PK, $l(t) = l_1 v(t)$

(b) AJD POR, $l(t) = l_1 v(t)$

(c) PK from Eq. (8), $l(t) = l_1 v(t)$

(d) Exp-Affine PK, $l(t) = l_0$

(e) AJD POR, $l(t) = l_0$

(f) PK from Eq. (8), $l(t) = l_0$

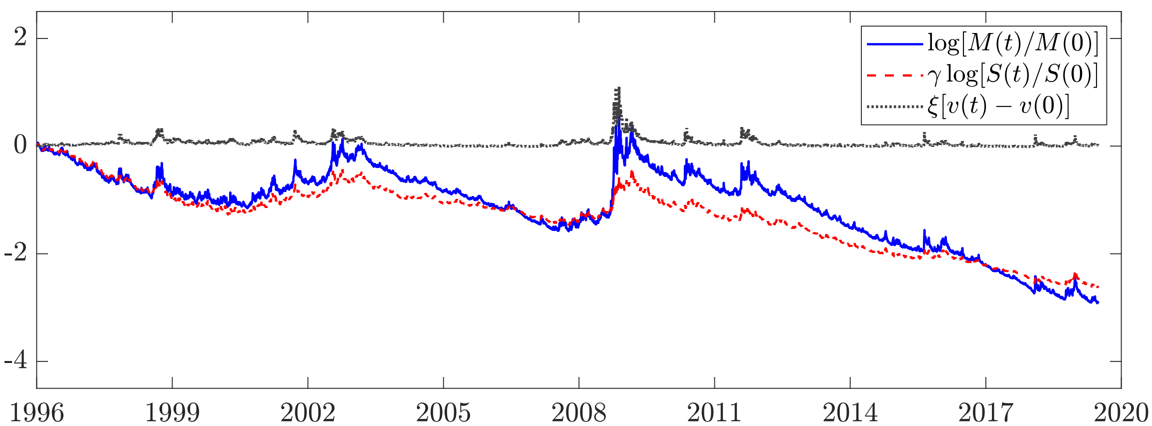
(g) Exp-Affine PK, Two-Factor

(h) Affine POR, Two-Factor

(i) PK from Eq. (8), Two-Factor

Notes: We plot the implied 1-month log marginal pricing kernels for a model with one stochastic volatility factor and jumps in returns and for a model with two stochastic volatility factors. Panels A, B, and C report on specifications with stochastic jump intensity, and Panels D, E, and F report on specifications with constant jump intensity. Panels G, H, and I report on the two-factor stochastic volatility model. For each model, we report results for the exponential-affine pricing kernel, the more general pricing kernel, and the affine POR specification.

Figure 8: 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: Additional Tables and Figures

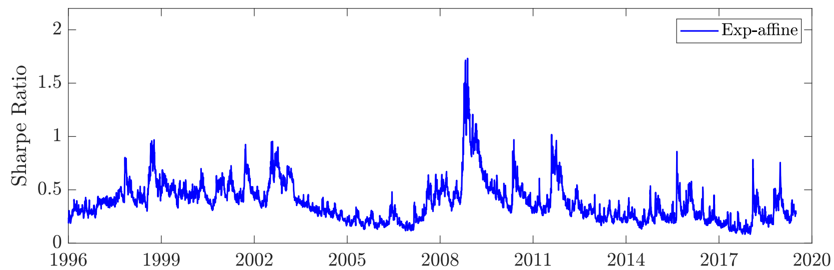
Table A.1: Joint MLE Estimation. Robustness Exercises

	Minimum Variance	Filtered Variance		
		Exp-Affine	Affine POR	PK in Eq. (8)
Restrictions	None	None	None	None
	(1)	(2)	(3)	(4)
Risk Preference Parameters				
γ		1.2631		1.3548
		(0.4011)		(0.2831)
ξ		0.9979		0.9996
		(0.2964)		(0.1905)
α				-0.0060
				(0.0033)
Implied Risk Premia Parameters				
μ_0			0.0358	-0.0040
			(0.0269)	
μ or μ_1	2.1846	1.9214	2.0523	2.0141
	(0.8685)		(0.9575)	
λ_0			0.0290	0.0041
			(0.0076)	
λ or λ_1	-1.5266	-1.5092	-1.9208	-1.5703
	(0.0961)		(0.0798)	
P-Parameters				
κ	3.0155	3.8284	4.2420	3.8892
	(0.1034)	(4.4E-08)	(5.9E-09)	(2.1E-08)
θ	0.0360	0.0323	0.0359	0.0328
	(0.0015)	(0.0007)	(0.0015)	(0.0010)
σ	0.6569	0.8230	0.8215	0.8288
	(0.0221)	(0.0111)	(0.0243)	(0.0165)
ρ	-0.7874	-0.8016	-0.8020	-0.8016
	(0.0099)	(0.0099)	(0.0106)	(0.0102)
η_0	-0.0078			
	(9.8E-06)			
η_1	0.9358			
	(0.0014)			
Q-Parameters				
κ^*	1.4890	2.3193	2.3212	2.3189
θ^*	0.0729	0.0533	0.0531	0.0532
Model Fit				
$\log L^R$	43,351.61	21,439.42	21,444.95	21,439.42
$\log L^O$	83,484.27	79,019.36	79,015.17	79,019.62
Weighted $\log L$	193,710.26	117,991.43	118,007.62	117,991.79
Vega-Weighted Option Price RMSE	0.0201	0.0229	0.0229	0.0229

