Characterizing the Conditional Pricing Kernel: A New Approach*

Hyung Joo Kim†

C. T. Bauer College of Business, University of Houston

November 9, 2021

Abstract

I propose a novel method to reliably estimate the conditional pricing kernel with the aid of conditioning variables. I find that the VIX, term spread, sentiment, and market return are conditioning variables that are informative about the kernel. The conditional pricing kernel estimate exhibits significant time variation: the more favorable future market expectations, the higher the kernel estimate in the negative future return region. During bad times, the conditional equity premium inferred from the conditional kernel estimate is entirely attributable to compensation for the left tail of the market return distribution. This observation is economically plausible and contrasts with findings for the unconditional kernel estimate. The out-of-sample return forecast based on the conditional kernel estimate outperforms the forecast based on the unconditional kernel estimate.

*I am extremely grateful to my committee chairs Hitesh Doshi and Kris Jacobs, and members Paola Pederzoli and Sang Byung Seo. I also thank Rohit Allena, Antonio Gargano, Mohammad Ghaderi, Sangeun Ha, Seung Joo Lee, Rauli Susmel, and seminar participants at the Virtual Derivatives Workshop and University of Houston for valuable comments and advice.

†C. T. Bauer College of Business, University of Houston. E-mail: hjkim@bauer.uh.edu

1 Introduction

The pricing kernel (or stochastic discount factor) is a critical 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. Extant studies have primarily focused on the characteristics of unconditional pricing kernel estimates or the time series average of conditional pricing kernel estimates, but less on conditional estimates. 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 stressed the importance of the conditional pricing kernel, as investors' conditional information and beliefs play crucial roles in pricing assets (see, e.g., Harvey, 1989; Nagel and Singleton, 2011).

In this paper, I show how to robustly estimate the conditional pricing kernel based on forward-looking measures without strong parametric model assumptions. I adapt the method of Linn, Shive, and Shumway (2018), who provide an unconditional pricing kernel estimate, by using conditioning variables that are informative about the kernel. My estimate of the conditional pricing kernel enables the exploration of its time series characteristics as well as the implications for conditional risk premia and out-of-sample predictability of stock returns in a unified framework.

I begin the estimation with 11 candidate conditioning variables that are potentially informative about the kernel. I use four risk measures: the VIX, model-free risk-neutral skewness and kurtosis of Bakshi, Kapadia, and Madan (2003), and left-tail risk index of Bollerslev, Todorov, and Xu (2015). I also use five economic indicators: the term spread, Chicago Fed National Activity Index (CFNAI), Aruoba, Diebold, and Scotti (ADS, 2009) index, industrial production, and real consumption growth. The remaining two variables are the investor sentiment index and stock market return. However, using all the conditioning variable candidates simultaneously in characterizing the conditional pricing kernel is not ideal. To filter the most informative variables, I first conduct a univariate analysis with each of the variables. Statistical tests show that the VIX, term spread, market return, and sentiment are the most informative variables. Then, I re-estimate the conditional pricing kernel using a multivariate specification. I examine the (average) shape of the conditional pricing kernel estimate, how its shape changes over time, how the realized pricing kernel varies over time, and the kernel estimates' economic properties and implications.

First, I show that the monotonicity of the average conditional kernel estimate is not statistically rejected. A large body of literature on empirical pricing kernels discusses the pricing kernel puzzle, which refers to the observation that empirical pricing kernels are U-shaped in the market return, as opposed to the traditional intuition that the kernel should be downward sloping.[1,2] Recently, Linn, Shive, and Shumway (2018) and Barone-Adesi, Fusari, Mira, and Sala (2020) refute the puzzle and empirically show the monotonicity of their forward-looking pricing kernels.[3] They argue that the puzzle appears when the pricing kernel estimation is not on a forward-looking basis. My empirical pricing kernel is also based on forward-looking measures, and its monotonicity is not rejected, consistent with the traditional intuition.

Second, I find significant time variation in the shape of the conditional pricing kernel estimate, specifically in the left-tail region of the market return. The results imply that investors' beliefs about negative future market returns are more sensitive to conditional information than the beliefs about positive future market returns. Moreover, the pricing kernel estimate in the negative return region is noticeably lower when a conditioning variable contains negative information about the future market than when it contains positive information. This indicates that for negative future returns, investors are less patient to consume when they have unfavorable information about the financial market or economy. Another interesting finding is that the realized pricing kernel estimates exhibit a business-cycle pattern and significantly react to financial market shocks. In contrast, these observations are absent in the realized pricing kernel implied by a cross-sectional factor model.

