Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Characterizing the Conditional Pricing Kernel: A New Approach

Hyung Joo Kim

2026-059

Please cite this paper as:
Kim, Hyung Joo (2026). "Characterizing the Conditional Pricing Kernel: A New Approach," Finance and Economics Discussion Series 2026-059. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2026.059.

Characterizing the Conditional Pricing Kernel: A New Approach*

Hyung Joo Kim†

Abstract

I propose a novel method to reliably estimate the conditional pricing kernel by incorporating conditioning variables. The VIX and the term spread are most informative variables for identifying state prices. The conditional kernel estimate exhibits significant time variation: the more favorable market expectations, the higher state prices in negative return states. During bad times, the equity premium implied by the conditional kernel is fully attributable to compensation for left-tail scenarios, in contrast to findings from the unconditional kernel. Lastly, the conditional kernel estimate yields superior out-of-sample option pricing performance compared to the unconditional kernel estimate.

*I would like to thank Rohit Allena, Fousseni Chabi-Yo, Hitesh Doshi, Xiang Fang, Gustavo Freire, Antonio Gargano, Mohammad Ghaderi, Sangeun Ha, Jimmy Hilliard, Kris Jacobs, Seung Joo Lee, Dong Hwan Oh, Paola Pederzoli, Sang Byung Seo, Rauli Susmel, and seminar participants at the New Zealand Finance Meeting, the EEA-ESEM Conference, the IAAE Annual Conference, the Virtual Derivatives Workshop, Concordia University, Cornerstone Research, Federal Reserve Board of Governors, Korea Development Institute, Korea Insurance Research Institute, Saint Mary's University, University of Houston, and Wilfrid Laurier University for valuable comments and advice. All OptionMetrics via WRDS and Datastream data were obtained by the author prior to employment at the Federal Reserve Board, while the author was at the University of Houston. The analysis and conclusions set forth are those of the author and do not indicate concurrence by the Federal Reserve Board or other members of its staff.

†Federal Reserve Board; hyungjoo.kim@frb.gov

1 Introduction

The pricing kernel (or stochastic discount factor) is a central concept in financial economics, and its properties affect the pricing of all financial assets. Starting with Aït-Sahalia and Lo (2000) and Jackwerth (2000), the literature has extensively documented the properties of an empirical pricing kernel, defined as the ratio of the risk-neutral density to the physical density of market returns, discounted by the risk-free rate. Most existing studies focus on investigating unconditional pricing kernel estimates or the time-series average of conditional pricing kernel estimates, but relatively little is known about the conditional pricing kernel. Reliably estimating the conditional pricing kernel is challenging due to a lack of data at each point in time. Despite these well-known challenges, prior studies have emphasized the importance of the conditional pricing kernel, as investors' conditional information and beliefs play a crucial role in asset pricing (see, e.g., Harvey, 1989; Nagel and Singleton, 2011).

In this paper, I propose a method to robustly estimate the conditional pricing kernel based on forward-looking measures without strong parametric model assumptions. I extend the method of Linn, Shive, and Shumway (2018), who estimate an unconditional pricing kernel, by incorporating conditioning variables that are informative about the kernel. This method allows the kernel to vary over time in response to changing economic states. As a result, the conditional pricing kernel estimate enables the exploration of its time-series characteristics, implications for conditional risk premia, and out-of-sample asset pricing tests within a unified framework.

I begin the empirical investigation with 12 candidate conditioning variables that are potentially informative about the kernel. These variables include four measures of risk: the VIX, model-free risk-neutral skewness and kurtosis of Bakshi, Kapadia, and Madan (2003), and the left-tail risk index of Bollerslev, Todorov, and Xu (2015). I also use six economic indicators: the term spread, Chicago Fed National Activity Index (CFNAI), Aruoba, Diebold, and Scotti (ADS, 2009) index, industrial production growth, real consumption growth, and credit spread. The remaining two variables are investor sentiment and the lagged stock market return. Since estimating the conditional pricing kernel using all 12 conditioning variables simultaneously is impractical, I first identify the most informative variables by estimating the kernel with each variable separately. Statistical tests reveal that the VIX, sentiment, term spread, industrial production growth, credit spread, and market return are informative. I then re-estimate the conditional pricing kernel using multiple informative variables. I document the (average) shape of the conditional pricing kernel estimate, investigate its time variation, and explore the kernel's economic properties and implications.

First, I find that the monotonicity of the average conditional kernel estimate is not statistically rejected up to 8% of the monthly market return. A large body of literature on empirical pricing kernels discusses the pricing kernel puzzle, where empirical pricing kernels appear U-shaped in market returns, contradicting the traditional expectation of a downward-sloping kernel.[1,2] However, Linn, Shive, and Shumway (2018) and Barone-Adesi, Fusari, Mira, and Sala (2020) show that the puzzle disappears when both risk-neutral and physical densities are forward-looking. The pricing kernel estimates in this paper are also forward-looking. I confirm that the monotonicity of the unconditional kernel estimate is not rejected, consistent with the traditional intuition, but I observe mixed evidence for the conditional kernel estimate. The monotonicity of its time-series average holds only up to 8% of the market return and is rejected in deep right tails. Turning to the conditional kernel estimate at each point in time, a noticeable U-shaped pattern is also observed depending on market conditions.

Second, I find significant time variation in the shape of the conditional pricing kernel estimate, especially in the left tail of the market return distribution. The results suggest that investors' beliefs about negative future market returns are more sensitive to conditional information than beliefs about positive returns. Moreover, the pricing kernel estimate in the negative return region is substantially lower when a conditioning variable signals negative market outlooks than positive

[1] Empirical pricing kernel generally refers to an estimated pricing kernel projected on market returns. I use the terms empirical pricing kernel and pricing kernel estimate interchangeably.

[2] See, for instance, Bakshi, Madan, and Panayotov (2010), Chabi-Yo (2012), Christoffersen, Heston, and Jacobs (2013), Hens and Reichlin (2013), Song and Xiu (2016), Cuesdeanu and Jackwerth (2018b), Almeida and Freire (2022), and Driessen, Koëter, and Wilms (2025) for more extensive discussion of the pricing kernel puzzle.

outlooks. This indicates that for negative future returns, investors are less patient to consume when they have unfavorable information about the financial market or economy, or that investors are more fearful of negative future returns with favorable information than with unfavorable information.

My estimates of the time series of the risk-neutral densities, physical densities, and pricing kernels enable me to study additional economic implications. I calculate conditional risk premia implied by the conditional pricing kernel estimate. I compare these conditional measures with those implied by an unconditional pricing kernel estimate that does not specify any conditioning variables. The equity premium is better captured when conditional information is incorporated, while higher-moment risk premia are similar across the two approaches.

Third, I contribute to the literature on conditional asset pricing by analyzing the time-varying sources of the conditional equity premium, inspired by Beason and Schreindorfer (2022) and Chabi-Yo and Loudis (2022).[3] During bad times, my unconditional pricing kernel estimate partially captures a contribution of the left tail of the market return distribution to the conditional equity premium, compared to the conditional pricing kernel estimate, which is able to capture a significant contribution. This analysis highlights the value of conditional asset pricing and the conditional pricing kernel.

Fourth, I compare the out-of-sample performance of option pricing using the conditional kernel estimate versus the unconditional one. By the fundamental theorem of asset pricing, all financial assets can be priced once the pricing kernel is known. I test whether the Euler equation holds for option returns and delta-hedged option portfolio returns and find that the conditional kernel estimate provides superior out-of-sample performance. This result partially stems from better capturing the state prices for deep tails conditionally, or better matching out-of-the-money (OTM) call and put prices.[4]

[3]These studies examine the return ranges that contribute most to the equity premium. In this paper, I distinguish the economic implications of the conditional and unconditional pricing kernel estimates by investigating the sources of the conditional equity premium.

[4]Bakshi, Madan, and Panayotov (2010) and Bondarenko (2014) discuss overpriced OTM call and put prices (or those negative returns), respectively.

For implementation, my empirical method builds on a statistical property of cumulative distribution functions (CDFs) known as the universality of the uniform, where a CDF follows the uniform distribution between zero and one. The conditional physical density can be recovered using risk-neutral density estimates and a given pricing kernel. The realized physical CDFs should unconditionally satisfy uniform moment conditions. Linn, Shive, and Shumway (2018) use a generalized method of moments (GMM) based on these moments to estimate an unconditional pricing kernel. They refer to the approach as the conditional density integration (CDI) method.

My adaptation of the CDI method addresses two key criticisms in the literature on estimating the projected pricing kernel on market returns. First, while the risk-neutral density can be reliably inferred from option prices, the physical density is typically obtained from historical data or through a parametric model.[5] This method creates a timing mismatch: the historical physical density is backward-looking, whereas the risk-neutral density is forward-looking. Second, both the historical-data- and model-based conditional pricing kernel estimates highly depend on the estimation specification. The CDI method avoids both issues because it uses options data to find the physical density and is independent of model assumptions about the return dynamics.

I improve the CDI method by adding a constraint to satisfy a property of a probability density: the integral of the physical density over its domain must equal one. Despite the importance of this fundamental property for conditional physical density, this property is not enforced in the original CDI method. If violated, the core relation among the pricing kernel, risk-neutral density, and physical density breaks down. Hence, I incorporate the constraint by minimizing the error between the integrated value and one.[6] This improvement also allows me to exploit the estimated densities in additional analyses, such as studying the conditional risk premia.

[5] See, for example, Aït-Sahalia and Lo (2000), Jackwerth (2000), Rosenberg and Engle (2002), Chabi-Yo, Garcia, and Renault (2007), Barone-Adesi, Engle, and Mancini (2008), Chabi-Yo (2012), Christoffersen, Heston, and Jacobs (2013), and Cuesdeanu and Jackwerth (2018b).

[6] Cuesdeanu and Jackwerth (2018a) address this issue by calculating the physical density as the ratio of the risk-neutral density to the pricing kernel and then normalizing the physical density by its integrated value. However, if this normalization is not simultaneously accounted for during estimation, the normalized physical density may not equal the ratio of the risk-neutral density to the pricing kernel.

I further modify the CDI method by using observable economic or financial variables as conditional information. To this end, I approximate the pricing kernel using a polynomial function, which offers both flexibility and tractability. I also augment the GMM moment conditions by including the Euler equations based on the market return and risk-free rate. The conditional pricing kernel is then estimated via two-step GMM.

This paper empirically contributes to the estimation of the forward-looking conditional pricing kernel. Linn, Shive, and Shumway (2018) estimate the pricing kernel using the VIX as a GMM instrumental variable or estimate it for two subperiods (high- and low-VIX periods) and refer to these as conditional estimates. My definition of the conditional pricing kernel differs in that I estimate a fully time-varying pricing kernel. Barone-Adesi, Fusari, Mira, and Sala (2020) also estimate a time-varying kernel but rely on parametric return dynamics, which are then adjusted using options data.[7] In contrast, my approach does not assume parametric retun dynamics.

Closely related to this paper, a contemporaneous study by Schreindorfer and Sichert (2025) estimates the conditional pricing kernel by using the VIX as a conditioning variable via maximum likelihood estimation and finds similar time-series patterns of the kernel estimate.[8] While their focus is on equilibrium models that can explain the observed patterns, my focus is on estimating the pricing kernel with multiple conditioning variables and examining the economic implications (e.g., conditional risk premia) of the time variation in the pricing kernel.

Lastly, this paper contributes to the literature on conditional risk premia. My conditional risk premia estimates are especially important because they reflect investors' forward-looking expectations, which are not adequately reflected in historical risk premia. Prior studies calculate the ex-ante conditional equity premium by imposing model restrictions or exploiting dividends/earnings information (see, e.g., Pástor and Stambaugh, 2001; Fama and French, 2002; Donaldson, Kamstra, and Kramer, 2010; Duan and Zhang, 2014). The conditional variance risk premium has also been

[7] Other related examples include Gagliardini, Gourieroux, and Renault (2011) and Polkovnichenko and Zhao (2013).

[8] Similarly, Schneider (2022) uses a simple variance as a conditioning variable in studying the stochastic discount factor in the context of Ross (2015) recovery

widely studied.[9] In this paper, I provide a tractable and semi-nonparametric approach for computing conditional risk premia of any return moment, by estimating both the risk-neutral and physical conditional return distributions.

The rest of the paper proceeds as follows. Section 2 reviews existing methods for estimating the conditional pricing kernel and their limitations. Section 3 describes the data, and Section 4 details the proposed estimation procedure. Section 5 presents the estimation results and discusses their implications and out-of-sample analysis. Section 6 concludes.

2 Estimating the Pricing Kernel: Existing Approaches

The fundamental theorem of asset pricing states that the no-arbitrage condition is equivalent to the existence of a risk-neutral density (or state price density). This statement is equivalent to the existence of a nonnegative pricing kernel, which is the state price per unit physical density discounted by the risk-free rate. Asset prices are then determined by the expectation of future payoffs multiplied by the pricing kernel under the physical probability measure. Economically, the pricing kernel conveys information about investors' beliefs about future states. Therefore, knowing the pricing kernel as precisely as possible is vital to better understand the economy and financial markets.

How can the pricing kernel be characterized? By definition, one can directly calculate the pricing kernel provided that both the risk-neutral and physical probability densities are known. On the other hand, in consumption-based equilibrium models, the pricing kernel is expressed as a function of consumption growth and other risk factors (see, e.g., Rubinstein, 1976; Campbell and Cochrane, 1999; Bansal and Yaron, 2004; Wachter, 2013). Such models are usually calibrated to match the stock market data or other financial market data because it is impossible to find the exact pricing kernel via its definition without fully specified risk-neutral and physical joint densities

[9] See, among others, Bollerslev, Tauchen, and Zhou (2009), Carr and Wu (2009), Todorov (2010), Bollerslev, Gibson, and Zhou (2011), Andersen, Fusari, and Todorov (2015), and Aït-Sahalia, Karaman, and Mancini (2020).

of the economy's state variables. Therefore, prior studies often focus on an empirical pricing kernel projected on the market return measured by the S&P 500 index return and estimate the kernel directly through its definition.

2.1 The Projected Pricing Kernel

To compute the pricing kernel projected on market returns, the risk-neutral density is extracted from index options data following Breeden and Litzenberger (1978), and the physical density can be estimated using the index return data. Moreover, if the empirical objective is to analyze the aggregate stock market, studying the projected pricing kernel is innocuous, as the orthogonal component of the pricing kernel is uncorrelated with the return on the aggregate market. Even if one is interested in studying the aggregate wealth in the economy, the use of the projected pricing kernel is still valid under the assumption that the stock market is assumed to be perfectly correlated with wealth.[10]

More formally, the projected pricing kernel ($M_{t,t+\tau}$) between time t and $t+\tau$ is defined as

$$M_{t,t+\tau}(r_{t+\tau}) = \frac{1}{1+r_{f,t}\tau}\frac{f_t^*(r_{t+\tau})}{f_t(r_{t+\tau})}, \tag{1}$$

where τ is the investment horizon, $r_{f,t}$ is the annualized simple risk-free rate, $r_{t+\tau}$ is the monthly market return between time t and $t+\tau$, and f_t^* and f_t are the time-t conditional risk-neutral and physical densities of $r_{t+\tau}$. Following Breeden and Litzenberger (1978), the risk-neutral density, f_t^*, can be obtained using:

$$f_t^*(r_{t+\tau}=r) = (1+r_{f,t}\tau)\, S_t \frac{\partial^2 Call_t(K,\tau)}{\partial K^2}\bigg|_{K=(1+r)S_t}, \tag{2}$$

where S_t is the index value at time t, and $Call_t(K,\tau)$ indicates the index call option price with strike price K and time to maturity τ at time t.[11] In particular, I focus on the 30-day horizon,

[10] For a more detailed discussion, see Barone-Adesi, Fusari, Mira, and Sala (2020).