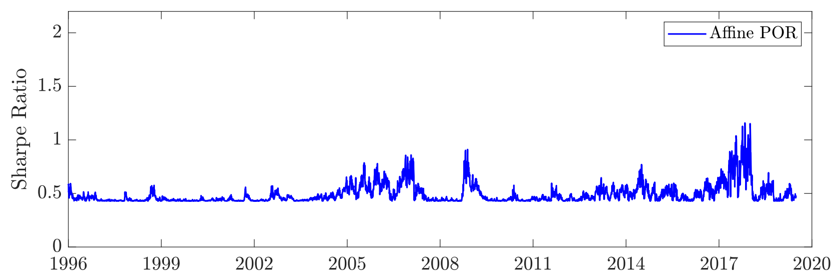
Notes: We report on joint maximum likelihood estimation using the 1996-2019 sample and either a model with a minimum variance level or a variance filtering technique. Column (1) reports on the minimum variance model, column (2) on the exponential-affine pricing kernel (PK) specification, column (3) on the affine price of risk (POR) specification, and column (4) on the more general PK from equation (8). We report the log likelihood and the vega-weighted option price RMSE for each specification. The risk preference parameters γ, ξ, and α (or the risk premia parameters μ_0, μ_1, λ_0, and λ_1 for columns (1) and (3)) as well as the P-parameters are estimated; the other parameters are implied by the respective restrictions. Robust standard errors are in parentheses.