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

[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), and Driessen, Koëter, and Wilms (2019) for more extensive discussion of the pricing kernel puzzle.

[3]I use the term "forward-looking pricing kernel" when both the risk-neutral and physical densities are estimated in forward-looking measures.

My estimates of the time series of the risk-neutral densities, physical densities, and pricing kernels enable me to look into additional implications. Using the conditional risk-neutral and physical densities, I calculate the conditional equity, variance, skewness, and kurtosis risk premia inferred from the conditional pricing kernel estimate. I also compare these conditional measures with those inferred from an unconditional pricing kernel estimate unaccompanied by any conditioning variables. The result shows that the equity premium is captured better when conditional information is incorporated, whereas the other higher moments risk premia are similar for both approaches.

Third, I contribute to the literature on conditional asset pricing by investigating the time-varying sources of the conditional equity premium, inspired by Beason and Schreindorfer (2020) and Chabi-Yo and Loudis (2021).[4] 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 with the conditional pricing kernel estimate, which is able to capture a significant contribution. This analysis of the equity premium confirms the importance of conditional asset pricing and the conditional pricing kernel.

Fourth, the out-of-sample analysis for stock return forecasts also emphasizes the importance of conditional information. I compare the out-of-sample performances of various predictors following Goyal and Welch (2008) and Elliott, Gargano, and Timmermann (2013). Well-known predictors, such as the variance risk premium and dividend-price ratio, do not significantly outperform the conditional equity premium inferred from my conditional kernel estimate. More importantly, the out-of-sample performance of the conditional equity premium is significantly improved when it is inferred from the conditional kernel estimate rather than from the unconditional kernel estimate.

For implementation, my empirical methodology builds on that of Linn, Shive, and Shumway (2018). They use a statistical property of a cumulative distribution function (CDF) known as the universality of the uniform, in which a CDF, as a random variable, follows the uniform distribution between zero and one. The conditional physical density can be recovered using the risk-neutral

[4] They investigate the range of market returns for which investors require a large portion of the equity premium. I study the sources of the conditional equity premium to distinguish between economic implications from my conditional and unconditional pricing kernel estimates.

density estimates for a given pricing kernel, and the realized physical CDF values should unconditionally satisfy the moment conditions of the uniform distribution. Linn, Shive, and Shumway (2018) use a generalized method of moments (GMM) estimation based on these unconditional moments to find an unconditional pricing kernel estimate. They refer to this method as the conditional density integration (CDI) method.

My adaption of the CDI method helps circumvent two criticisms raised in the literature on estimating the projected pricing kernel on market returns. While the risk-neutral density can be reliably estimated using option prices, the physical density is obtained using historical return data or a parametric return dynamic.[5] When the estimation is based on the historical data, there is a timing mismatch between the physical and risk-neutral densities. The former is backward-looking, whereas the latter is forward-looking. Moreover, both the historical-data- and model-based conditional pricing kernel estimates highly depend on the specification. The CDI method is not subject to these criticisms because it uses options data to find the physical density and is independent of model assumptions.

I improve the CDI method by adding a constraint to satisfy a property of a probability density — the integral of the physical density over the domain should be equal to one. Despite the importance of this fundamental property for conditional physical density, the CDI method does not take it into account. This property must hold; otherwise, the relation among the risk-neutral density, physical density, and pricing kernel is violated. Hence, I include the constraint as a form of minimizing errors between the integrated values and one.[6] The inclusion of this constraint enables me to exploit the physical and risk-neutral densities for additional analyses such as studying the conditional risk premia.

I further modify the CDI method with the aid of observable economic or financial variables as

[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) also try to resolve this issue by normalizing the ratio of the risk-neutral density to the pricing kernel. However, although the normalization allows the ratio to be integrated to one, the relation among the risk-neutral density, physical density, and pricing kernel would be violated.

conditional information. Moreover, I approximate the pricing kernel by a higher-order polynomial function, which provides flexibility and tractability for incorporating the conditioning variables. Lastly, rather than including only the CDF's statistical moments in the GMM orthogonality conditions, I also include the Euler equations with the market return and the risk-free rate. The conditional empirical pricing kernel is then obtained through the GMM estimation.