[11] Note that option strike prices are based on ex-dividend prices. This fact implies that the risk-neutral

i.e., $\tau = 30$ days. Since 30-day maturity options are not always observed at each time, I interpolate/extrapolate the 30-day Black-Scholes implied volatilities and then calculate the corresponding call option prices.[12] Using the call option prices over a wide range of strike prices, one can non-parametrically construct the conditional risk-neutral density by calculating the second derivative in equation (2). Appendix A provides more detailed empirical steps.

In contrast to the risk-neutral density estimation, there are a variety of ways to estimate the physical density. A simple approach is to use the non-overlapping τ-horizon historical return data and to apply a kernel density estimation, as in Aït-Sahalia and Lo (2000) and Jackwerth (2000). Another method is to use a parametric return dynamic. After the return dynamic is estimated using the realized market return data, the physical distribution is determined by the parameters of the dynamics and the shocks to market returns. A common assumption is to use the GARCH process (see, e.g., Rosenberg and Engle, 2002; Barone-Adesi, Engle, and Mancini, 2008; Christoffersen, Heston, and Jacobs, 2013).

These approaches to find the physical density have, however, been criticized. A criticism of using the historical return data is that such estimated physical density is backward-looking, while the risk-neutral density is forward-looking. The timing of the two probability measures is therefore subject to a mismatch. Since the pricing kernel represents what investors believe about future states, the two probability densities should be based on concurrent information, which ideally should be forward-looking. A criticism of using a parametric return dynamic is that the result depends on the model assumption and estimation specification. If the conditional physical density is considered, model dependence becomes a more serious issue. Aït-Sahalia and Lo (2000) also argue that the behavior of financial data is often poorly captured by prevalent parametric return dynamics.

---

density f_t^* is a function of the ex-dividend market return, and thus the pricing kernel $M_{t,t+\tau}$ is.

[12] Implied volatilities are obtained from OptionMetrics, and they are estimated by the Black and Scholes (1973) formula. I provide a detailed description of the data in Section 3.

2.2 Existing Estimates of the Conditional Pricing Kernel

To illustrate the shortcomings of existing approaches, I consider four different implementations of conditional physical densities: 1) a rolling estimation with the historical return data following Jackwerth (2000); 2) a parametric estimation with the Heston and Nandi (2000) GARCH model; 3) a parametric estimation with the Heston (1993) stochastic volatility (SV) model; and 4) a parametric estimation with the stochastic volatility and jump (SVJ) model of Bates (2000) and Pan (2002). Appendix B describes the dynamics and estimation procedure for all four cases in more detail. Using the conditional risk-neutral density calculated by equation (2) and the above conditional physical densities, I obtain pricing kernel estimates based on equation (1).

Figure 1 depicts the average conditional pricing kernel estimates for these four specifications. I estimate the conditional pricing kernel in every month and average over the sample. When the physical density is estimated via rolling-window historical return data (Panel A) or using the GARCH dynamic (Panel B), their empirical pricing kernels are, on average, U-shaped. When it is estimated using the SV model (Panel C), however, the increasing pattern in the deep right tail is not as pronounced as the patterns in Panels A and B. Finally, when the model also includes a Poisson jump (Panel D), the average kernel estimate is nonmonotonic but not U-shaped.

Figure 2 depicts the four estimates of the conditional pricing kernel in four specific months. Panel A (September 2002) and Panel C (September 2014) represent normal periods, while Panel B (September 2008) and Panel D (March 2020) are examples of recession periods. The former is around the time of the Lehman Brothers' collapse, and the latter is at the peak of the COVID-19 pandemic crisis. The estimates obtained using the four different methods widely differ in each of the four months.

In sum, Figures 1 and 2 demonstrate that the conditional pricing kernel estimates deliver different implications across the above four approaches. Closer scrutiny shows that the shapes of the conditional pricing kernel estimates obtained by these approaches are distinctive; that is, the conditional beliefs or expectations of investors are highly reliant on the estimation specifications or model assumptions. Therefore, it is problematic to study conditional implications using the existing approaches as the best specification is unknown. This motivates my approach, which characterizes the forward-looking conditional pricing kernel without using specific assumptions on the return dynamics.

3 Data

All variables are sampled at the monthly frequency at the end of each month, and the sample period is from January 1996 to December 2020.

To construct the risk-neutral density (hence, the pricing kernel), I use out-of-the-money (OTM) S&P 500 call and put options from OptionMetrics. I retain options with maturities between 5 and 365 days and apply the following filters:

1. Discard options with volume or open interest less than five contracts.
2. Discard options with bid price less than $3/8.
3. Discard options with data errors — options for which the bid price exceeds offer price, or where a negative price is implied through put-call parity.

I collect the S&P 500 index and its monthly returns data from CRSP and the continuously compounded risk-free rate at various maturities from OptionMetrics. I also obtain the S&P 500 cum-dividend returns data from Bloomberg. Since the analysis is based on the 30-day horizon or maturity, I interpolate the 30-day risk-free rate from the data at the end of each month.

I consider 12 different conditioning variables in my estimation of the conditional pricing kernel. First, risk measures may contain essential conditional information because investors' decisions likely depend on the level of risk. A large body of literature uses time-varying conditional volatility or future volatility as an important state variable to describe characteristics of the pricing kernel (see, e.g., Chabi-Yo, 2012; Christoffersen, Heston, and Jacobs, 2013; Song and Xiu, 2016). The VIX is a good proxy for measuring conditional volatility. I obtain the VIX data from the Federal Reserve Bank of St. Louis.[13] As argued by Chabi-Yo (2012), higher-moment risks of future returns are other

[13]Source: Chicago Board Options Exchange, CBOE Volatility Index: VIX [VIXCLS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/VIXCLS

potential factors impacting the pricing kernel. Since such risks are persistent, currently observable moment risks are also likely to be informative. I use the model-free risk-neutral skewness and kurtosis of Bakshi, Kapadia, and Madan (2003). As investors are more sensitive to and fearful of left-tail risks, a left-tail risk factor may be more informative. I additionally use the left-tail risk index of Bollerslev, Todorov, and Xu (2015), obtained from Viktor Todorov's website.

Besides market and economic risk, investor sentiment also represents crucial conditional information to characterize the pricing kernel, as argued by Shefrin (2001) and Han (2008). For the sentiment measure, I use the Investors Intelligence bull/bear ratio survey data, which I obtain from Datastream. I also include various economic indicators. I obtain the 10-year minus 2-year Treasury term spread, a leading economic indicator, from the Federal Reserve Bank of St. Louis. For economic growth proxies, I obtain the CFNAI from the Federal Reserve Bank of Chicago, the ADS business conditions index from the Federal Reserve Bank of Philadelphia, and industrial production and real consumption per capita growth from the Federal Reserve Bank of St. Louis. The credit spread is defined as the difference between Moody's Baa corporate bond yield and Aaa corporate bond yield, downloaded from the Federal Reserve Bank of St. Louis.[14] Lastly, Chernov (2003) argues that the lagged market return contains important conditional information; hence, I include the lagged monthly S&P 500 return. Table 1 reports descriptive statistics for the conditioning variables.

4 Estimating the Conditional Pricing Kernel: A New Approach

In this section, I explain how I estimate the forward-looking conditional pricing kernel. I first describe the conditional density integration (CDI) method of Linn, Shive, and Shumway (2018), which is designed to estimate the unconditional pricing kernel using GMM. I then explain how I adapt the CDI method by incorporating the conditioning variables and applying additional mod-

[14]Source: Moody's, Moody's Seasoned Aaa and Baa Corporate Bond Yields [AAA, BAA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/AAA and https://fred.stlouisfed.org/series/BAA

ifications. I use conditioning variables because a GMM estimation with the conditional moments at each point in time is challenging, given that only one market return is realized. I refer to my approach as the modified CDI method.

4.1 The CDI Method and the Use of Conditioning Variables

To facilitate the exposition, I define a scaled pricing kernel as

$$m_{t,t+\tau}(r_{t+\tau}) = \frac{f_t^*(r_{t+\tau})}{f_t(r_{t+\tau})}. \tag{3}$$

That is, the scaled pricing kernel is the pricing kernel divided by the risk-free discount factor. I make the relatively general assumption that the scaled pricing kernel is a nonlinear function of future market returns, $r_{t+\tau}$. Given the scaled pricing kernel, the conditional physical density, f_t, is then represented as $f_t^*/m_{t,t+\tau}$.

The fundamental idea of the CDI method is to use the statistical property that the conditional physical CDF, $F_t(r) = \int_{-\infty}^{r} f_t(r_{t+\tau})dr_{t+\tau}$, follows a uniform distribution between zero and one. For each $F_t(r)$, the conditional k-th moment should be $1/(1+k)$ by the properties of the uniform distribution. Then, by the law of iterated expectation, the unconditional k-th statistical moment of the CDF is written as

$$E\left[\left(\int_{-\infty}^{r} \frac{f_t^*(r_{t+\tau})}{m_{t,t+\tau}(r_{t+\tau})} dr_{t+\tau}\right)^k\right] = \frac{1}{1+k}, \tag{4}$$

for $k \geq 1$, and Linn, Shive, and Shumway (2018) use the cubic B-spline of order $q \geq 4$ to approximate the nonlinear scaled pricing kernel. They conduct a one-step GMM estimation with the q moment conditions in equation (4) by choosing $k = 1$ to q.

In this paper, I use an n-th order polynomial function approximation to describe the pricing kernel rather than cubic B-splines because the polynomial approximation facilitates the use of conditioning variables in the specification of the pricing kernel.[15] I approximate the scaled pricing

[15]For other studies that approximate the pricing kernel with a polynomial function or its variate, see,

kernel as

$$m_{t,t+\tau}(r_{t+\tau}) = \sum_{i=0}^{n} b_{i,t} r_{t+\tau}^{i}. \tag{5}$$

Note that if the loading $b_{i,t}$ is not time-varying but a constant b_i, equation (5) corresponds to the unconditional scaled pricing kernel. To guarantee the positivity of the pricing kernel, I impose a penalty on the GMM objective function if the scaled pricing kernel estimate has a non-positive value over an acceptable range of the 1-month market return.[16]

It is important to verify that the pricing kernel estimation results do not depend on the approximation methods. Thus, I re-estimate the unconditional pricing kernel using both the cubic B-spline with order seven (the main specification of the CDI method) and the polynomial approximation with order two, and I then compare the results from the two approaches.[17] The left side of Figure 3 summarizes the results. Panel A is based on the cubic B-splines method, and Panel B is based on the polynomial approximation. The empirical pricing kernels are highly similar. The bootstrapped 90% confidence intervals also show that both are not significantly increasing in the right tail, consistent with Linn, Shive, and Shumway (2018).

To incorporate the conditioning variables observable at time t into the pricing kernel, I specify the coefficient $b_{i,t}$ as a function of the conditioning variables. Let $z_{j,t}$ denote the j-th conditioning variable. Then, similar to Lettau and Ludvigson (2001) and Nagel and Singleton (2011), I specify the time-varying coefficient $b_{i,t}$ as a linear function of the time-t observable variables $\{z_{j,t}\}_{j=1}^{m}$:

$$b_{i,t} = \beta_{i,0} + \sum_{j=1}^{m} \beta_{i,j} z_{j,t}, \tag{6}$$

where m is the number of conditioning variables. Moreover, for the conditional pricing kernel

for instance, Rosenberg and Engle (2002), Sandulescu and Schneider (2020), and Schreindorfer and Sichert (2025).

[16]For the main results, I use the range from -100% to 30%. The estimation results remain intact, although I choose a wider range of the market return. Moreover, an alternative way to guarantee the positivity is approximating the log pricing kernel as a polynomial function, and I obtain consistent results under this setup.

[17]A statistical analysis for choosing the maximum order of the polynomial is provided in Section 4.3 below.

estimation, I use the following moment conditions of the physical CDF instead of equation (4):

$$E\left[\left(\left(\int_{-\infty}^{r}\frac{f_t^*(r_{t+\tau})}{m_{t,t+\tau}(r_{t+\tau})}dr_{t+\tau}\right)^k-\frac{1}{1+k}\right)\otimes X_t\right]=0, \tag{7}$$

where X_t is the set of instrumental variables, consisting of the constant and conditioning variables, $\{z_{j,t}\}_{j=1}^m$. The entire time series of the probability densities and the projected pricing kernel are thus described by the set of parameters $\{\beta_{i,j}\}$ for $i = 0, \cdots, n$ and $j = 0, \cdots, m$ as well as the conditioning variables. In this sense, the estimation can be regarded as semi-nonparametric or (almost) model-free because I do not impose any model restrictions on the economy and the distribution of market returns.

4.2 The Modified CDI Method

In addition to using conditioning variables and a polynomial approximation for the pricing kernel, I make two more modifications to the original CDI method of Linn, Shive, and Shumway (2018). First, I improve their method by considering a particular property of the conditional physical density at each time t: the integral of the implied conditional physical density over the support should be equal to one. For simplicity, I call this property the P-density restriction. Moreover, I include the Euler equation in the GMM orthogonality conditions to explain financial asset prices (at least for the market portfolio and risk-free asset).

The CDI method does not examine the conditional physical densities. Since the GMM estimation in the CDI method relies only on the moments of the conditional physical CDFs, as in equation (7), the P-density restriction may not be well captured. On the right side of Figure 3, I depict the histograms of the integral, denoted by CDF(∞) on the x-axis. Panel A shows that many samples obtained from the B-splines estimation significantly deviate from one, and this also occurs in the density samples from the polynomial approximation in Panel B. I, therefore, take into account the P-density restriction as a constraint.

In estimation, I deal with the P-density restriction by including it in the set of GMM orthogonality conditions. Specifically, I represent the restriction in the form of orthogonality conditions as follows:

$$E\left[\begin{pmatrix} \int_{-\infty}^{\infty} \frac{f_t^*(r_{t+\tau})}{m_{t,t+\tau}(r_{t+\tau})} dr_{t+\tau} - 1 \\ \left(\int_{-\infty}^{\infty} \frac{f_t^*(r_{t+\tau})}{m_{t,t+\tau}(r_{t+\tau})} dr_{t+\tau} - 1\right)^2 \end{pmatrix} \otimes X_t\right] = 0. \tag{8}$$

Note that the first orthogonality condition in equation (8) can well capture the P-density restriction unconditionally but may not well satisfy it conditionally. To reduce the conditional restriction errors, I append the second orthogonality condition. In other words, the variance of the errors should be zero, implying that the errors should be all zero. Therefore, I am able to incorporate the restriction in the GMM estimation.

Furthermore, the unconditional empirical pricing kernel estimated by the CDI method may lack economic significance as it is purely based on matching the statistical moments of the physical CDFs. I employ the Euler equation with the market return and the risk-free rate for the last GMM orthogonality conditions.[18] They are mathematically expressed as follows:

$$E\left[\left(M_{t,t+\tau}(r_{t+\tau}) \begin{pmatrix} 1 + r_{t+\tau}^{cd} \\ 1 + r_{f,t}\tau \end{pmatrix} - 1\right) \otimes X_t\right] = 0, \tag{9}$$

where $r_{t+\tau}^{cd}$ is the cum-dividend market return. Although the Euler equation is in principle defined under the conditional expectation, the law of iterated expectations results in equation (9), which can be incorporated into GMM estimation.

The modified CDI method refers to the GMM estimation with all the moment conditions in equations (7), (8), and (9).[19] I estimate the parameters $\{\beta_{i,j}\}$, which specify the conditional empirical pricing kernel, via the two-step efficient GMM. I use the maximum $k = 3$ because it is

[18] Note that the empirical pricing kernel in this paper is the projected pricing kernel on the market return. Therefore, the orthogonal component of the true pricing kernel to the market return is not captured via the empirical pricing kernel. In this sense, using test assets other than the market portfolio and risk-free asset might be inappropriate if their returns potentially have significant correlations with the orthogonal component.