Figure A.1: Time Series of Annualized Sharpe Ratios

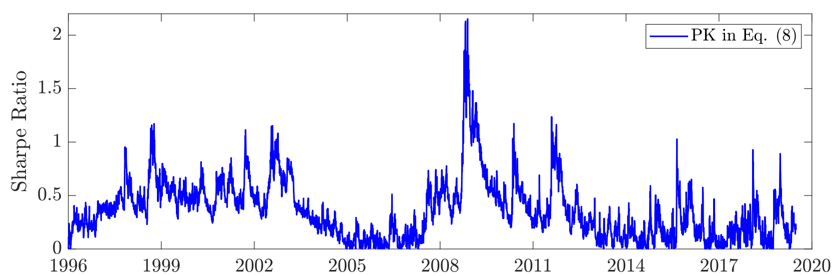
(a) Exponential-Affine PK



(b) Affine Price of Risk

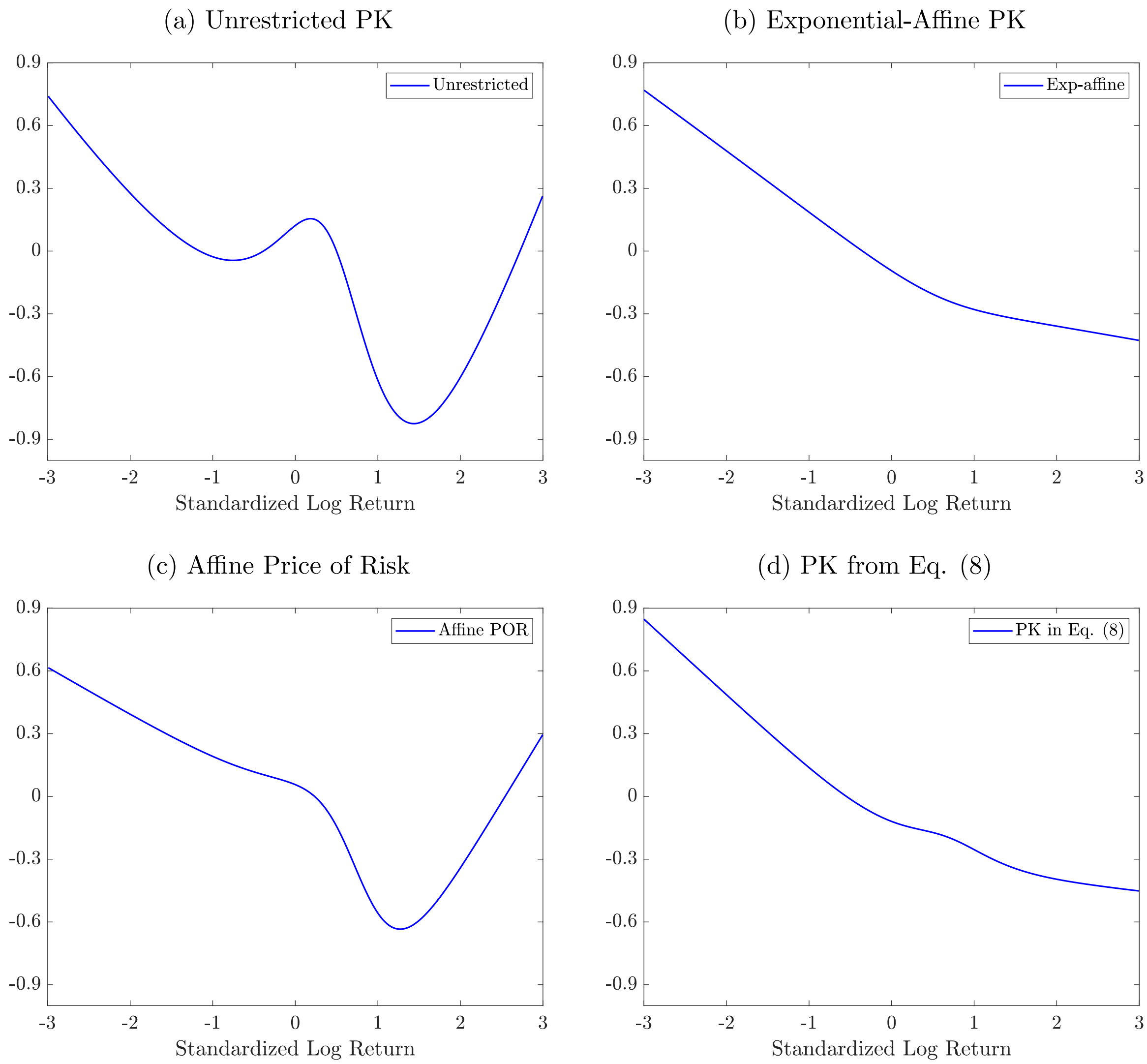


(c) PK from Eq. (8)



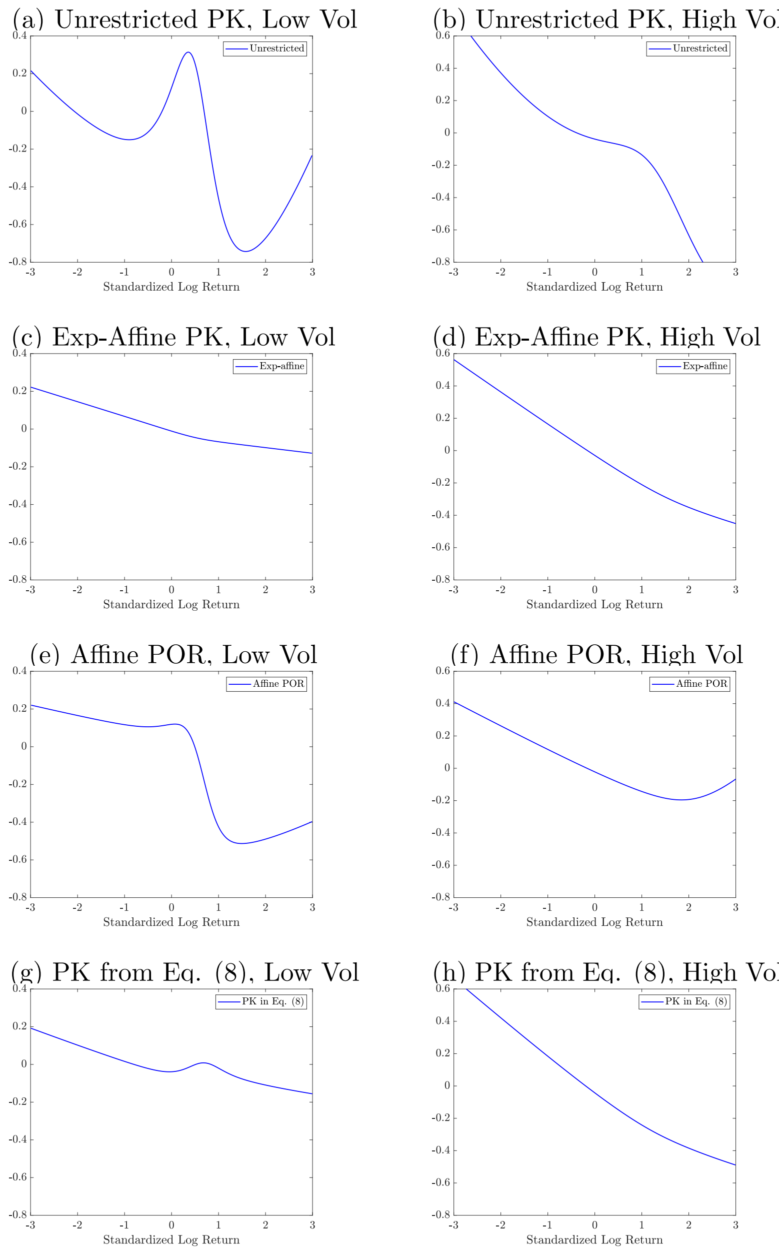
Notes: We plot the time series of the annualized Sharpe ratios for the following three specifications: the unrestricted exponential-affine pricing kernel (Panel A), the affine price of risk specification (Panel B), and the unrestricted pricing kernel from equation (8) (Panel C). Parameter values are from Table 3.