This paper contributes to the estimation of the forward-looking conditional pricing kernel. Linn, Shive, and Shumway (2018) additionally estimate the pricing kernel using the VIX as a GMM instrumental variable or estimate it for two subperiods (high- and low-VIX periods). While they call these outcomes conditional estimates, my definition of conditional pricing kernel differs in the sense that I implement a time-varying pricing kernel. Barone-Adesi, Fusari, Mira, and Sala (2020) also estimate the time-varying pricing kernel by assuming a parametric return dynamic and then adjusting the physical density with options data.[7] Unlike their approach, however, my estimation does not assume a parametric return dynamic. In addition, my estimation approach in conjunction with conditioning variables allows me to easily examine the economic implications of the time variation of the pricing kernel.

My work is also related to the literature on calculating conditional risk premia. Studying conditional risk premia is vital, as investors' expectations are reflected in them rather than in historical risk premia. Prior studies suggest various methods to 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). Moreover, the conditional variance risk premium has been extensively studied.[8] In this paper, I provide a tractable and semi-nonparametric approach to calculate the conditional risk premia of any moments of returns by obtaining the risk-neutral and physical conditional return distributions.

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

[8]Some examples include, but are not limited to, 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).

Lastly, the findings of this paper are related to the broader literature on conditional factor models. Some important papers include Jagannathan and Wang (1996), Lettau and Ludvigson (2001), and Santos and Veronesi (2006). Their underlying assumption is that the pricing kernel is a function of future economic states and current state information, and this paper shares the same idea. However, I assume that the empirical pricing kernel is a nonlinear function of future market returns, whereas the above papers use a linear function of market returns. Moreover, by looking at economic implications for the equity premium, I show why conditional information is critical in asset pricing.

The rest of the paper proceeds as follows. Section 2 explains how the conditional pricing kernel can be estimated through existing approaches, and discusses their limitations. Section 3 describes the data, and Section 4 explains how I estimate the conditional pricing kernel in this paper. Section 5 reports the estimation results and discusses their implications. 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 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 Projected Pricing Kernel

In the projected pricing kernel calculation, the risk-neutral density can be 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.[9]

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. Using the call option prices over a wide range

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

of strike prices, one can 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.

2.2 Examples of Conditional Pricing Kernel Estimates

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.

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 for the different models 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. 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 use 11 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. As argued by Chabi-Yo (2012), higher-moment risks of future returns are other 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 measures, 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. Lastly, Chernov (2003) argues that the lagged market return contains important conditional information; hence, I include the 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 CDI method of Linn, Shive, and Shumway (2018), which is designed to estimate the unconditional pricing kernel using a GMM. I then explain how I adapt the CDI method by incorporating the conditioning variables and applying additional modifications. 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.

The 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.[10] 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.[11] I approximate the scaled pricing 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.[12]

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 three, and I then compare the results from the two approaches. 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

[10]For each $F_t(r)$, the conditional k-th moment should be $1/(1+k)$ by the properties of the uniform distribution. Then, equation (4) is obtained using the law of iterated expectation.

[11]For other studies that approximate the pricing kernel with a polynomial function, see, for instance, Rosenberg and Engle (2002) and Sandulescu and Schneider (2020).

[12]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.

right tail, consistent with Linn, Shive, and Shumway (2018). Appendix C shows that the result changes little for different choices of the degree of the polynomial function.

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. 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, my 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 unconditional moments of the conditional physical CDFs, as in equation (4), 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 an orthogonality condition as follows:

$$E\left[\left(\int_{-\infty}^{\infty}\frac{f_t^*(r_{t+\tau})}{m_{t,t+\tau}(r_{t+\tau})}dr_{t+\tau}-1\right)^2\right]=0. \tag{7}$$

In other words, the variance of the errors should be zero, implying that the errors should all be 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.[13] They are mathematically expressed as follows:

$$E\left[\left(M_{t,t+\tau}(r_{t+\tau})\begin{pmatrix}1+r_{t+\tau}\\1+r_{f,t}\end{pmatrix}-1\right)\otimes X_t\right]=0, \tag{8}$$

where X_t is the set of instrumental variables, consisting of the constant and conditioning variables, $\{z_{j,t}\}_{j=1}^m$.

The modified CDI method refers to the GMM estimation with all the moment conditions in equations (4), (7), and (8). In particular, I estimate the parameters $\{\beta_{i,j}\}$, which specify the conditional empirical pricing kernel, via the two-step efficient GMM. The number of parameters to be estimated is $(n+1)\times(m+1)$. The orthogonality conditions consist of $2\times(m+1)$ Euler equations, $(n+1)\times(m+1)-2\times(m+1)$ moment conditions in equation (4), and one constraint in equation (7). To keep the estimation as simple as possible, I use the degree of the polynomial

[13] 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.

$n = 3$ and the number of conditioning variables $m \leq 3$.