[19] One advantage of this GMM approach is that it can be flexibly extended by adding relevant moment conditions and/or restrictions. In this paper, however, I only keep the most necessary conditions as in these three equations.

sufficient to maintain the number of orthogonality conditions greater than that of parameters in both the unconditional and conditional pricing kernel estimations. The number of parameters to be estimated is $(n+1) \times (m+1)$. The orthogonality conditions consist of $3 \times (m+1)$ moment conditions in equation (7), $2 \times (m+1)$ constraints in equation (8), and $2 \times (m+1)$ Euler equations.

To compare the modified CDI method with the original CDI method, Panel C of Figure 3 shows the unconditional pricing kernel estimation result via the modified CDI method with no conditioning variables included. I use the second-order polynomial approximation as in Panel B. The left-hand graph shows the pricing kernel estimate, which strictly decreases in the market return. The right-hand histogram shows the integrals of the implied conditional physical densities. All numbers are close to one, in contrast to the results in Panels A and B. This result confirms that the P-density restriction is well-managed via equation (8). Henceforth, the unconditional pricing kernel estimate refers to the one estimated using this approach, *i.e.*, the modified CDI method without conditioning variables.[20]

4.3 The Choice of Polynomial Function

To find an optimal degree n of the polynomial function, I statistically compare estimation results for the unconditional pricing kernel with different choices of n, varying from four to one. For each choice of n, I obtain the estimates of $\beta_{i,0}$ for $0 \leq i \leq n$ and then calculate the Wald statistic for testing the restriction of $\beta_{n,0} = 0$.

Table 2 summarizes the parameter estimates. Having $\beta_{0,0}$ near one implies that the pricing kernel is estimated to be around one when the next-period market return is zero. The slope coefficient $\beta_{1,0}$ is estimated to be negative in all the cases and statistically significant in most cases. This result indicates that the unconditional pricing kernel is monotonically decreasing around the

[20]Note that an unconditional pricing kernel estimate can also be obtained via the conditional pricing kernel estimation. If I have information about the joint distribution of the conditioning variables, it is possible to find the corresponding unconditional measure of the pricing kernel. However, since this unconditional measure relies on the conditioning variables, it is not the same as the unconditional estimate obtained through the modified CDI method without conditioning variables. In this paper, I do not calculate the former unconditional pricing kernel estimate; I only use the latter.

zero rate of market return, regardless of the choice of n.

When $n = 4$ or 3, some of the parameter estimates are insignificant at 10%. Their corresponding Wald statistics are not statistically different from zero, implying that the hypothesis of $\beta_{n,0} = 0$ is not rejected. In other words, the polynomial approximation with $n = 4$ or 3 is not distinguishable from the approximation with the max order of $n - 1$. However, when $n = 2$, all the parameter estimates and the Wald statistic are significant at 10%. Each parameter plays a vital role in building the unconditional pricing kernel, and reducing the max order of n from two to one is not innocuous. Therefore, I choose $n = 2$ as the main specification for the remaining unconditional pricing kernel implementation. For simplicity, I maintain the same max order for conditional pricing kernel analyses.

5 Estimation Results

Since the conditioning variables that are most informative about the pricing kernel are unknown, I first estimate the conditional pricing kernel with one conditioning variable at a time. I refer to the resulting estimates as the "univariate conditional empirical pricing kernels." Then, I determine which variables are the most informative about the pricing kernel by inspecting the statistical significances of the parameter estimates as well as the over-identification test statistics.

After selecting a set of preferred conditioning variables, I estimate the conditional pricing kernel by incorporating those multiple conditioning variables. I use the term "multivariate conditional empirical pricing kernel" to indicate the resulting estimate.[21] Then, using the time series of the estimated pricing kernel and the probability densities, I investigate their time variation and the implications for the conditional risk premia.

[21] In this paper, the term "univariate" or "multivariate" refers to the number of conditioning variables but not the number of state variables describing the future states of the economy.

5.1 Univariate Conditional Pricing Kernel Estimates

Table 3 summarizes the parameter estimates for the univariate conditional pricing kernel. Each row presents estimation results for each of the 12 conditioning variables. The parameter $\beta_{i,j}$ represents the coefficient of the j-th conditioning variable in the time-varying coefficient $b_{i,t}$ of $r^i_{t+\tau}$. For instance, $\beta_{0,1}$ determines the variation in the level of the pricing kernel with respect to the change in a conditioning variable. The parameter $\beta_{1,1}$ captures the variation in the slope of the pricing kernel when the market return is zero. When $j = 0$, it corresponds to the constant term in the time-varying coefficients. In the last column, I report the over-identification test statistic (J statistic) of Hansen (1982) and its p-value for each GMM estimation in the bracket.

Note that at the zero rate of market return, the value of the empirical pricing kernel is $\beta_{0,0} + \beta_{0,1} z_{j,t}$. The estimate of $\beta_{0,0}$ is around one for all the specifications. This result implies that at the zero rate of return, the pricing kernel is set around one, and each conditioning variable further adjusts it. The magnitude of this estimate ($\beta_{0,0}$) is economically sensible in the sense that the pricing kernel is interpreted as the stochastic discount factor. When I use the VIX as the conditioning variable, the estimate of $\beta_{0,1}$ is -0.41 and statistically significant. This result indicates that the overall level of the pricing kernel decreases by 0.04 as the VIX increases by 0.1 (or 10%).

The estimate of $\beta_{1,0}$ is significantly negative in most cases, and some of the $\beta_{1,1}$ estimates are also significant. The negative $\beta_{1,0}$ implies that the unconditional component of the slope at the zero rate of return is negative; that is, the pricing kernel is unconditionally downward sloping around the zero rate of return. Furthermore, for conditioning variables that are measures of risk (the VIX, risk-neutral skewness and kurtosis, and left-tail risk index), the significantly positive estimates of $\beta_{1,1}$ indicate that the lower the risk, the steeper the pricing kernel estimates. For example, when the conditioning variable is the VIX, a 0.1 (or 10%) difference in the VIX yields a slope difference of 0.851 at the zero rate of return. When the expectation about the future financial market or economy (as measured by the sentiment, term spread, or credit spread) is favorable or when the current market conditions (as measured by the ADS index, industrial production growth, consumption growth, or monthly market return) are good, the slope of the pricing kernel becomes steeper, supported by significant $\beta_{1,1}$ estimates.

The over-identification test results also provide an important message. When the VIX is the conditioning variable, the J statistic is not rejected at the 10% level. When the pricing kernel estimate is conditional on the term spread, the J statistic is not rejected at the 5% level. Specifications with the conditioning variables of sentiment, industrial production growth, credit spread, or market return show their J statistics are not rejected at the 1% level. For all other conditioning variables, the test of the over-identification restriction rejects at the 1% level. In other words, when conditioning variables other than the VIX, sentiment, term spread, industrial production growth, credit spread, and market return are used, the GMM over-identification orthogonality restrictions are highly rejected.

Considering the statistical significance of the parameter estimates and the results of the over-identification restriction test, I choose the VIX, sentiment, term spread, industrial production growth, credit spread, and market return as conditioning variables that are most informative about the pricing kernel. Also, it is worthwhile to mention that this paper deals with the forward-looking estimate of the pricing kernel, and the VIX, sentiment, term spread, and credit spread are regarded as forward-looking measures.

Figure 4 depicts conditional empirical pricing kernels with a single conditioning variable. Panels A through F show the estimation results based on the VIX, sentiment, term spread, industrial production growth, credit spread, and market return, respectively. The left-hand graphs are the time series average of the pricing kernel, and the right-hand graphs are the sensitivities of the pricing kernel with respect to each conditioning variable. In the sensitivity analysis, I vary each variable between the 5th percentile (low) and the 95th percentile (high) of its time series data, and I then plot the fitted pricing kernel estimates.

The average pricing kernel estimates in all four panels look similar, specifically in the left tail. However, the empirical pricing kernel in Panel A slightly increases in the right tail, whereas those in the other panels strictly decrease. Nevertheless, the monotonicity of the pricing kernel in Panel A is not rejected at the 10% significance level up to 5% of the market return.

The sensitivity analysis shows apparent variation in the shape of the conditional pricing kernel estimates. In all panels, the variation is more conspicuous in the left tail than in the right tail, perhaps because investors react more strongly to negative shocks than positive shocks to the market and economy. Hence, reactions to a shock are far more sensitive to the conditional information when the shock is negative to the market return than when the shock is positive.

These results raise the question how the variations in the empirical pricing kernels should be interpreted. For instance, in the negative return region, the pricing kernel associated with the high VIX is lower and less steep than that associated with the low VIX, which is consistent with the model implications from Christoffersen, Heston, and Jacobs (2013) and the result of Schreindorfer and Sichert (2025). This observation indicates that for negative future return states, investors become more patient to consume when the VIX is lower. One possible explanation is as follows. During low VIX periods, as volatility risk is low, highly negative return states are less likely to be expected in the near future, compared to high VIX periods. That is, investors would be more fearful of negative returns. If those states are realized, it will be a significant negative shock to investors. Therefore, they are willing to save more for negative states during low VIX periods, pushing the pricing kernel upward and steepening it.

Furthermore, the pricing kernel estimate in the negative return region is lower and less steep during the low term spread period than during the high term spread period. When the term spread is relatively higher, investors are less likely to expect highly negative return states in the near future. Hence, during such periods, the realization of negative return states will be a substantial negative shock to investors. As a result, they become more patient to consume but rather save more, forcing the pricing kernel upward and steepening it. Similar logic can be applied to the pricing kernel estimates associated with the other four conditioning variables (sentiment, industrial production growth, credit spread, and market return).

5.2 Multivariate Conditional Pricing Kernel Estimates

The univariate conditional pricing kernel estimates in Table 3 show that the VIX, sentiment, term spread, industrial production growth, credit spread, and market return are state variables that are informative about the pricing kernel. I start with using two conditioning variables in the multivariate estimation because the optimal set and number of conditioning variables are unknown. Subsequently, I extend the estimation with three conditioning variables and evaluate if its result improves statistically.

Note that the univariate estimation results indicate that the VIX is the most informative variable. I, therefore, consider the following five bivariate estimation specifications relying on (1) the VIX and sentiment, (2) the VIX and term spread, (3) the VIX and industrial production growth, (4) the VIX and credit spread, and (5) the VIX and market return. Each specification contains nine free parameters.

Panel A of Table 4 reports the resulting parameter estimates with two conditioning variables. The overall estimation results are similar for all five cases. The parameter estimate of $\beta_{1,0}$, which is the constant component of the time-varying coefficient of $r_{t+\tau}$, is significantly negative in all specifications, implying that the unconditional component of the pricing kernel is decreasing around the central region of the return distribution. Moreover, in specification (2), the estimates of both $\beta_{1,1}$ and $\beta_{1,2}$ are significantly positive and negative, respectively. In other specifications, however, only $\beta_{1,1}$, which is the sensitivity to the VIX, is statistically significant, although the signs of both $\beta_{1,1}$ and $\beta_{1,2}$ estimates are consistent with the results from the univariate estimation. On the other hand, the $\beta_{2,2}$ estimate is statistically significant when one of the conditioning variables is the industrial production growth or market return, but any of the sensitivities to the sentiment or credit spread are not significant. These observations are presumably due to the fact that the sentiment and credit spread are highly correlated with the VIX (-0.625 and 0.552, respectively), as seen in Table 1. In contrast, the term spread provides additional forward-looking information with its low correlation with the VIX, and the industrial production growth and market return are marginally informative as they represent current market conditions.

I conduct two Wald tests to see if each bivariate conditional pricing kernel shows statistically better in-sample performance than the unconditional pricing kernel and if each bivariate conditional pricing kernel performs better than the univariate conditional pricing kernel using the VIX. The former (latter) tests the restriction of $\beta_{i,1} = \beta_{i,2} = 0$ ($\beta_{i,2} = 0$) for all i, and I denote its statistic as Wald_{unc} ($\text{Wald}_{\text{cond}}$). In Table 4, Wald_{unc}'s, as well as their p-values in the bracket, show that conditioning variables improve the estimation of the pricing kernel in all specifications. On the other hand, $\text{Wald}_{\text{cond}}$'s are statistically significant at the 5% significance level only in three specifications where the term spread, industrial production growth or market return is an additional conditioning variable. Note that Hansen's J statistics show that the over-identification test for only the specification with the term spread is not rejected at the 5% level. Based on the significance of parameter estimates, over-identification test, and Wald tests, I choose the VIX and term spread as the most informative set of conditioning variables for the bivariate estimation.

Figure 5 depicts the average conditional pricing kernel estimates for all five specifications, which look pretty similar, particularly in the negative future return region. Although some of them are U-shaped and the others are monotonically decreasing within the range from -10% to 10% of a monthly market return, their 90% confidence intervals still provide comparable implications. Within the same range, the monotonicity is not rejected at least up to 8% of the market return (which is slightly smaller than its two standard deviations) in all five panels but is rejected in deep right tails except in Panel C.

I also estimate the pricing kernel by using three conditioning variables and compare it with the bivariate (the VIX and term spread) kernel estimate. I fix two variables, the VIX and term spread, and then include each of marginally informative variables inferred from the bivariate estimation. There are two trivariate estimation specifications, where conditioning variables are (1) the VIX, term spread, and industrial production growth, and (2) the VIX, term spread, and market return. Panel B of Table 4 summarizes the estimation results. None of the two specifications have statistically significant estimates of $\beta_{i,2}$ for all $i = 0,\ 1,$ and 2. Hansen's J statistics reject the null hypothesis at the 1% level. In this panel, $\text{Wald}_{\text{cond}}$ is for testing the restriction of $\beta_{i,3} = 0$; that is, it is comparing each pricing kernel estimate with the bivariate specification using the VIX and term spread. None of the $\text{Wald}_{\text{cond}}$'s is significant even at the 10% level, implying that the bivariate pricing kernel estimate is statistically indistinguishable from trivariate estimates.Therefore, I use the bivariate conditional pricing kernel with the VIX and term spread for the remaining analyses and discussion.

Panel A of Figure 6 shows the results of the sensitivity analysis with respect to the two conditioning state variables, the VIX and term spread. Consistent with the univariate estimates, the differences across the levels of the state variables are noticeable in the left tail region. In line with the similarities and differences between the parameter estimates in Tables 3 and 4, Figures 4 and Panel A of Figure 6 graphically indicate that all the sensitivity patterns based on the multivariate estimation are equivalent to those from the univariate estimation.

Panel B of Figure 6 depicts the time series of the bivariate empirical pricing kernel. The shape of the empirical pricing kernel varies considerably over time. Its time variation in the left tail of the return distribution is particularly pronounced, consistent with observations in Panel A. The pricing kernel estimate is relatively very low in the negative return region during economically bad times, such as the 2008 financial crisis and the 2020 COVID-19 crisis. At the -10% level of the future monthly return, their pricing kernel estimates are even lower than one. These estimates indicate that investors are still willing to discount future dollars during such periods in contrast with the estimates during normal times.

5.3 Conditional Risk Premia

Since I estimate the forward-looking conditional pricing kernel, the time series of the forward-looking conditional physical density can be constructed as a natural by-product of the analysis based on equation (3). The time series data of the risk-neutral and physical densities enable the direct calculation of the s-th moment forward-looking risk premium at time t, $\text{RP}_t^{(s)}$, as well as its statistical moments:

$$
\mathrm{RP}_t^{(s)} = \begin{cases} E_t[r_{t+\tau}^s] - E_t^*[r_{t+\tau}^s] & \text{if} \quad s = 1 \\ E_t[(r_{t+\tau} - \mu_t)^s] - E_t^*[(r_{t+\tau} - \mu_t^*)^s] & \text{if} \quad s = 2 \\ \frac{E_t[(r_{t+\tau} - \mu_t)^s]}{\sigma_t^s} - \frac{E_t^*[(r_{t+\tau} - \mu_t^*)^s]}{(\sigma_t^*)^s} & \text{if} \quad s = 3 \text{ or } 4, \end{cases} \tag{10}
$$

where μ_t, σ_t, μ_t^*, and σ_t^* are respectively the conditional physical mean, standard deviation, risk-neutral mean, and standard deviation, and

$$
\begin{aligned}
E_t[g(r_{t+\tau})] &= \int_{-\infty}^{\infty} g(r_{t+\tau}) f_t(r_{t+\tau}) dr_{t+\tau}, \\
E_t^*[g(r_{t+\tau})] &= \int_{-\infty}^{\infty} g(r_{t+\tau}) f_t^*(r_{t+\tau}) dr_{t+\tau},
\end{aligned}
$$

for an integrable function, $g(\cdot)$.