Figure A.2: Implied Log Marginal Pricing Kernels. Six-Month Maturity



Notes: We plot the implied 6-month log marginal pricing kernels for the following four specifications: independent estimation of the P- and Q-parameters (Panel A), the unrestricted exponential-affine (Panel B), the affine price of risk specification (Panel C), the unrestricted pricing kernel from equation (8) (Panel D). Parameter values for Panel A are from Table 2, and those for Panels B, C, and D are from Table 3. The x-axis represents the log return in deviation from its mean and divided by its standard deviation.

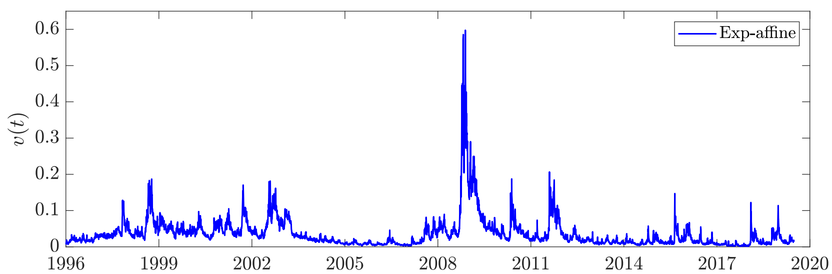
Figure A.3: Implied Log Marginal Pricing Kernels Conditional on the Variance Level



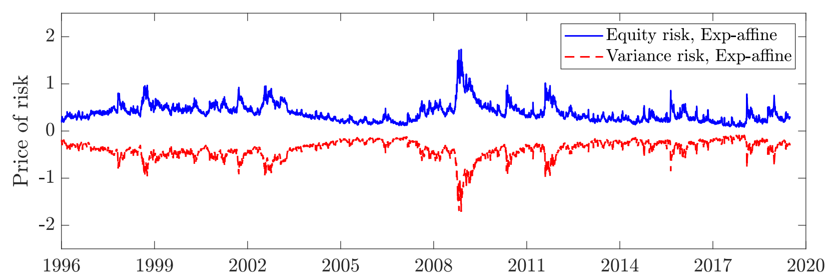
Notes: We plot the implied 1-month log marginal pricing kernels for the following four specifications: independent estimation of the P- and Q-parameters (Panels A and B), the unrestricted exponential-affine pricing kernel (Panels C and D), the affine price of risk specification (Panels E and F), and the unrestricted pricing kernel from equation (8) (Panels G and H). The left panels are conditional on low variance (conditional variance set at 0.01), while the right panels are conditional on high variance (conditional variance set at 0.1). The parameter values used for Panels A and B are from Table 2, and those used for Panels B through H are from Table 3. The x-axis represents the log return in deviation from its mean and divided by its standard deviation.

Figure A.4: Time Series of the Variance and Equity Price of Risk

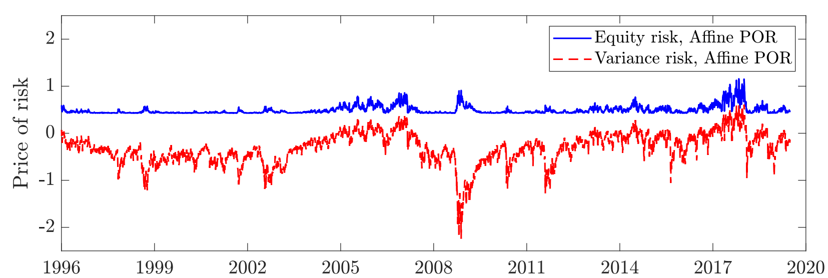
(a) Stochastic Variance, Exponential-Affine PK