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. 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 (7). Henceforth, the unconditional pricing kernel estimate refers to the one estimated by this approach, *i.e.*, by the modified CDI method without conditioning variables.[14]

5 Estimation Results

Since I do not know which conditioning variables are most informative about the pricing kernel, 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 them. I use the term "multivariate conditional empirical pricing kernel" to indicate the resulting estimate. 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.

[14] 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.

5.1 Univariate Conditional Pricing Kernel Estimates

Table 2 summarizes the parameter estimates for the univariate conditional pricing kernel. Each row presents estimation results for each of the 11 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 two columns, I report the over-identification test statistic (J statistic) of Hansen (1982) and its p-value for each GMM estimation.

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 -0.68 and statistically significant. This result indicates that the overall level of the pricing kernel decreases by 0.068 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 the conditioning variables that are measures of risk (the VIX, risk-neutral 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 0.929 at the zero rate of return. When the sentiment, term spread, and market return are used as the conditioning variable, $\beta_{1,1}$ is negative and significant. When the expectation about the future financial market or economy (as measured by the sentiment or term spread) is favorable or when the current market conditions are good (as measured by the monthly market return), the slope of the pricing kernel becomes steeper.

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 5% level but rejected at the 10% level. When the pricing kernel estimate is conditional on the sentiment, term spread, industrial production, consumption growth, or market return, the J statistic is not rejected at the 1% level but at the 5% 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, consumption growth, 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, term spread, market return, and sentiment as conditioning variables that are most informative about the pricing kernel.

Figure 4 depicts the conditional empirical pricing kernels with a single conditioning variable. Panels A through D show the estimation results based on the VIX, term spread, market return, and sentiment, 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 6% 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 of Christoffersen, Heston, and Jacobs (2013). 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 with high VIX periods. 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 two conditioning variables (market return and sentiment).

5.2 Multivariate Conditional Pricing Kernel Estimates

The univariate conditional pricing kernel estimates in Table 2 show that the VIX, term spread, market return, and sentiment are state variables that are informative about the pricing kernel. Out of these four variables, I use a maximum of three conditioning variables in the multivariate estimation to avoid an abundance of free parameters. The VIX and term spread are the two most informative variables. I, therefore, consider the following two multivariate estimation specifications based on (1) the VIX, term spread, and market return, and (2) the VIX, term spread, and sentiment. Each specification contains 16 parameters.

Table 3 reports the resulting parameter estimates. Note that the parameter $\beta_{i,3}$ represents the coefficients of the market return in specification (1), but it represents the coefficients of the sentiment index in specification (2). The estimation results are similar for the two cases. The parameter estimate of $\beta_{1,0}$ (the constant component of the time-varying coefficient of R_m) is significantly negative in both specifications, implying that the unconditional component of the pricing kernel is decreasing around the central region of the return distribution. In addition, the estimates of $\beta_{1,1}$ and $\beta_{1,2}$ (sensitivities to the VIX and term spread) are significantly positive and negative, respectively, in both specifications (1) and (2). The signs of these parameter estimates are also consistent with the results from the univariate estimation, although their magnitudes are smaller. All coefficients of the market return and sentiment become insignificant in the multivariate estimations. These observations are presumably due to the fact that the market return and sentiment are negatively correlated with the VIX (-0.394 and -0.625, respectively), as seen in Table 1.

I conduct the Wald test to see if the conditional pricing kernel shows statistically better in-sample performance than the unconditional pricing kernel in which all coefficients of the conditioning variables are zero. I use the restriction of $\beta_{i,1} = \beta_{i,2} = \beta_{i,3} = 0$ for all i. In Table 3, the Wald statistics, as well as their p-values, show that conditioning variables improve the estimation of the pricing kernel at the 1% significance level in both specifications (1) and (2).

Figure 5 depicts the average conditional pricing kernel estimates (Panels A and C) as well as the average probability distributions (Panels B and D) for the two specifications. The two average conditional pricing kernels in Panels A and C are remarkably similar. They decrease up to the level of a future market return of 2% and increase slightly after that. The univariate estimation results in Panel A of Figure 4 suggest that the slight increase in the right tail is likely the effect of including the VIX as one of the conditioning variables. However, the increasing pattern is statistically insignificant at the 10% significance level. As anticipated, the average of both the risk-neutral (solid blue line) and the physical (dashed red line) densities are also similar to each other. For the remaining analyses and discussion about the conditional pricing kernel, I focus on specification (2) with the VIX, term spread, and sentiment as the conditioning variables, given that specifications (1) and (2) provide statistically indistinguishable results.[15] Additionally, this paper deals with the forward-looking estimate of the pricing kernel, and the VIX, term spread, and sentiment are all regarded as forward-looking measures, whereas the market return is not.