Table 5 summarizes the first four statistical moments of the risk-neutral (Panel A) and physical distributions of monthly returns.[22] For the physical distributions, I separately report the statistics obtained from the unconditional pricing kernel estimation (Panel B) and the conditional pricing kernel estimation (Panel C). Since the conditional mean, standard deviation, skewness, and kurtosis are calculated at each point in time, I report their time-series mean, standard deviation, and quantile distributions based on their 5% to 95% values.

The first moment of the risk-neutral density estimates, 0.21%, should be the monthly risk-free rate; it is overall lower than that of the physical density (from either the unconditional or conditional estimation). On the other hand, the magnitudes of the higher moments of the risk-neutral density exceed those of the physical densities. All the three distributions are negatively skewed and leptokurtic, which is confirmed in Figure 7 depicting the average risk-neutral (solid blue line) and physical (dotted black line - implied by the unconditional kernel, dashed red line - implied by the conditional kernel) probability distributions. Moreover, when comparing Panel B

[22]The market return data used in this calculation is the ex-dividend return. For better interpretation, dividends are adjusted, and thus the reported numbers in this table are based on the cum-dividend return.

with Panel C of Table 5, the first moment's time series statistics are fairly distinct from each other. The mean value of the first moment in Panel B and C are 0.96% and 0.84%, respectively, and the latter is closer to the sample average of the recent 25-year (from 1996 to 2020) monthly market return, 0.88%. The time series standard deviation of the expected return in Panel B is greater than twice that in Panel C. The unconditional estimate of the pricing kernel may put a significant restriction on the formation of reliable physical density estimates, although the Euler equation is one of the orthogonality conditions in the GMM setup.

As the conditional moments of the risk-neutral and physical densities are found, I can calculate the conditional equity, variance, skewness, and kurtosis risk premia based on both the unconditional pricing kernel and conditional pricing kernel by equation (10). Table 6 reports the descriptive statistics for the risk premia. Panel A presents the result from the unconditional pricing kernel estimates, and Panel B for the conditional pricing kernel estimates. The table contains the information about the time series mean, standard deviation, quantile distribution, and first-order autocorrelation (AR1). As inferred from Table 5, the time series average of the equity risk premium (ERP) in Panel A (0.76%) is estimated to be higher than its data counterpart, which is 0.67%, while the average of the conditional ERP in Panel B (0.63%) is closer to 0.67%. Meanwhile, the variance risk premium (VRP) looks reasonable, and the time series means are -10.36%^{2} and -9.35%^{2} in Panels A and B. For comparison, I obtain the updated VRP data of Bollerslev, Tauchen, and Zhou (2009) from Hao Zhou's website, and their average VRP is -14.40%^{2} for my sample period, which is not very different from the values in Panels A and B.Since the VRP values specifically depend on how the expected return variances are measured, the numbers in Panels A and B are sufficiently sizable and reasonable despite the difference.

Figure 8 shows the time series of the four risk premia based on the unconditional pricing kernel (solid blue line) and the conditional pricing kernel (red dashed line). Consistent with the descriptive statistics in Table 6, the implications of the unconditional and conditional pricing kernel estimates for the ERP are distinguishable, whereas those for the other risk premia are very similar. In Panel A, during normal times, both the level and variation of the conditional ERPs from the unconditional and conditional pricing kernel estimates are very similar. By contrast, during economic downturns, the difference between the two ERPs becomes very large. For example, at the peak of the 2008 financial crisis or the COVID-19 crisis, the difference increases up to nearly 2% per month. This large gap during bad times is the main reason for the higher average ERP in Panel A of Table 6. This implies that if conditional information is not adequately taken into account, the implied conditional ERP may be exaggerated during bad times. In other words, relevant conditional information helps capture not only the average but also the time series of the conditional ERP.

5.4 The Sources of the Conditional Equity Premium

It is also interesting to investigate the conditional ERP from the perspective of the sources of the ERP. Following Beason and Schreindorfer (2022), I evaluate how much each monthly return state contributes to the total ERP. Different from their approach, which uses the unconditional probability densities, I take advantage of the conditional probability densities to figure out the time variation in the sources of the ERP. This exercise is implemented by calculating the fraction of the conditional equity premium at the level of market return r as follows:

$$\text{Contribution}_t(r) = \frac{\int_{-\infty}^{r} r_{t+\tau}[f_t(r_{t+\tau}) - f_t^*(r_{t+\tau})]dr_{t+\tau}}{E_t[r_{t+\tau}] - E_t^*[r_{t+\tau}]}.$$

Figure 9 shows the contribution of monthly return states to the conditional ERP. Panel A presents the average contribution in the full sample period, and Panels B and C present the average contributions in low-VIX and high-VIX subsample periods.[23] The low-VIX (high-VIX) period consists of the subsample where the VIX is lower (higher) than or equal to the 10th (90th) percentile of the VIX data. When the full sample period is used, the negative returns contribute, on average, to the conditional ERP slightly more when the conditional information is incorporated (dashed red line) than when it is not (solid blue line). The gap between the contributions from the unconditional and conditional pricing kernel estimates during the low-VIX period becomes smaller than the

[23] I conduct the subsample analysis based on the VIX because it is the most informative conditioning variable.

average gap. However, the gap during the high-VIX period is significantly larger than the average gap. If no conditional information is taken into account, a 99% contribution is attained at around 15% of the market return. In contrast, a 99% contribution is achieved at around -20% of the market return if the conditional pricing kernel is used.

Instead of comparing the average contributions during the subsample periods, it is also possible to examine the time series return level for a specific percentage contribution to the conditional ERP. Figure 10 depicts this time series information for a 10% contribution (Panel A), 50% contribution (Panel B), and 99% contribution (Panel C). Panel A shows that the 10% contribution is made at similar return levels, regardless of whether the unconditional or conditional pricing kernel is specified. The exceptions are the peak of the 2008 financial crisis and the 2020 COVID-19 pandemic crisis. Panel C, meanwhile, exhibits that the return levels for the 99% contribution are significantly different depending on the use of the unconditional or conditional pricing kernel. Although they are highly correlated and not much different during the mid-2000s and mid-2010s, the divergence is considerably large during the 1998 LTCM collapse, 2002 stock market crash, 2008-2009 financial crisis, 2012 European debt crisis, and 2020 COVID-19 pandemic crisis. For instance, at the peak of the 2008 financial crisis, the nearly -30% return attains the 99% contribution to the ERP based on the conditional pricing kernel, while the 20% return attains the 99% contribution based on the unconditional pricing kernel. Not only is the level completely different, the patterns are too. During such economic or financial downturns, investors fear an increasing probability of large negative returns, and thus they require more risk compensation, specifically from the negative return states. In line with this, the observed time series pattern of the ERP from the conditional pricing kernel estimate in Panel C is economically plausible, but that from the unconditional pricing kernel estimate is not. Figures 9 and 10 confirm the message that different implications for the conditional ERPs between the unconditional and conditional empirical pricing kernels are magnified during economically bad times.

5.5 Out-of-Sample Evaluation: Option Returns

By the fundamental theorem of asset pricing, one should be able to price any financial assets using the Euler equation once the pricing kernel is given. As a result of estimation, I obtain either the unconditional pricing kernel or the conditional pricing kernel as a function of the future one-month return. This setup provides the following out-of-sample testable hypothesis:

$$E\left[M_{t,t+\tau}(r_{t+\tau})\left(1+r^x_{t+\tau}\right)\right]-1=0, \tag{11}$$

where $r^x_{t+\tau}$ is the τ-period return on an any asset x, and the expectation is computed by taking the time-series average. I use only the market return and the risk-free rate in the in-sample estimation. One candidate $r^x_{t+\tau}$, closely related to the aggregate market portfolio, to test the hypothesis in equation (11) is the option return. To conduct this out-of-sample test, I utilize the Volatility Surface data from OptionMetrics because the dataset provides interpolated option prices on a uniform grid of option maturities and Black-Scholes deltas. Since the estimation uses monthly returns measured at the end of each month, I also focus on 30-day maturity options. I specifically choose options with deltas of -0.3, -0.5, and -0.7 for puts and deltas of 0.3, 0.5, and 0.7 for calls as test assets. In addition to examining simple option returns, I also investigate delta-hedged option portfolio returns.[24]

Table 7 reports the out-of-sample testing results. Panel A presents the results for the unconditional empirical pricing kernel, and Panel B presents those for the conditional empirical pricing kernel. In each panel and for each of option return and delta-hedged return tests, I report the average error corresponding to the left-hand side of equation (11). If this average error is not statistically different from zero, the null hypothesis is not rejected.

Mixed evidence is observed for call option returns when the unconditional kernel is applied.

[24] For the delta-hedged call return calculation, I follow Almeida, Freire, and Hizmeri (2025):

$$\frac{\max(S_{t+\tau}-K,0)-Call_t(K,\tau)-\Delta(S_{t+\tau}-S_t)-(Call_t(K,\tau)-\Delta S_t)r_{f,t}\tau}{S_t}.$$

While it fails to explain out-of-the-money (delta = 0.3) call returns, it explains returns for near-the-money (delta = 0.5) and in-the-money (delta = 0.7) calls. On the other hand, the null hypothesis is not statistically rejected for any of the three call returns when the conditional empirical pricing kernel is used. The magnitudes of the average errors under the conditional kernel are also much smaller than those under the unconditional kernel. Similar patterns are observed for delta-hedged call returns, except that the out-of-the-money delta-hedged call return is also well explained under the unconditional kernel. This outperformance of the conditional pricing kernel can be linked to the time variation of kernel estimates. Although the conditional kernel estimate does not show a pronounced increasing pattern in the right tail on average, depending on market conditions, there are times when it sufficiently exhibits an increasing pattern, which helps to price out-of-the-money calls higher and consequently explains the negative call returns demonstrated in the existing literature, leading to improvement in call returns.

I find different implications for delta-hedged put returns and outright put returns. The delta-hedged put return results show a clearer difference between the two types of the pricing kernels. While the null hypothesis is rejected at the 10% significance level for all three delta-hedged put returns under the unconditional kernel, it is not rejected under the conditional kernel. On the other hand, the Euler equation does not hold for outright put option returns in either Panel A or B; the null hypothesis is statistically rejected.[25] It is important to note again, however, that the pricing kernels are not estimated by optimally matching option returns. It is also well-known that matching out-of-the-money put option prices and returns is particularly difficult even in-sample unless complex parametric models, such as those with stochastic volatility and jumps, are considered (see, for example, Hu and Liu, 2022).

[25] A potential explanation of this result is insufficient increases of the pricing kernel in the deep left tail to capture highly negative put option returns out-of-sample (OOS). This inference might be due to a result of approximating the pricing kernel as a second-order polynomial, which somewhat restricts its shape. Since the main goal of this paper is to show the time variation of the pricing kernel through a parsimonious functional form, we acknowledge that this restriction may not sufficiently explain OOS outright put option returns.

6 Conclusion

Accurate estimation of the pricing kernel is crucial for understanding asset prices and investors' beliefs. While the literature emphasizes the importance of conditional information, the conditional pricing kernel has received limited attention due to its complexity relative to the unconditional pricing kernel. Prior studies estimate the conditional kernel using historical data or specific model assumptions, but their results are highly sensitive to these specifications. In this paper, I propose a novel approach to semi-nonparametrically estimate the forward-looking conditional pricing kernel.

I use conditioning variables to capture the time variation of the pricing kernel. For its implementation, I modify the CDI method that is proposed by Linn, Shive, and Shumway (2018) to estimate the unconditional pricing kernel. Using the modified CDI method with one conditioning variable, I estimate the conditional pricing kernel and find that the VIX, sentiment, term spread, industrial production growth, credit spread, and market return are informative about the 1-month forward pricing kernel. Using these selected conditioning variables, I also estimate multivariate conditional pricing kernels and find that the combination of the VIX and term spread is the most informative. The estimated kernels are, on average, downward sloping up to two standard deviations of the return, and they show significant time variation, which is consistent with economic intuition; for example, the pricing kernel estimate in negative return states is lower when conditional information contains unfavorable market expectations than when it contains favorable expectations.

Finally, I examine the conditional risk premia derived from the pricing kernel estimates, with or without conditional information, and find evidence bolstering the importance of incorporating the conditioning variables. The empirical results show significant differences between the unconditional and conditional approaches, specifically with respect to the equity risk premium. The conditional kernel estimate yields a more plausible time series pattern for the equity premium. Furthermore, the conditional kernel estimate outperforms the unconditional kernel estimate in explaining option returns and delta-hedged returns out-of-sample. I conclude that considering relevant conditional information is critically important to understand investors' beliefs and asset prices.

References

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 Andrew W. Lo, 2000, Nonparametric risk management and implied risk aversion, *Journal of Econometrics* 94, 9–51.

Almeida, Caio, and Gustavo Freire, 2022, Demand in the option market and the pricing kernel, Working paper, Princeton University.

Almeida, Caio, Gustavo Freire, and Rodrigo Hizmeri, 2025, 0DTE asset pricing, Working paper, Princeton University.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015, The risk premia embedded in index options, *Journal of Financial Economics* 117, 558–584.

Aruoba, S Borağan, Francis X Diebold, and Chiara Scotti, 2009, Real-time measurement of business conditions, *Journal of Business & Economic Statistics* 27, 417–427.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bakshi, Gurdip, Dilip Madan, and George Panayotov, 2010, Returns of claims on the upside and the viability of U-shaped pricing kernels, *Journal of Financial Economics* 97, 130–154.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long-run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barone-Adesi, Giovanni, Robert F. Engle, and Loriano Mancini, 2008, A GARCH option pricing model with filtered historical simulation, *Review of Financial Studies* 21, 1223–1258.

Barone-Adesi, Giovanni, Nicola Fusari, Antonietta Mira, and Carlo Sala, 2020, Option market trading activity and the estimation of the pricing kernel: A Bayesian approach, *Journal of Econometrics* 216, 430–449.

Bates, David S., 2000, Post-'87 crash fears in the S&P 500 futures option market, *Journal of Econometrics* 94, 181–238.

Beason, Tyler, and David Schreindorfer, 2022, Dissecting the equity premium, *Journal of Political Economy* 130, 2203–2222.

Black, Fischer, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–654.

Bollerslev, Tim, Michael Gibson, and Hao Zhou, 2011, Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities, *Journal of Econometrics* 160, 235–245.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Bollerslev, Tim, Viktor Todorov, and Lai Xu, 2015, Tail risk premia and return predictability, *Journal of Financial Economics* 118, 113–134.

Bondarenko, Oleg, 2014, Why are put options so expensive?, *The Quarterly Journal of Finance* 4, 1450015.

Breeden, Douglas T., and Robert H. Litzenberger, 1978, Prices of state-contingent claims implicit in option prices, *Journal of Business* 51, 621–651.

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 Liuren Wu, 2009, Variance risk premiums, *Review of Financial Studies* 22, 1311–1341.

Chabi-Yo, Fousseni, 2012, Pricing kernels with stochastic skewness and volatility risk, *Management Science* 58, 624–640.

Chabi-Yo, Fousseni, Rene Garcia, and Eric Renault, 2007, State dependence can explain the risk aversion puzzle, *Review of Financial Studies* 21, 973–1011.

Chabi-Yo, Fousseni, and Johnathan Loudis, 2022, A decomposition of conditional risk premia and implications for representative agent models, Working paper, University of Notre Dame.

Chernov, Mikhail, 2003, Empirical reverse engineering of the pricing kernel, *Journal of Econometrics* 116, 329–364.

Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2013, Capturing option anomalies with a variance-dependent pricing kernel, *Review of Financial Studies* 26, 1963–2006.

Cuesdeanu, Horatio, and Jens Carsten Jackwerth, 2018a, The pricing kernel puzzle in forward looking data, *Review of Derivatives Research* 21, 253–276.

Cuesdeanu, Horatio, and Jens Carsten Jackwerth, 2018b, The pricing kernel puzzle: Survey and outlook, *Annals of Finance* 14, 289–329.

Donaldson, R. Glen, Mark J. Kamstra, and Lisa A. Kramer, 2010, Estimating the equity premium, *Journal of Financial and Quantitative Analysis* 45, 813–846.

Driessen, Joost, Joren Koëter, and Ole Wilms, 2025, Horizon effects in the pricing kernel: How investors price short-term versus long-term risks, *Journal of Financial and Quantitative Analysis, Forthcoming.*

Duan, Jin-Chuan, and Weiqi Zhang, 2014, Forward-looking market risk premium, *Management Science* 60, 521–538.

Fama, Eugene F., and Kenneth R. French, 2002, The equity premium, *Journal of Finance* 57, 637–659.

Gagliardini, Patrick, Christian Gourieroux, and Eric Renault, 2011, Efficient derivative pricing by the extended method of moments, *Econometrica* 79, 1181–1232.

Han, Bing, 2008, Investor sentiment and option prices, *Review of Financial Studies* 21, 387–414.

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Harvey, Campbell, 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289–317.

Hens, Thorsten, and Christian Reichlin, 2013, Three solutions to the pricing kernel puzzle, *Review of Finance* 17, 1065–1098.

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., and Saikat Nandi, 2000, A closed-form GARCH option valuation model, *Review of Financial Studies* 13, 585–625.

Hu, Guanglian, and Yuguo Liu, 2022, The pricing of volatility and jump risks in the cross-section of index option returns, *Journal of Financial and Quantitative Analysis* 57, 2385–2411.

Jackwerth, Jens Carsten, 2000, Recovering risk aversion from option prices and realized returns, *Review of Financial Studies* 13, 433–451.

Lettau, Martin, and Sydney C. Ludvigson, 2001, Consumption, aggregate wealth and expected stock returns, *Journal of Finance* 56, 815–849.

Linn, Matthew, Sophie Shive, and Tyler Shumway, 2018, Pricing kernel monotonicity and conditional information, *Review of Financial Studies* 31, 493–531.

Nagel, Stefan, and Kenneth J. Singleton, 2011, Estimation and evaluation of conditional asset pricing models, *Journal of Finance* 66, 873–909.

Pan, Jun, 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.

Pástor, Ľuboš, and Robert F. Stambaugh, 2001, The equity premium and structural breaks, *Journal of Finance* 56, 1207–1239.

Polkovnichenko, Valery, and Feng Zhao, 2013, Probability weighting functions implied in options prices, *Journal of Financial Economics* 107, 580–609.

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.

Sandulescu, Mirela, and Paul Schneider, 2020, Mispricing and uncertainty in international markets, Swiss Finance Institute Research Paper.

Schneider, Paul, 2022, Ross recovery with time series information and economic constraints, Swiss Finance Institute Research Paper.

Schreindorfer, David, and Tobias Sichert, 2025, Conditional risk and the pricing kernel, *Journal of Financial Economics* 171, 104106.

Seo, Sang Byung, and Jessica A. Wachter, 2019, Option prices in a model with stochastic disaster risk, *Management Science* 65, 3449–3469.

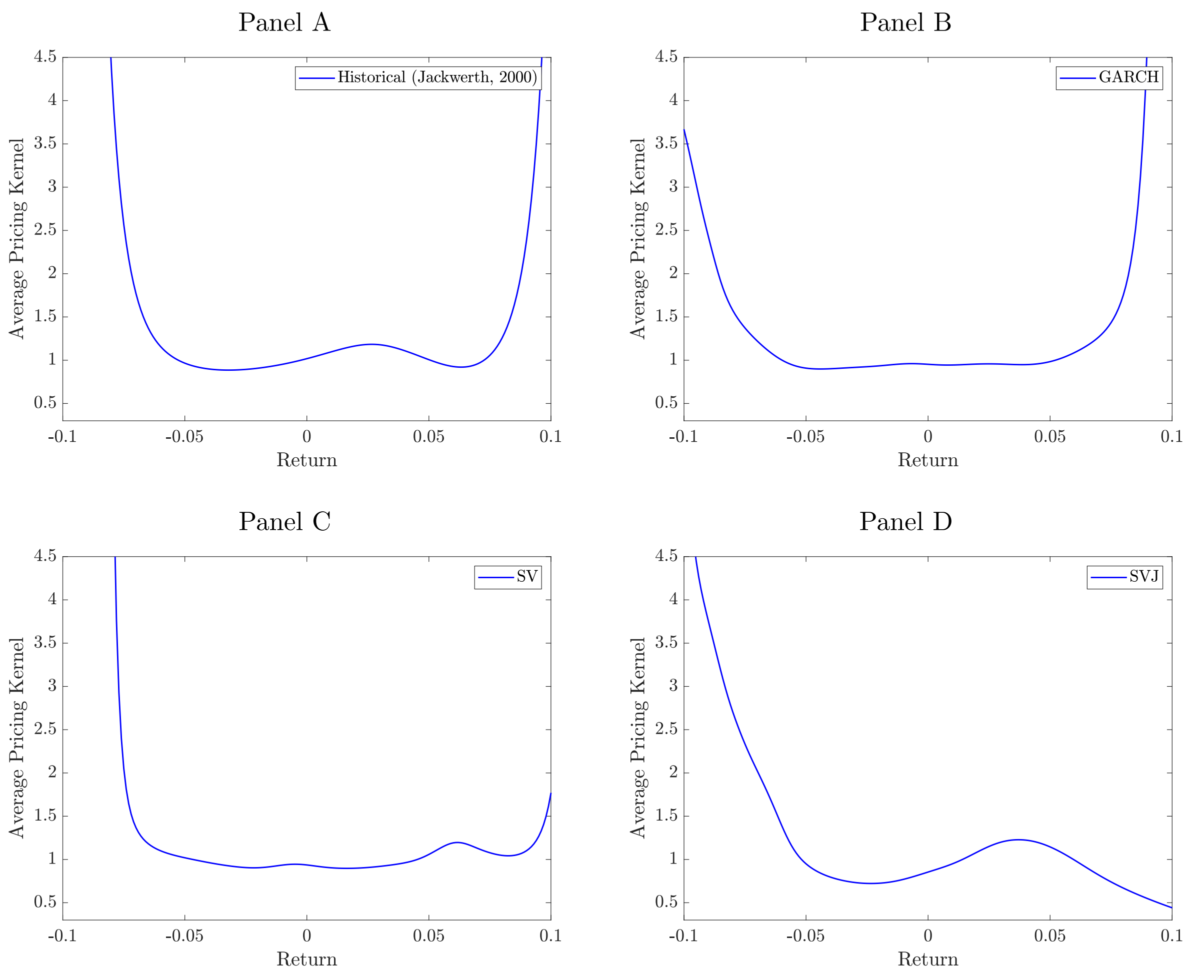
Shefrin, Hersh, 2001, On kernels and sentiment, Working paper.

Song, Zhaogang, and Dacheng Xiu, 2016, A tale of two option markets: Pricing kernels and volatility risk, *Journal of Econometrics* 190, 176–196.

Todorov, Viktor, 2010, Variance risk-premium dynamics: The role of jumps, *Review of Financial Studies* 23, 345–383.

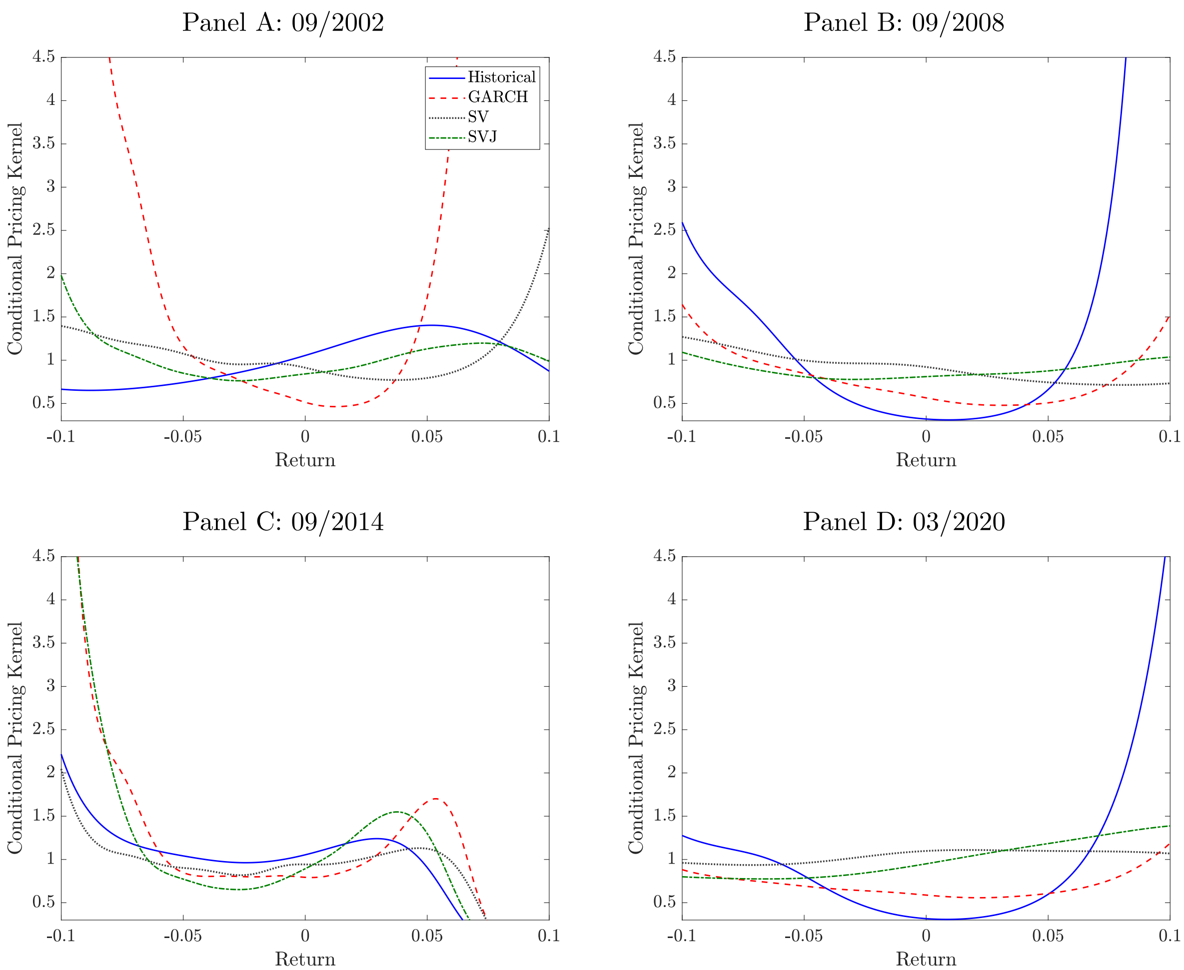
Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *Journal of Finance* 68, 987–1035.

Figure 1: Average Conditional Empirical Pricing Kernels - Existing Approaches



Notes: This figure shows the average of the conditional empirical pricing kernels where the physical densities are estimated using the historical data (Panel A), GARCH model (Panel B), stochastic volatility model (Panel C), and stochastic volatility and jump model (Panel D).

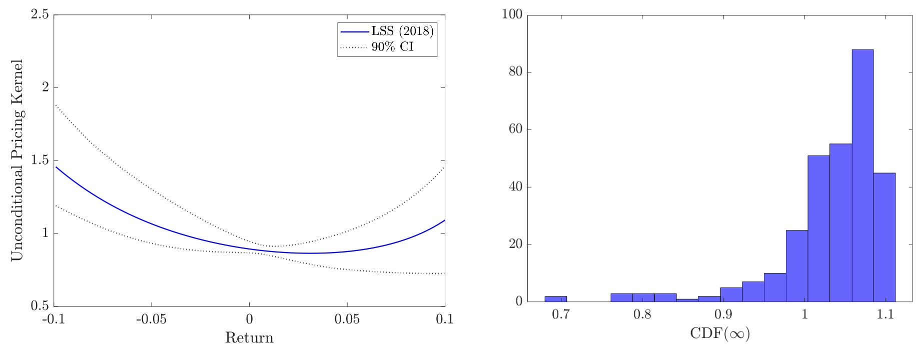
Figure 2: Conditional Empirical Pricing Kernels - Existing Approaches



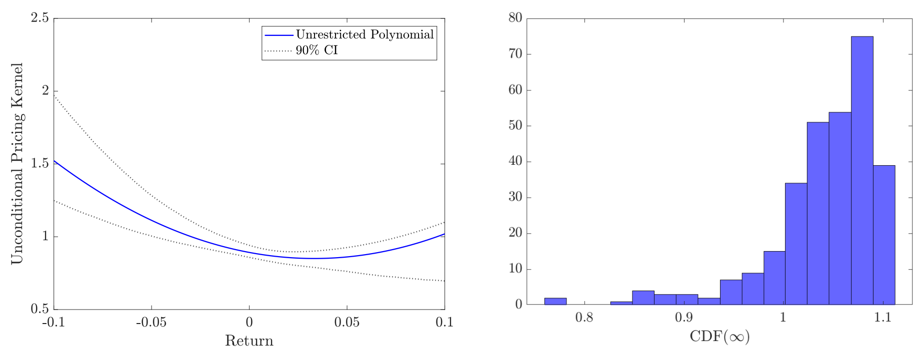
Notes: This figure shows the conditional empirical pricing kernels estimated using the historical data, GARCH model, stochastic volatility (SV) model, and stochastic volatility with jumps (SVJ) model on specific months: 09/2002 (Panel A), 09/2008 (Panel B), 09/2014 (Panel C), and 03/2020 (Panel D).