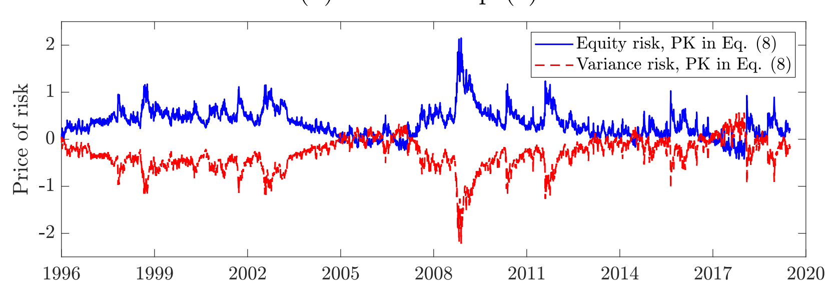
(b) Exponential-Affine PK



(c) Affine Price of Risk



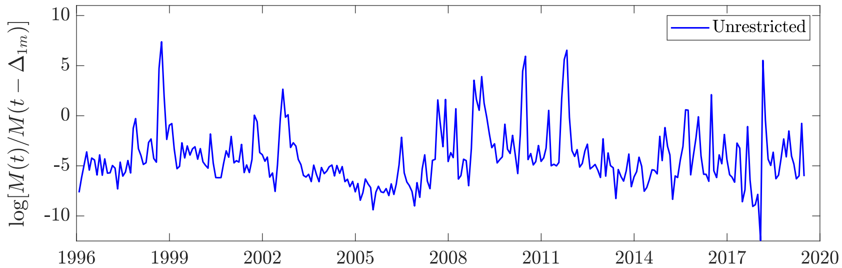
(d) PK from Eq. (8)



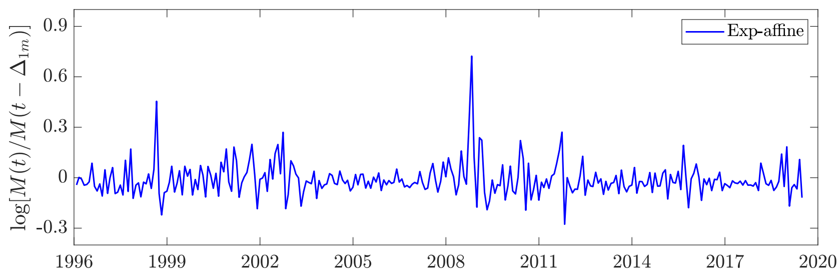
Notes: We plot the time series of the stochastic variance for the unrestricted exponential-affine (Panel A) and the prices of equity and variance risks for the following three specifications: the unrestricted exponential-affine (Panel B), the affine price of risk specification (Panel C), and the unrestricted pricing kernel from equation (8) (Panel D). Parameter values are from Table 3.