Figure 3: Unconditional Empirical Pricing Kernels

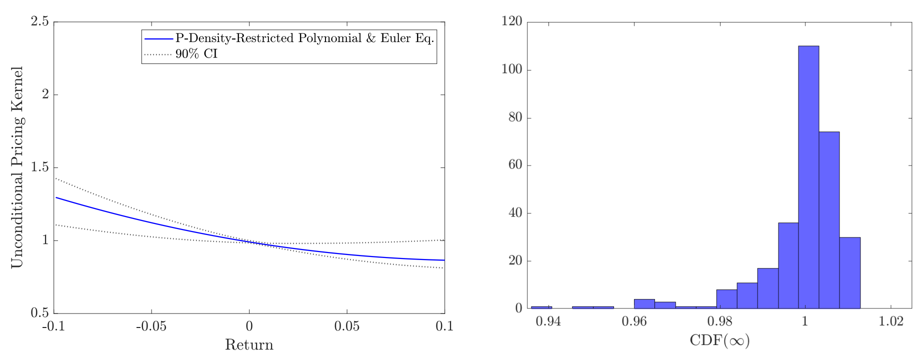
Panel A: Linn, Shive, and Shumway (2018)



Panel B: Unrestricted Polynomial EPK

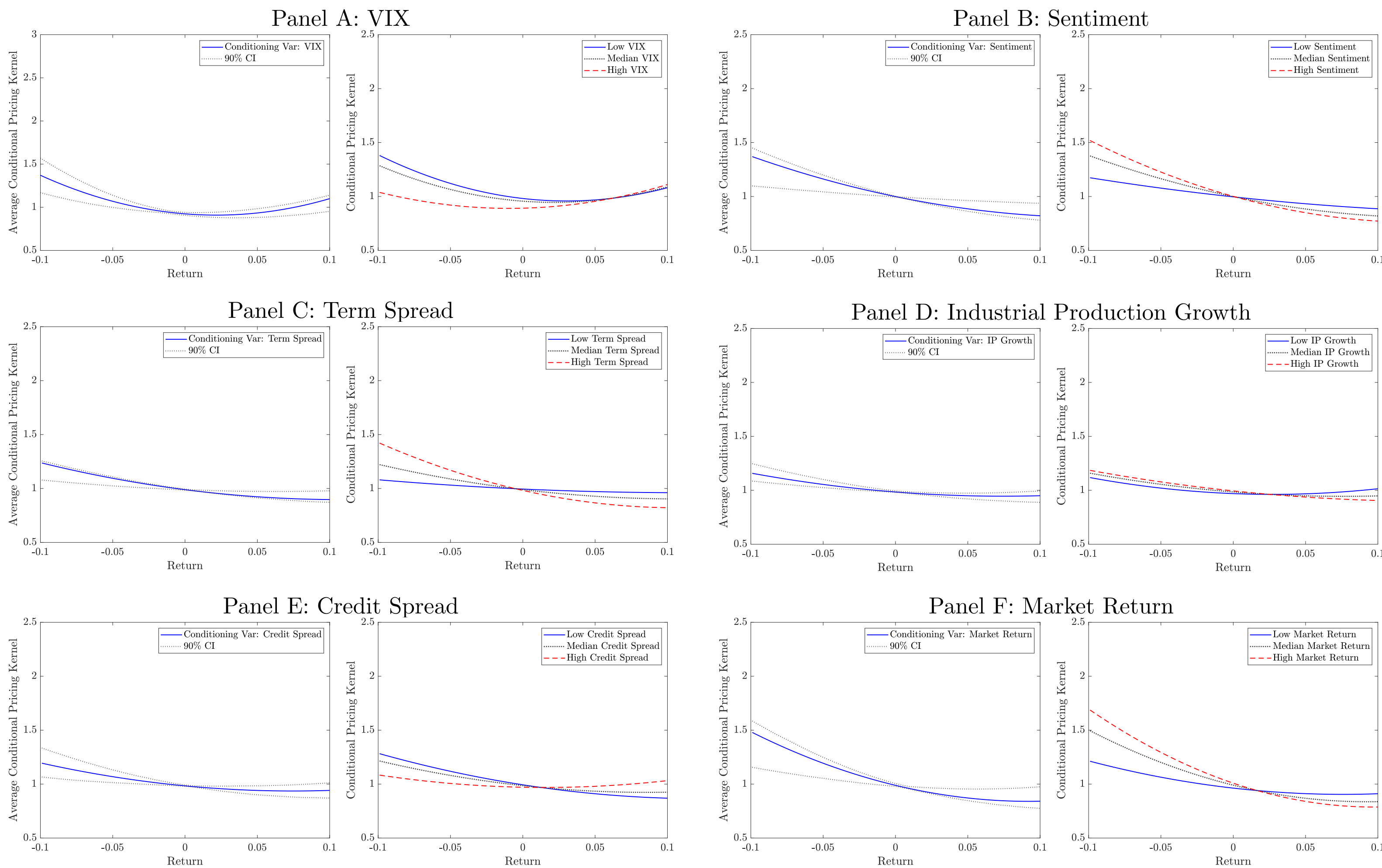


Panel C: Polynomial EPK with the P-Density Restriction and Euler Equations



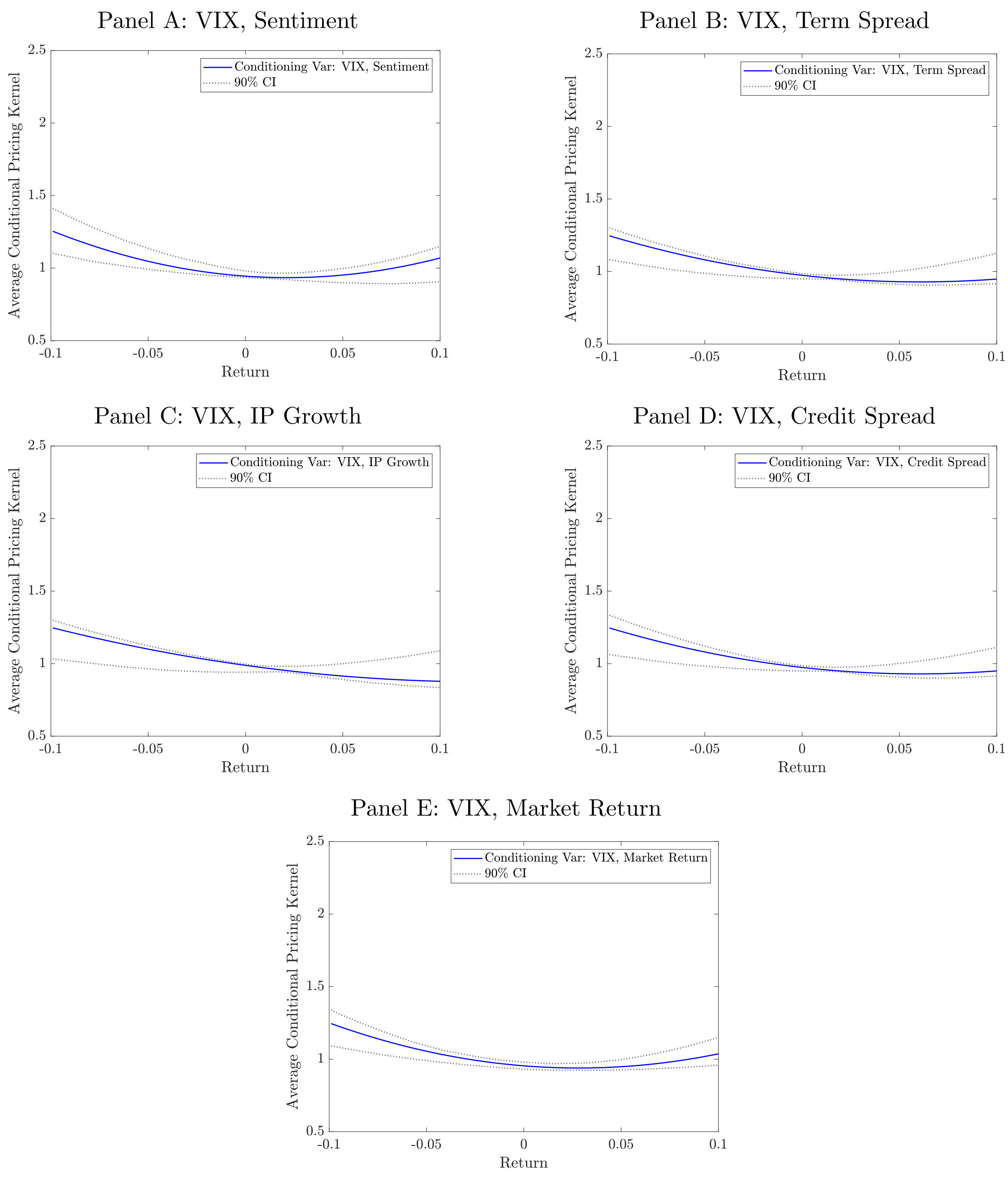
Notes: This figure shows the forward-looking unconditional empirical pricing kernels (EPKs). Panel A follows the CDI method of Linn, Shive, and Shumway (2018), Panel B describes the unrestricted polynomial pricing kernel estimation, and Panel C describes the polynomial pricing kernel estimation considering the P-density restriction and Euler equations. The left side of each panel presents the empirical pricing kernel (solid blue line) with its point-wise bootstrapped 90% confidence interval (dotted black lines). The right side of each panel presents the histogram of the integral of the physical densities over the support.

Figure 4: Univariate Conditional Empirical Pricing Kernels



Notes: This figure shows the forward-looking conditional empirical pricing kernels with one of six conditioning variables: the VIX (Panel A), sentiment (Panel B), term spread (Panel C), industrial production growth (Panel D), credit spread (Panel E), and market return (Panel F). The left side of each panel is the average empirical conditional pricing kernel (solid blue line) with its point-wise bootstrapped 90% confidence interval (dotted black lines). The right side of each panel presents how the conditional pricing kernel changes when each conditioning variable varies from the 5th percentile to the 95th percentile of its data.

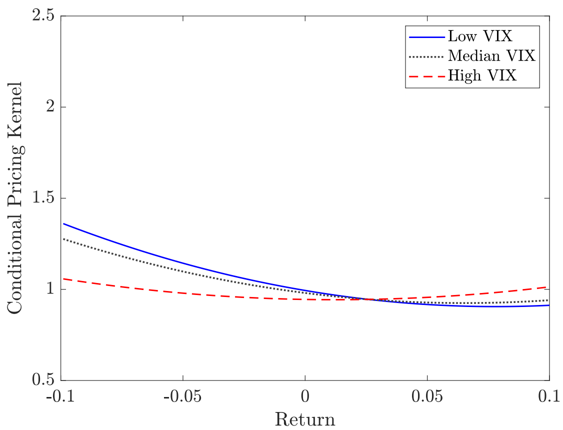
Figure 5: Bivariate Conditional Empirical Pricing Kernels

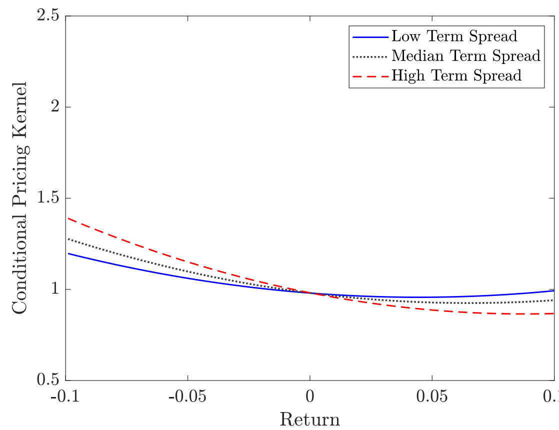


Notes: This figure shows the average forward-looking conditional empirical pricing kernels with five sets of two conditioning variables: the VIX and sentiment (Panel A), the VIX and term spread (Panel B), the VIX and industrial production growth (Panel C), the VIX and credit spread (Panel D), and the VIX and market return (Panel E). The dotted black lines are the point-wise bootstrapped 90% confidence interval.

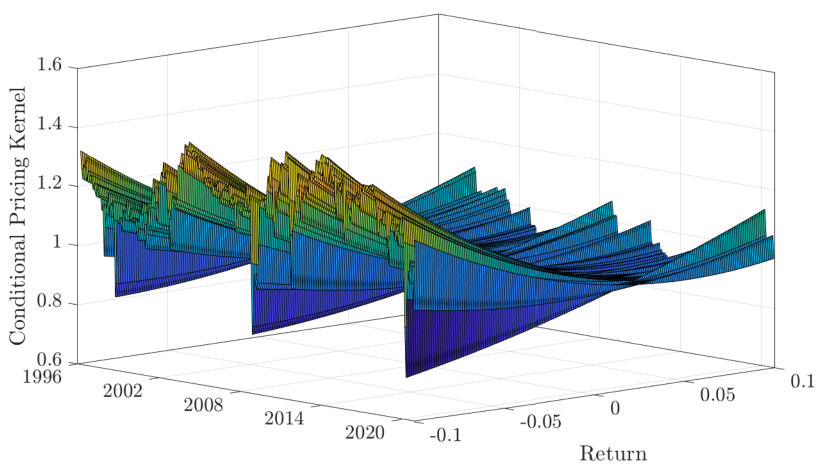
Figure 6: Characterizing Variation in the Bivariate Conditional Empirical Pricing Kernel

Panel A: Sensitivity to Each Conditioning Variable



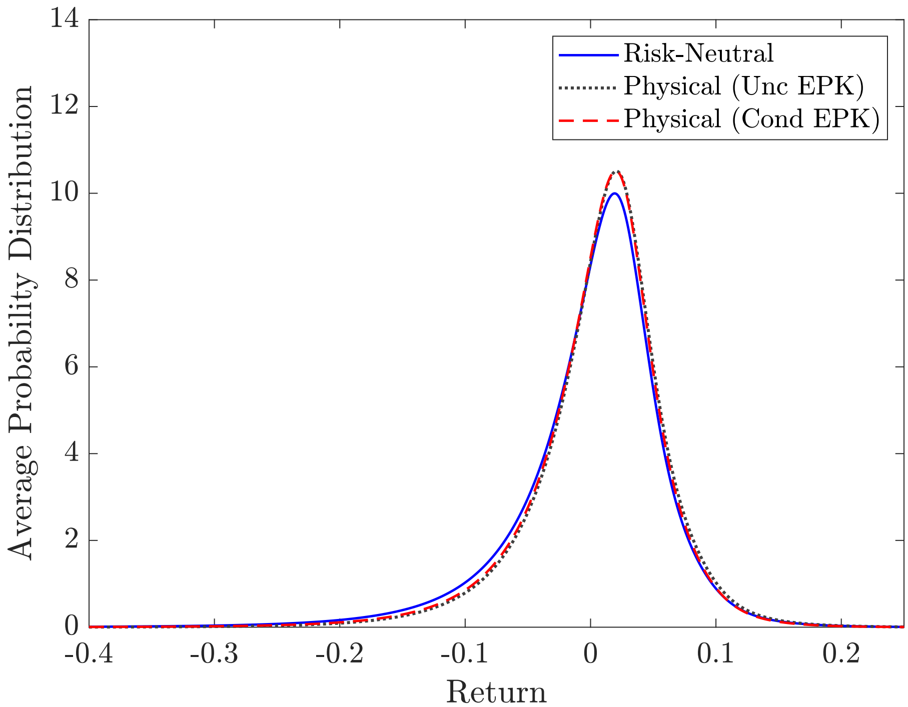


Panel B: Time Series of the Conditional Pricing Kernel



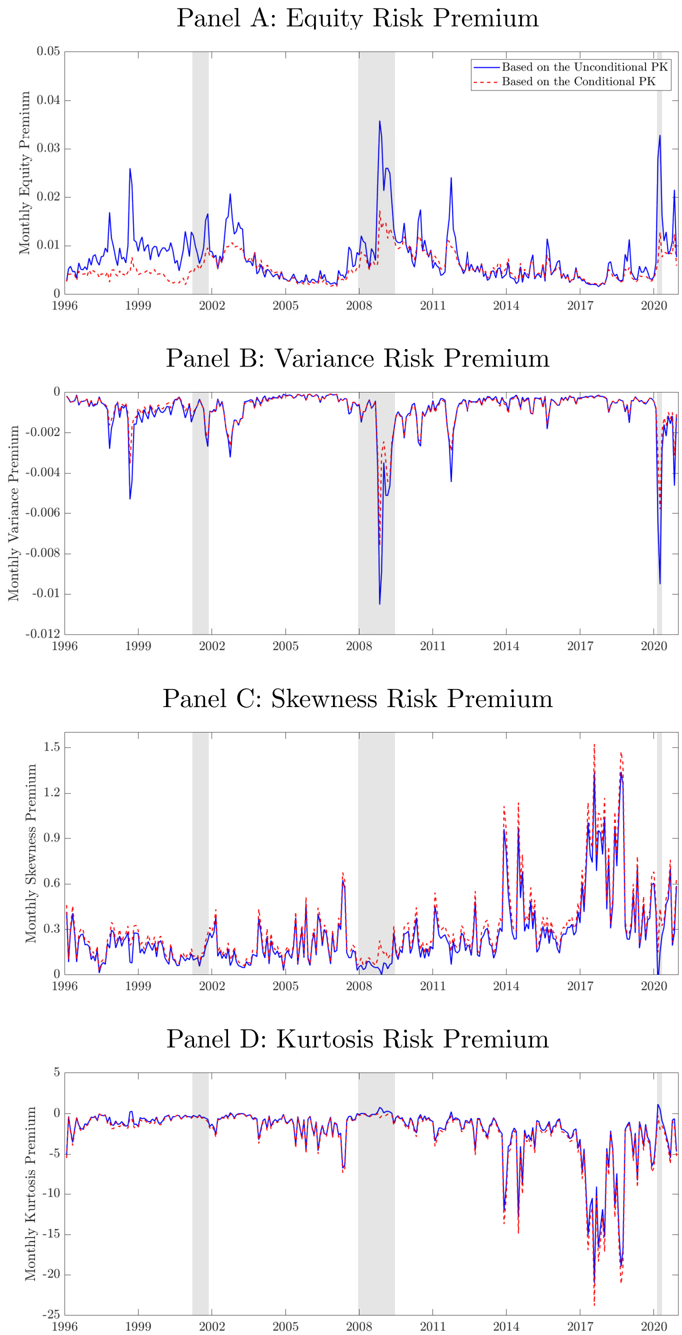
Notes: Panel A of this figure shows how the bivariate (the VIX and the term spread) conditional empirical pricing kernel changes when the VIX (left) or the term spread (right) varies from the 5th percentile to the 95th percentile of its data while the other variable is fixed at its median value. Panel B describes the time series of the conditional pricing kernel for the January 1996 to December 2020 sample period.