Figure A.5: Monthly Log Pricing Kernels

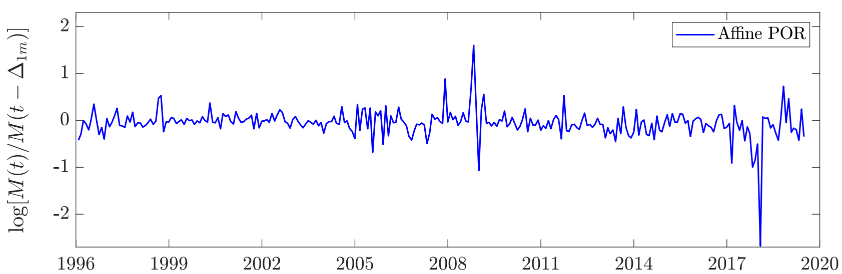
(a) Unrestricted PK



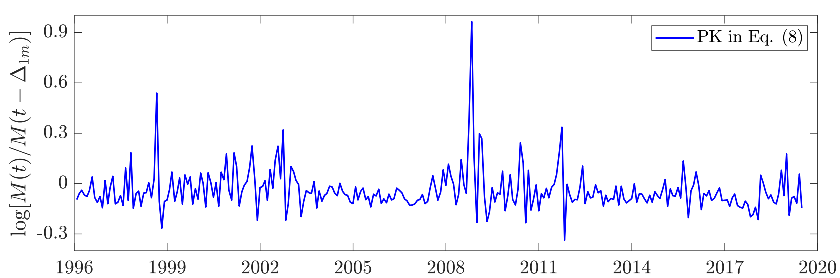
(b) Exponential-Affine PK



(c) Affine Price of Risk

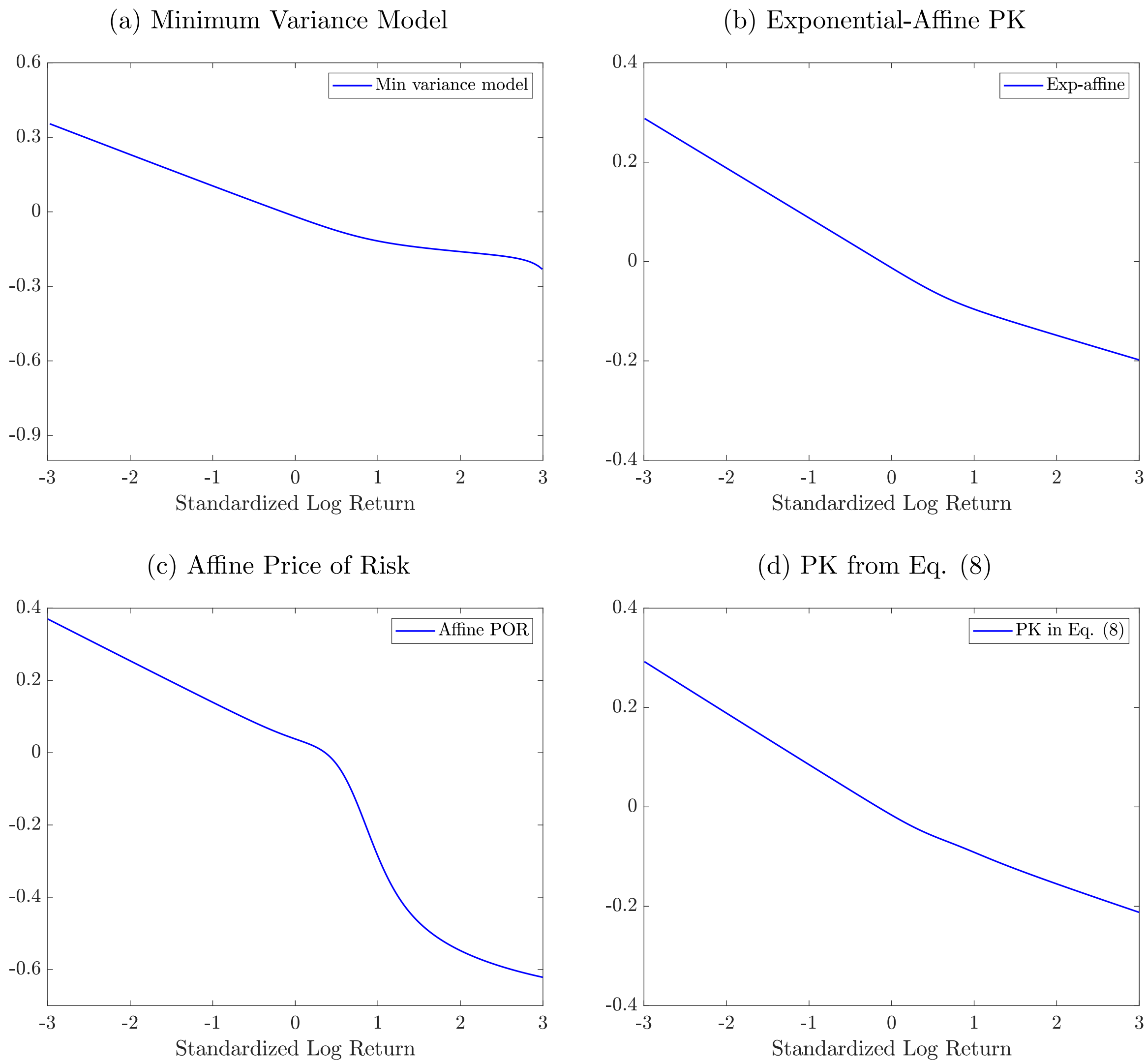


(d) PK from Eq. (8)



Notes: We plot the time series of the monthly log pricing kernels for the following four specifications: independent estimation of the P- and Q-parameters (Panel A), the unrestricted exponential-affine (Panel B), the affine price of risk specification (Panel C), the unrestricted pricing kernel from equation (8) (Panel D). The parameter values used for Panel A are from Table 2, and those used for Panels B, C, and D are from Table 3.

Figure A.6: Implied Log Marginal Pricing Kernels. Various Robustness Exercises



Notes: Panel A presents results for the option pricing model with a minimum variance level. Panels B-D plot the implied 1-month log marginal pricing kernels based on the alternative parameter estimates in Table A.1 obtained using variance filtering, for the following three specifications: the unrestricted exponential-affine (Panel B), the affine price of risk specification (Panel C), and the unrestricted pricing kernel from equation (8) (Panel D). The x-axis represents the log return in deviation from its mean and divided by its standard deviation.

Online Appendix

A The Pricing Kernel and Risk Premia

Given the pricing kernel in equation (3) or (5), we use the fact that $U(t)M(t)$ is a martingale for any asset U to calculate risk premia. Applying Ito's lemma to the martingale condition $E[d(U(t)M(t))] = 0$ shows that expected returns are determined by their covariance with the pricing kernel:

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

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

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

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

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

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

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

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

Likewise, the pricing kernel with the affine prices of risk in equation (5) is consistent with affine risk premiums $\mu_0 + \mu_1 v(t)$ and $\lambda_0 + \lambda_1 v(t)$, where

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

B No-Arbitrage Restrictions for the Pricing Kernel in Equation (8)

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

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

Applying Ito's lemma to equation (8) 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