Figure 7: Average Probability Distributions



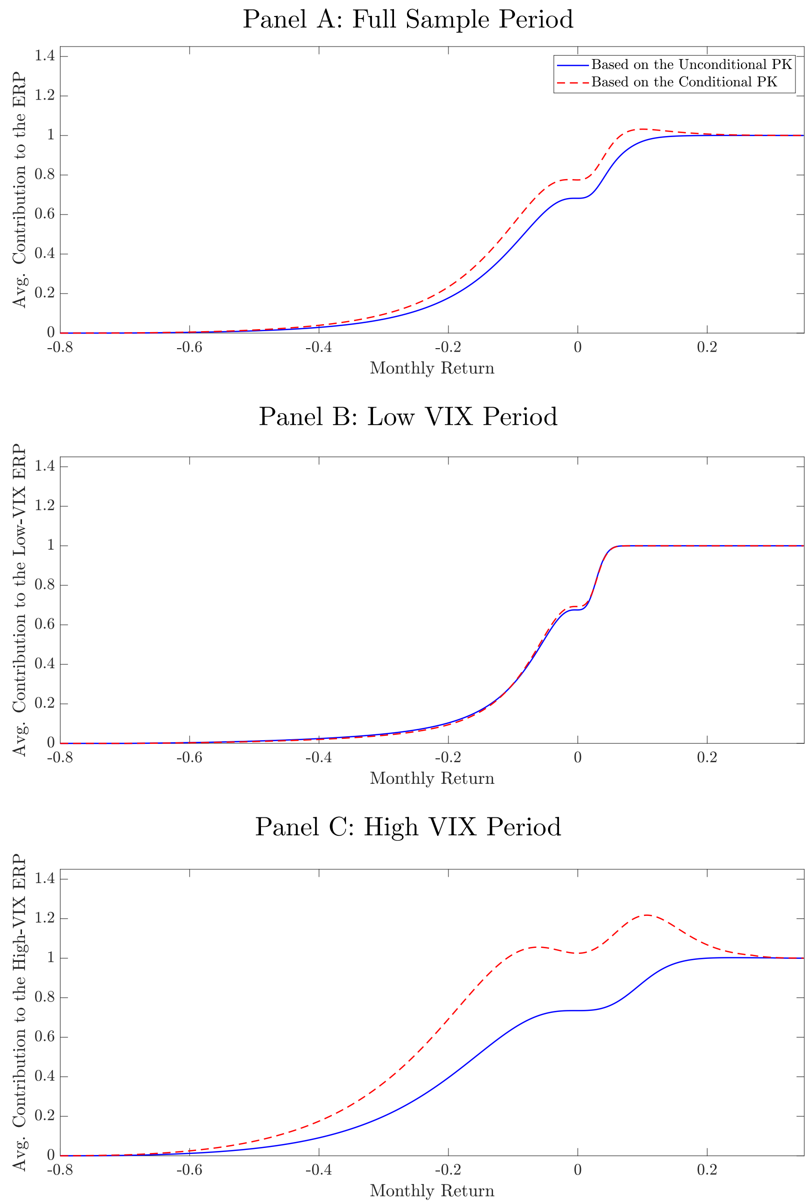
Notes: This figure shows the average risk-neutral (solid blue line) and physical probability densities. The dotted black line indicates the physical density implied by the unconditional empirical pricing kernel, and the dashed red line represents the one implied by the conditional empirical pricing kernel with two conditioning variables, the VIX and term spread.

Figure 8: Conditional Monthly Risk Premia



Notes: This figure describes the time series of the forward-looking conditional monthly risk premia inferred from the unconditional (solid blue line) and conditional (dashed red line) pricing kernel estimates. Panels A through D present the equity, variance, skewness, and kurtosis risk premia, respectively. The sample period is from January 1996 to December 2020.

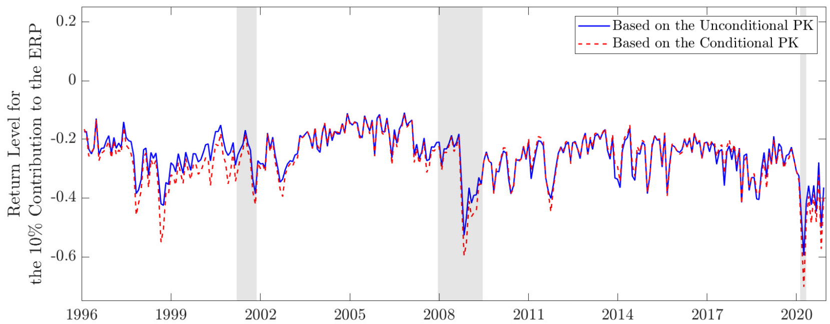
Figure 9: The Sources of the Conditional Equity Risk Premium



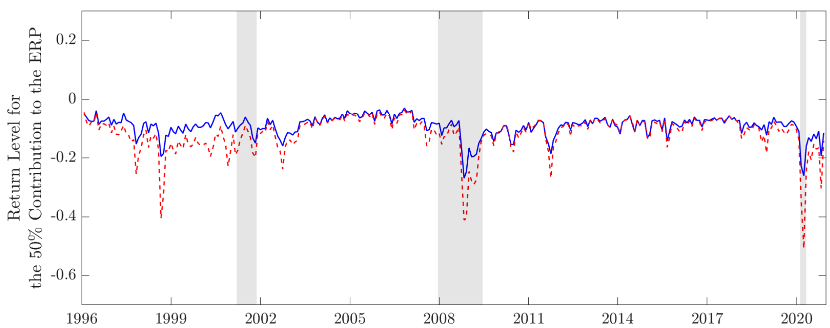
Notes: This figure shows the sources of the conditional equity risk premium during the full sample period (Panel A), low-VIX period (Panel B), and high-VIX period (Panel C). The low-VIX (high-VIX) period consists of the subsample where the VIX is lower (higher) than or equal to the 10th (90th) percentile of the VIX data. In each panel, the average contributions to the ERP based on the unconditional (solid blue line) and conditional (dashed red line) pricing kernel estimates are plotted.

Figure 10: The Sources of the Conditional Equity Risk Premium: Time Variation

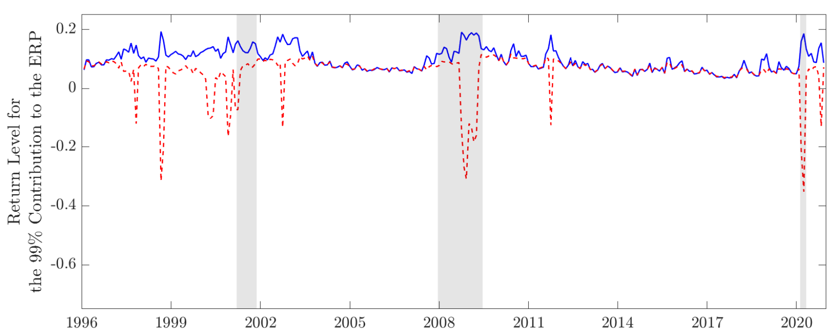
Panel A: 10% Contribution



Panel B: 50% Contribution



Panel C: 99% Contribution



Notes: This figure describes the time series of the sources of the conditional equity risk premium based on the unconditional (solid blue line) and conditional (dashed red line) pricing kernel estimates. Panels A through C present the return levels for the 10%, 50%, and 99% contributions to the ERP, respectively. The sample period is from January 1996 to December 2020.

Table 1: Conditioning Variables: Descriptive Statistics

	VIX	Skewness	Kurtosis	LeftTailRisk	Sentiment	TermSpread	CFNAI	ADS	IPGrowth	ConsGrowth	CreditSpread	MktReturn
Mean	0.204	-2.470	17.485	6.862	0.223	0.011	-0.093	-0.260	0.001	0.001	0.010	0.009
StdDev	0.080	1.037	14.377	2.814	0.139	0.009	1.259	2.089	0.011	0.010	0.004	0.046
Correlation												
VIX	1.000											
Skewness	0.358	1.000										
Kurtosis	-0.363	-0.961	1.000									
LeftTailRisk	0.559	-0.278	0.193	1.000								
Sentiment	-0.625	-0.499	0.504	-0.215	1.000							
TermSpread	0.152	0.085	-0.091	0.270	0.013	1.000						
CFNAI	-0.249	-0.093	0.101	-0.298	0.206	-0.070	1.000					
ADS	-0.356	-0.091	0.105	-0.341	0.254	-0.057	0.734	1.000				
IPGrowth	-0.189	-0.084	0.096	-0.120	0.166	-0.027	0.922	0.713	1.000			
ConsGrowth	-0.141	-0.027	0.045	-0.091	0.120	-0.027	0.893	0.750	0.762	1.000		
CreditSpread	0.552	0.180	-0.170	0.471	-0.403	0.350	-0.333	-0.304	-0.307	-0.117	1.000	
MktReturn	-0.394	-0.178	0.191	-0.262	0.436	-0.020	-0.014	0.191	-0.033	0.011	-0.094	1.000

Notes: This table summarizes the time series statistics (mean, standard deviation, and correlations) of 12 conditioning variables that are used in constructing the conditional pricing kernel. The conditioning variables include the VIX, risk-neutral skewness and kurtosis of Bakshi, Kapadia, and Madan (2003), left-tail risk index of Bollerslev, Todorov, and Xu (2015), Investors Intelligence sentiment index, 10-year minus 2-year Treasury term spread, CFNAI, ADS index, industrial production growth, real consumption growth, credit spread (difference between Moody's Baa and Aaa 20-year corporate bond yields) and S&P 500 market return.

Table 2: Unconditional Empirical Pricing Kernels: Parameter Estimates

Max Order (n)	$\beta_{0,0}$	$\beta_{1,0}$	$\beta_{2,0}$	$\beta_{3,0}$	$\beta_{4,0}$	Wald Statistic
4	0.99	-1.86	6.91	-0.00	0.00	0.00
	(28.66)	(-0.81)	(0.30)	(-0.00)	(0.00)	[1.000]
3	0.99	-2.22	9.49	-0.00		0.00
	(128.56)	(-3.33)	(1.53)	(-0.00)		[1.000]
2	0.98	-2.25	13.00			3.59
	(120.81)	(-3.28)	(1.90)			[0.058]
1	1.01	-1.24				
	(226.79)	(-19.46)				

Notes: This table reports the parameter estimates of the unconditional pricing kernel with a different maximum order (n) of polynomials. The maximum order varies from four to one. The GMM t-statistics are displayed in parentheses. The last column reports the Wald statistic for testing the restriction of $\beta_{n,0} = 0$. Its p-value is displayed in the bracket.

Table 3: Univariate Conditional Empirical Pricing Kernels: Parameter Estimates

	Parameter Estimates for the Time-Varying Coefficients						
	Intercept		$r_{t+\tau}$		$r^2_{t+\tau}$		
Cond. Var.	$\beta_{0,0}$	$\beta_{0,1}$	$\beta_{1,0}$	$\beta_{1,1}$	$\beta_{2,0}$	$\beta_{2,1}$	Hansen's J
VIX	1.03	-0.41	-2.50	8.51	28.78	-30.92	13.00
	(113.33)	(-10.45)	(-2.64)	(6.90)	(9.69)	(-3.87)	[0.112]
Skewness	1.00	-0.00	-0.60	0.38	0.85	-0.93	42.71
	(305.06)	(-0.28)	(-0.66)	(2.76)	(0.15)	(-0.36)	[0.000]
Kurtosis	0.97	0.00	-1.80	0.10	7.34	0.11	27.93
	(10.48)	(1.22)	(-6.99)	(8.71)	(2.71)	(0.65)	[0.001]
LeftTailRisk	1.03	-0.00	-4.72	0.30	12.42	-0.33	21.72
	(45.10)	(-8.00)	(-9.82)	(9.14)	(2.17)	(-0.97)	[0.006]
Sentiment	1.00	0.01	-1.65	-4.97	4.43	24.27	17.37
	(67.66)	(0.56)	(-3.61)	(-3.33)	(2.07)	(3.80)	[0.027]
TermSpread	0.99	-0.47	-0.69	-91.67	3.06	437.80	15.22
	(232.69)	(-2.31)	(-1.64)	(-1.83)	(2.17)	(1.78)	[0.055]
CFNAI	0.98	-0.00	-0.84	-0.15	6.29	0.32	33.62
	(28.76)	(-1.33)	(-2.11)	(-1.34)	(8.17)	(1.53)	[0.000]
ADS	1.01	0.00	-3.04	-0.29	6.78	-0.03	25.86
	(58.61)	(0.68)	(-4.90)	(-9.74)	(1.68)	(-0.18)	[0.001]
IPGrowth	0.98	0.81	-1.02	-28.06	7.22	-147.73	16.78
	(221.89)	(2.52)	(-1.70)	(-2.33)	(7.47)	(-1.95)	[0.032]
ConsGrowth	0.98	1.61	-1.61	-33.08	8.65	-130.20	22.07
	(20.59)	(4.39)	(-1.96)	(-7.55)	(4.32)	(-0.99)	[0.005]
CreditSpread	1.01	-2.60	-3.34	208.25	8.11	44.85	16.61
	(124.13)	(-10.95)	(-3.57)	(7.92)	(2.76)	(0.22)	[0.034]
MktReturn	0.99	0.30	-3.05	-19.45	16.74	85.51	17.72
	(185.25)	(2.15)	(-8.63)	(-5.77)	(8.29)	(2.93)	[0.024]

Notes: This table reports the parameter estimates of the conditional pricing kernel with each of the conditioning variables from Table 1. The GMM t-statistics are displayed in parentheses. The last column reports Hansen's J statistic for the over-identification test as well as its p-value in the bracket.