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

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

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

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

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

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

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

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

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

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

Applying Ito's lemma to equation (12) gives

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

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

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

In order to find the restrictions on 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 (C.1),

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

By rearranging this equation,

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

Since equation (C.2) must hold for any values of v, it implies that (A) and (B) must be zero. Therefore, we have the following restrictions on 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, by equations (A.4) and (C.1), we find the expression for the equity and variance risk premium parameters μ and λ as functions of the preference parameters:

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

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

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

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

$$\kappa=\kappa^*-\lambda \quad \text{and} \quad \theta=\kappa^*\theta^*/\kappa. \tag{C.4}$$

D Parameter Restrictions in the Joint Option and Returns Likelihood

The following table summarizes the restrictions used when estimating the joint likelihoods in Tables 3 and 4. As mentioned above, for the case in column (8) there are no restrictions. We implement these cases by letting either 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.

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

E A Closed-Form Expression for the Joint Probability Distribution

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

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

where $\Delta = 1/252$. Note that $Q(\log R(t+\Delta), v(t+\Delta)|v(t))$ and $P(\log R(t+\Delta), v(t+\Delta)|v(t))$ are the risk-neutral and physical joint probability distributions of $\log R(t+\Delta)$ and $v(t+\Delta)$ conditional on $v(t)$. In this implementation, we have to use the joint probability rather than the marginal probability of the log return described in equation (27), because we need to use the realization of both $\log R(t+\Delta)$ and $v(t+\Delta)$.[40] We can also find a quasi closed form expression (up to numerically computing double integrals for the joint probability).

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

When x and v evolve according to

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

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

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

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

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

Suppose g has the following functional form:

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

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

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

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

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

and the terminal conditions of the ODEs are $G(0) = 0$ and $H(0) = i\phi_v$ inferred from equation (E.3). This system of ODEs expressed in equation (E.5) has the following closed-form solution:

$$\begin{aligned} H(\tau) &= \frac{A_2(D_m X - iA_2 Y)\phi_v - 2iA_1(-D_m - 2A_3\phi_v + (Y-1)(D_m - 2A_3\phi_v))}{iA_2 D_m X - A_2^2 Y + 4A_1A_3Y - 2A_3D_mX\phi_v}, \\ G(\tau) &= r\phi_x i\tau \\ &+ \frac{1}{4A_3}\Bigg[-2aA_2\tau - \frac{2ia\tau(A_2^2 - 4A_1A_2)}{D_m} + 2ia\arctan\left(\frac{A_2A_3X^2\phi_v}{A_2^2(Y-1) - A_1A_3Y^2 + A_3^2X^2 + \phi_v^2}\right) \\ &+ \frac{4aD_p}{D_m}\arctan\left(\frac{iA_2^2 + 2A_2A_3X\phi_v - 2iA_3(A_1Y - A_3X\phi_v^2)}{D_p(A_2 + 2iA_3\phi_v)}\right) \\ &- \frac{4iaD_p}{D_m}\operatorname{arctanh}\left(\frac{D_p}{A_2 + 2iA_3\phi_v}\right) + a\log(D_p) \\ &- a\log\left(A_2^4(Z-1) + A_1^2A_3^2Y^4 + A_2^2A_3^2X^2Z\phi_v^2 + A_3^4X^4\phi_v^4 - 2A_1A_3Y^2(A_2^2(Y-1) + A_3^2X^2\phi_v^2)\right)\Bigg], \end{aligned}$$

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,

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

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

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

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

Numerical integration of equation (E.7) 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).[41] Hence, we numerically compute equation (E.7) as

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

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

F Alternative Shock Decomposition and Restrictions on Affine Risk Premiums

We examine alternative representation of the stock and variance dynamics as follows:

$$\begin{aligned} dS(t)/S(t) &= rdt + \sqrt{1-\rho^2}\sqrt{v(t)}du_1^*(t) + \rho\sqrt{v(t)}dz_2^*(t), \\ dv(t) &= \kappa^*(\theta^* - v(t))dt + \sigma\sqrt{v(t)}dz_2^*(t), \end{aligned}$$

where $u_1^*(t)$ and $z_2^*(t)$ are independent. This representation allows us to easily study the effect of equity risk independent of variance risk or the effect of variance risk independent of equity risk

[41] 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.

under the affine price of risk specification. We consider the following changes of measure:

$$\begin{aligned} du_1^*(t) &= du_1(t) + \frac{a}{\sqrt{v(t)}}dt + b\sqrt{v(t)}dt, \\ dz_2^*(t) &= dz_2(t) + \frac{c}{\sqrt{v(t)}}dt + d\sqrt{v(t)}dt. \end{aligned}$$

Now it is straightforward to see how the affine structure of the equity risk premium (ERP) or variance risk premium (VRP) can be restricted through the above representation. Suppose $c = 0$; that is, the price of $z_2(t)$ risk is purely proportional to $\sqrt{v(t)}$. Then, the ERP and the VRP are expressed as:

$$\begin{aligned} ERP &= a\sqrt{1-\rho^2} + \left(b\sqrt{1-\rho^2} + d\rho\right) v(t), \\ VRP &= d\sigma v(t). \end{aligned}$$

The ERP has a constant term, whereas the VRP does not.

This representation is isomorphic to our main representation in equation (1), but from an empirical perspective, this shock decomposition is more transparent to highlight the source of the constant term in the risk premium. Using the notation from our paper, we can write $ERP = \mu_0 + \mu_1 v(t)$ and $VRP = \lambda_0 + \lambda_1 v(t)$. In the above representation,

$$\mu_0 = a\sqrt{1-\rho^2}, \quad \mu_1 = b\sqrt{1-\rho^2} + d\rho, \quad \lambda_0 = 0, \quad \lambda_1 = d\sigma.$$

With the isomorphic characteristic, we can relate this result to our original representation as follows:

$$\pi_{1,0} = \frac{a}{\sqrt{1-\rho^2}}, \quad \pi_{2,0} = -\frac{a\rho}{\sqrt{1-\rho^2}}, \quad \pi_{1,1} = \frac{b}{\sqrt{1-\rho^2}}, \quad \pi_{2,1} = d - \frac{b\rho}{\sqrt{1-\rho^2}}.$$

G An Option Pricing Model with a Variance Floor

We additionally consider a specification in which the stock return variance has a floor, v_0. We express the variance as $v_0 + v(t)$, and a state variable $v(t)$ follows a square-root process as below:

$$\begin{aligned} dS(t)/S(t) &= rdt + \sqrt{v_0 + v(t)}dz_1^*(t), \\ dv(t) &= \kappa^*(\theta^* - v(t))dt + \sigma\sqrt{v(t)}dz_2^*(t). \end{aligned} \tag{G.1}$$

In order to preserve affine structure, we assume that $Cov\left(\frac{dS(t)}{S(t)}, dv(t)\right) = \rho\sigma v(t)dt$. Furthermore, in this model, we only consider a specification, where the price of risk is proportional to $\sqrt{v_0 + v(t)}$ or $\sqrt{v(t)}$ and does not contain a term inversely proportional to $\sqrt{v(t)}$. The resulting change of measure is represented as

$$\begin{aligned} dz_1^*(t) &= dz_1(t) + \mu\sqrt{v_0 + v(t)}dt, \\ dz_2^*(t) &= dz_2(t) + \lambda\sqrt{v(t)}dt, \end{aligned}$$

and we obtain the following physical dynamics:

$$\begin{aligned} dS(t)/S(t) &= \left[r + \mu(v_0 + v(t))\right]dt + \sqrt{v_0 + v(t)}dz_1(t), \\ dv(t) &= \kappa(\theta - v(t))dt + \sigma\sqrt{v(t)}dz_2(t), \end{aligned} \tag{G.2}$$

where κ and θ are given as in equation (C.4). Ito's lemma also implies the following dynamic for $\log S(t)$:

$$d\log S(t) = \left[r + \left(\mu - \frac{1}{2}\right)(v_0 + v(t))\right]dt + \sqrt{v_0 + v(t)}dz_1(t).$$

For pricing options, we should have the characteristic function. Suppose the characteristic function has a form of

$$h_\tau^{CH}(\phi|S(t), v(t)) = e^{E + Dv(t) + i\phi\log S(t)}. \tag{G.3}$$

Following Heston (1993), we obtain

$$E = C - \frac{v_0}{2}(\phi^2 - \phi)\tau,$$

where the solutions of C and D are given by equation (25).

In implementation, the minimum variance parameter v_0 is arbitrarily set to 0.0006 because the historically lowest value of the annualized 5-minute realized variance is 0.000625 (or 0.025 in volatility). The parameter estimates are reported in column (1) of Table A.1, and the implied marginal pricing kernel is depicted in Panel A of Figure A.6. The weighted $\log L$ is slightly higher than those of our main specifications in Table 3, but the vega-weighted option price RMSE is not distinguishable.

Under this new model, the equity premium is given by $\mu v_0 + \mu v(t)$. For comparison, recall that Panel F of Figure 5 corresponds to the model specification where the equity premium is $\mu_0 + \mu_1 v(t)$ and the variance premium is $\lambda_1 v(t)$, and its implied marginal pricing kernel is not economically intuitive. In contrast, Panel A of Figure A.6 presents a more economically intuitive pricing kernel. The difference primarily arises from the fact that this minimum-variance model does not require a price of risk inversely proportional to $v(t)$. This result confirms that what really matters is the form of the pricing kernel or prices of risk but not the mere existence of a constant term in risk premia itself.