Table 4: Multivariate Conditional Empirical Pricing Kernels: Parameter Estimates

Conditioning Variables			Intercept				$r_{t+\tau}$				$r^2_{t+\tau}$						
$z_{1,t}$	$z_{2,t}$	$z_{3,t}$	$\beta_{0,0}$	$\beta_{0,1}$	$\beta_{0,2}$	$\beta_{0,3}$	$\beta_{1,0}$	$\beta_{1,1}$	$\beta_{1,2}$	$\beta_{1,3}$	$\beta_{2,0}$	$\beta_{2,1}$	$\beta_{2,2}$	$\beta_{2,3}$	Hansen's J	Wald_{unc}	$\text{Wald}_{\text{cond}}$
Panel A. Two Conditioning Variables																	
VIX	Sent		1.06	-0.54	-0.02		-3.32	12.16	-0.50		21.29	-6.92	9.54		24.91	5,372.5	0.80
			(264.80)	(-6.23)	(-0.33)		(-2.68)	(2.97)	(-0.14)		(2.67)	(-0.65)	(0.38)		[0.015]	[0.000]	[0.851]
VIX	TS		1.02	-0.22	0.05		-2.72	9.25	-60.54		16.03	-24.22	127.50		20.72	1,219.5	15.24
			(99.10)	(-2.99)	(0.05)		(-3.95)	(3.42)	(-1.67)		(5.97)	(-3.31)	(1.13)		[0.055]	[0.000]	[0.002]
VIX	IPGr		1.02	-0.17	0.24		-3.53	8.19	-1.51		10.46	-13.78	-51.37		22.40	1,777.7	8.79
			(144.22)	(-8.49)	(0.85)		(-9.10)	(8.60)	(-0.09)		(3.85)	(-2.16)	(-2.47)		[0.033]	[0.000]	[0.032]
VIX	CS		1.04	-0.19	-2.31		-3.92	7.60	86.78		15.74	-31.07	324.03		26.38	5,392.6	0.97
			(139.66)	(-1.35)	(-0.95)		(-7.98)	(1.79)	(0.90)		(4.93)	(-2.00)	(0.83)		[0.009]	[0.000]	[0.810]
VIX	MktR		1.05	-0.45	-0.19		-2.98	9.37	-0.33		15.58	12.90	80.43		22.62	741.4	8.77
			(120.62)	(-9.63)	(-0.85)		(-2.82)	(5.38)	(-0.03)		(4.03)	(1.07)	(1.94)		[0.031]	[0.000]	[0.033]
Panel B. Three Conditioning Variables																	
VIX	TS	IPGr	1.02	-0.26	-0.32	0.06	-2.84	8.12	-11.81	-18.52	21.14	-33.46	66.28	-17.53	33.82	5,897.7	5.61
			(215.00)	(-5.49)	(-0.48)	(0.21)	(-4.85)	(3.61)	(-0.24)	(-1.73)	(8.15)	(-4.16)	(0.50)	(-0.28)	[0.006]	[0.000]	[0.132]
VIX	TS	MktR	1.03	-0.23	-0.03	-0.15	-3.41	8.85	-14.34	-0.01	7.98	6.01	-58.81	52.30	35.00	1,442.1	6.02
			(99.54)	(-4.02)	(-0.04)	(-1.00)	(-4.91)	(4.22)	(-0.33)	(-0.00)	(2.65)	(0.59)	(-0.31)	(2.85)	[0.004]	[0.000]	[0.111]

Parameter Estimates for the Time-Varying Coefficients (spans the Intercept, $r_{t+\tau}$ and $r^2_{t+\tau}$ columns)

Notes: This table reports the parameter estimates of the multivariate conditional pricing kernel. Panel A includes the results for five sets of two conditioning variables: (1) the VIX and sentiment, (2) the VIX and term spread, (3) the VIX and industrial production growth, (4) the VIX and credit spread, and (5) the VIX and market return. Panel B includes the results for two sets of three conditioning variables: (6) the VIX, term spread, and industrial production growth, and (7) the VIX, term spread, and market return. The GMM t-statistics are displayed in parentheses. Hansen's J statistic for testing the over-identification and the Wald statistics for testing parameter restrictions, as well as their p-values, are also reported. Wald_{unc} is for comparing each conditional estimate with the unconditional estimate, *i.e.*, testing $\beta_{i,1} = \beta_{i,2}$ $(= \beta_{i,3}) = 0$. $\text{Wald}_{\text{cond}}$ is for testing if each additional conditioning variable is statistically informative, *i.e.*, testing $\beta_{i,2} = 0$ in Panel A and $\beta_{i,3} = 0$ in Panel B.

Table 5: Monthly Risk-Neutral and Physical Distributions: Descriptive Statistics

	Mean (%)	StdDev (%)	Skewness	Kurtosis
Panel A. Risk-Neutral Distribution				
Time-Series Mean	0.21	5.79	-1.64	10.01
Time-Series StdDev	0.18	2.22	0.73	7.58
Quantile (5%)	0.02	3.29	-3.12	4.04
Quantile (50%)	0.16	5.37	-1.53	7.63
Quantile (95%)	0.49	9.97	-0.75	26.28
Panel B. Physical Distribution (from the Unconditional EPK)				
Time-Series Mean	0.96	5.00	-1.38	7.84
Time-Series StdDev	0.58	1.76	0.51	4.36
Quantile (5%)	0.33	2.88	-2.39	4.01
Quantile (50%)	0.84	4.63	-1.33	6.58
Quantile (95%)	2.06	8.28	-0.67	16.61
Panel C. Physical Distribution (from the Conditional EPK)				
Time-Series Mean	0.84	5.03	-1.33	7.49
Time-Series StdDev	0.28	1.94	0.49	4.00
Quantile (5%)	0.41	2.82	-2.30	3.84
Quantile (50%)	0.85	4.62	-1.28	6.37
Quantile (95%)	1.30	8.69	-0.64	15.63

Notes: This table reports the descriptive statistics for the conditional risk-neutral and physical densities. Panel A is for the risk-neutral distribution, Panel B is for the physical distribution inferred from the unconditional empirical pricing kernel, and Panel C is for the physical distribution inferred from the conditional empirical pricing kernel. The conditional mean (%), standard deviation (%), skewness, and kurtosis are first calculated, and their time series means, standard deviations, and quantile distributions are displayed.

Table 6: Monthly Risk Premia: Descriptive Statistics

	ERP (%)	VRP (%^{2})	SRP	KRP
Panel A. Unconditional Empirical Pricing Kernel				
Time-Series Mean	0.76	-10.36	0.26	-2.17
Time-Series StdDev	0.55	13.24	0.23	3.29
Quantile (5%)	0.24	-30.36	0.06	-9.47
Quantile (50%)	0.61	-6.19	0.20	-1.04
Quantile (95%)	1.83	-2.07	0.76	0.01
AR1	0.79	0.70	0.70	0.71
Panel B. Conditional Empirical Pricing Kernel				
Time-Series Mean	0.63	-9.35	0.31	-2.53
Time-Series StdDev	0.29	8.95	0.25	3.65
Quantile (5%)	0.29	-26.14	0.08	-10.69
Quantile (50%)	0.57	-6.47	0.24	-1.32
Quantile (95%)	1.20	-2.43	0.84	-0.15
AR1	0.85	0.69	0.69	0.71

Notes: This table reports the descriptive statistics for the conditional risk premia. Panel A is based on the unconditional pricing kernel estimate, and Panel B is based on the conditional pricing kernel estimate. The conditional equity, variance, skewness, and kurtosis risk premia are first calculated, and their time series means, standard deviations, quantile distributions, and first-order autocorrelation (AR1) are displayed.

Table 7: Out-of-Sample Evaluation: Option Returns

Panel A. Unconditional Empirical Pricing Kernel						
Option Return						
	Put			Call		
Delta	-0.30	-0.50	-0.70	0.30	0.50	0.70
Avg. error	-0.413	-0.231	-0.143	-0.178	-0.089	-0.049
t-stat	(-4.16)	(-2.73)	(-2.12)	(-2.14)	(-1.53)	(-1.22)
Delta-Hedged Portfolio Return						
	Put			Call		
Delta	-0.30	-0.50	-0.70	0.30	0.50	0.70
Avg. error	-0.011	-0.011	-0.011	-0.009	-0.009	-0.010
t-stat	(-1.82)	(-1.68)	(-1.65)	(-1.37)	(-1.42)	(-1.54)
Panel B. Conditional Empirical Pricing Kernel						
Option Return						
	Put			Call		
Delta	-0.30	-0.50	-0.70	0.30	0.50	0.70
Avg. error	-0.446	0.267	-0.175	-0.130	-0.046	-0.014
t-stat	(-4.79)	(-3.34)	(-2.73)	(-1.47)	(-0.74)	(-0.33)
Delta-Hedged Portfolio Return						
	Put			Call		
Delta	-0.30	-0.50	-0.70	0.30	0.50	0.70
Avg. error	-0.007	-0.007	-0.007	-0.005	-0.005	-0.005
t-stat	(-1.64)	(-1.53)	(-1.53)	(-1.10)	(-1.14)	(-1.24)

Notes: This table reports the out-of-sample fit of option returns and delta-hedged option portfolio returns. Panel A is based on the unconditional pricing kernel estimate, and Panel B is based on the conditional pricing kernel estimate. In each panel, the time-series average of the Euler equation errors is reported for options with delta of -0.3, -0.5, and -0.7 for puts and 0.3, 0.5, and 0.7 for calls. The t-statistics are displayed in parentheses.

Appendix

A Risk-Neutral Density Estimation

I show detailed steps to construct the risk-neutral density. At each time t, the given data are the S&P 500 index price, traded OTM call and put options (prices, implied volatilities, strike prices, and maturities), and maturity-specific risk-free rates.

For this exercise, interpolation/extrapolation of the implied volatilities is necessary. This is because 1) call option prices at fine strike price grids are required to differentiate the prices numerically, 2) after filtering options data, some dates may not contain sufficient data to generate tails of the density, and 3) I focus on the 30-day horizon investment, but at each time, 30-day maturity options are not always observed.

Note that the traded data consists of option contracts with strike prices of every \$5. However, the interval of \$5 is not small enough to numerically calculate the derivative. I use the following grid points: \$1-spaced strike prices covering the moneyness range between $\min\{0.75, \min_i\{K_i/S_t\}\}$ and $\max\{1.25, \max_i\{K_i/S_t\}\}$, where $\{K_i/S_t\}$ is the set of moneyness in the data at time t. Then, I interpolate (and extrapolate for some data points) the 30-day implied volatilities with respect to moneyness and maturity over the moneyness grid points by following the second-degree polynomial regression method of Seo and Wachter (2019). Specifically, I run the regression below and find the fitted implied volatilities over the grid points at each time t:

$$IV(K/S, \bar{T}) = \theta_0 + \theta_1(K/S) + \theta_2(K/S)^2 + \theta_3\bar{T} + \theta_4\bar{T}^2 + \theta_5(K/S)\bar{T} + \theta_6(K/S)\bar{T}^2 + \epsilon_{K,\bar{T}},$$

where $\bar{T}$ is maturity. For applying equation (2) to find the risk-neutral densities, I convert the fitted implied volatilities into the Black-Scholes call option prices. After obtaining the risk-neutral density estimates, I test if they are well estimated in terms of the probability densities: the sum of the probabilities over the support should be equal to one. Figure A.1 verifies that the risk-neutral densities are estimated well enough.

B Existing Approaches to Conditional Physical Density

As discussed in Section 2.2, I estimate the conditional physical density with four different approaches. First of all, for kernel density estimation with the historical data, I follow Jackwerth (2000). That is, I use the Gaussian kernel density and the past 48 months' return data to estimate the conditional physical density at the end of each month.

For the Heston and Nandi (2000) GARCH dynamics, I follow the approach of Christoffersen, Heston, and Jacobs (2013) with monthly data. More specifically, a discrete-time physical return dynamics is assumed to follow

$$\begin{aligned} \ln S_{t+\tau} &= \ln S_t + r_{f,t}\tau + \left(\mu - \frac{1}{2}\right) h_{t+\tau} + \sqrt{h_{t+\tau}}\epsilon_{t+\tau}, \\ h_{t+\tau} &= \omega + \beta h_t + \alpha\left(\epsilon_t - \lambda\sqrt{h_t}\right)^2. \end{aligned}$$

I find the model parameter estimates via maximum likelihood estimation (MLE). To get the physical density, the distribution of $\epsilon_{t+\tau}$ needs to be specified, and the historical series of the monthly shocks is used rather than the exact normal distribution. By letting $\bar{\mu}_t = r_{f,t}\tau + \left(\mu - \frac{1}{2}\right) h_{t+\tau}$, $\epsilon_{t+\tau} = \left[\ln\left(S_{t+\tau}/S_t\right) - \bar{\mu}_t\right]/\sqrt{h_{t+\tau}}$ is provided. Let E denote the set of $\epsilon_{t+\tau}$ in the entire sample. Then, at each time t, the conditional physical density is calculated as

$$f_t(R_{t+\tau}) = f_t(\exp(\bar{\mu}_t + \sqrt{h_{t+\tau}}E)).$$

As the third method to find the conditional density, I consider the Heston (1993) SV model. Its dynamics is given as

$$\begin{aligned} d\ln S_t &= \left[r_{f,t} + \left(\mu - \frac{1}{2}\right) v_t\right] dt + \sqrt{v_t} d\epsilon_{1,t}, \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}d\epsilon_{2,t}, \end{aligned}$$

where v_t is the latent stochastic volatility, and ϵ_1 and ϵ_2 are correlated with the correlation coefficient

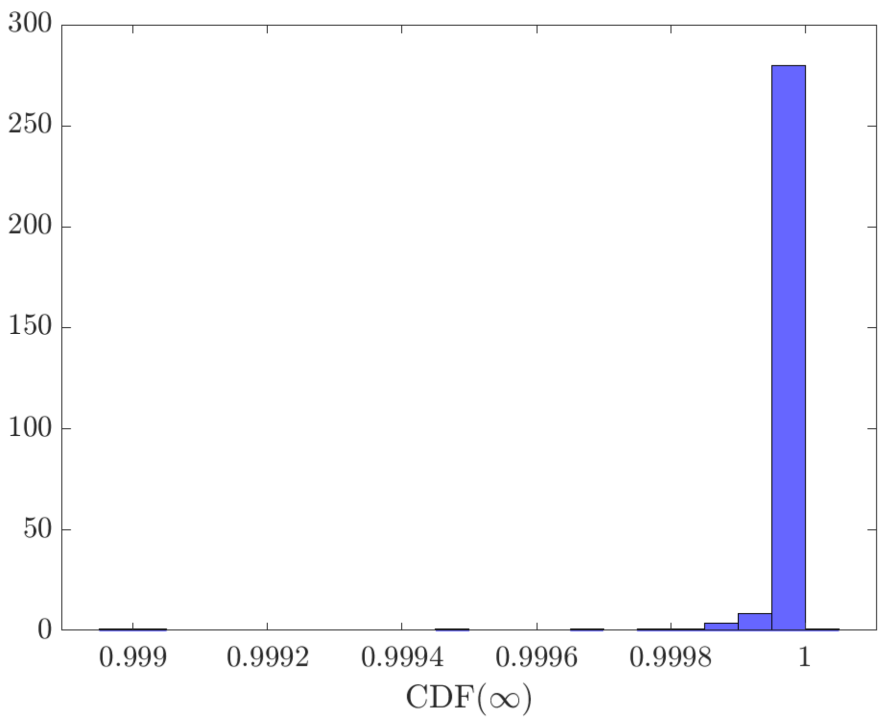
ρ. I apply the Euler discretization with $dt = 1/12$, and assume that the stochastic volatility is proxied by $v_t = \eta_0 + \eta_1 VIX_t^2$. I estimate the model by maximizing the likelihood function, and the rest of the procedure follows the same process as in the density estimation with the GARCH model.

For the stochastic volatility and jump model, I follow Bates (2000) and Pan (2002). The dynamics I use is

$$\begin{aligned} d\ln S_t &= \left[r_{f,t} + \left(\mu - \frac{1}{2}\right) v_t - \gamma_J \bar{\mu}_J\right] dt + \sqrt{v_t} d\epsilon_{1,t} + J dN_t, \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} d\epsilon_{2,t}, \end{aligned}$$

where the jump occurrence N_t follows a Poisson process with its constant intensity γ_J, and the jump size J follows a normal distribution with mean μ_J and variance σ_J^2. Then, the jump risk compensation is expressed as $\gamma_J \bar{\mu}_J$, where $\bar{\mu}_J = \exp\left(\mu_J + \frac{1}{2}\sigma_J^2\right) - 1$. Under this setup, I also proxy the stochastic volatility by a linear function of the VIX^2. However, the jump events are unobservable. In estimating the model via MLE, for simplicity, I assume that the maximum possible number of jump occurrences during the time interval dt is one. After the estimation, I run simulations 5,000 times to get the conditional return distribution.

Figure A.1: Validating the Risk-Neutral Density Estimates



Notes: This figure shows the histogram of the integrated values of the conditional risk-neutral density estimates over the support, where the integrated value is denoted by CDF(∞) on the x-axis. The number of samples in the histogram is 300, covering the monthly data from January 1996 to December 2020.