Figure 6 shows the results of the sensitivity analysis with respect to the three conditioning state variables, VIX, term spread, and sentiment. 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 2 and 3, Figures 4 and 6 graphically indicate that all the sensitivity patterns from the multivariate estimation are equivalent to those from the univariate estimation, but the differences between the pricing kernel estimates with low and high conditioning variables become smaller.

Figure 7 describes the time series of the multivariate 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 Figure 6. 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 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 The Time Series of the Realized Pricing Kernel

The pricing kernel estimates allow me to construct the time series of the pricing kernel, $M_{t,t+\tau}(r_{t+\tau})$. Note that the empirical pricing kernel itself is a function of the future market return. Since the realized market return data ($r_{t+\tau}$) is available, the value of the realized empirical pricing kernel can be found at each time $t+\tau$. Figure 8 compares the time series for two realized pricing kernel estimates with two kernels based on factor models. Panel A is the result inferred from the forward-looking unconditional empirical pricing kernel, and Panel B is that from the forward-

[15] Albeit not reported, all other analyses and implications remain the same even if I use specification (1).

looking conditional empirical pricing kernel. For comparison, Panel C presents the CAPM-implied pricing kernel, and Panel D is the Fama and French (1993) 3-factor-implied pricing kernel.[16]

Chernov (2003) and Ghosh, Julliard, and Taylor (2017) show that their nonparametrically filtered pricing kernels exhibit business-cycle patterns and significantly respond to financial market crashes.[17] Panels A and B in Figure 8 verify that the realized pricing kernels estimated by the modified CDI method also exhibit a similar business-cycle pattern and reaction to financial market downturns. The realized pricing kernels fluctuate more during the NBER recession periods (the shaded areas) compared with the other periods. Specifically, they abruptly spike around the 2008 financial crisis and the 2020 COVID-19 pandemic crisis. The pricing kernels also increase during the periods of the Long-Term Capital Management (LTCM) collapse in 1998, early 2000 recession, and stock market collapse in 2002. In contrast, the realized pricing kernels implied by the CAPM or Fama-French 3-factor model in Panels C and D do not exhibit these patterns. Granted, the empirical pricing kernels are projected pricing kernels. In other words, the time series graphs in Panels A and B do not reflect the component of the true pricing kernel that is orthogonal to the market return. This implies that once the orthogonal component is considered, the true realized pricing kernel could be more volatile than the projected kernels. Therefore, the implied pricing kernels from the factor models indeed fail to capture the time series properties of the kernel.

5.4 Implications: 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. The time series data of the risk-neutral and physical densities enable me to directly calculate the

[16] As in Cochrane (2001), a linear factor model can be converted to the pricing kernel as a linear function of the factors. I apply this method to the CAPM and the Fama-French 3-factor model. In order to estimate the prices of the factor risks, I use the 25 Fama-French Size and Book-to-Market portfolios as well as 10 industry portfolios for the cross-sectional test assets. The returns for the 35 portfolios are downloaded from Kenneth French's website.

[17] Chernov (2003) estimates asset-implied pricing kernels by specifying asset dynamics with two-factor stochastic variances. Ghosh, Julliard, and Taylor (2017) nonparametrically extract the time series of the pricing kernel via a relative entropy minimization. They both study the time series of realized pricing kernels but do not study how the shape of the pricing kernel estimates evolves over time.

s-th moment forward-looking risk premium at time t, $\mathrm{RP}_t^{(s)}$, as well as its statistical moments:

$$\mathrm{RP}_t^{(s)} = \begin{cases} E_t[r_{t+\tau}^s] - E_t^*[r_{t+\tau}^s] & \text{if} \quad s=1 \\ E_t[(r_{t+\tau}-\mu_t)^s] - E_t^*[(r_{t+\tau}-\mu_t^*)^s] & \text{if} \quad s=2 \\ \frac{E_t[(r_{t+\tau}-\mu_t)^s]}{\sigma_t^s} - \frac{E_t^*[(r_{t+\tau}-\mu_t^*)^s]}{(\sigma_t^*)^s} & \text{if} \quad s=3 \text{ or } 4, \end{cases} \tag{9}$$

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 4 summarizes the first four statistical moments of the risk-neutral (Panel A) and physical distributions. 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 should be close to the risk-free rate. Thus, it is overall lower than that of the physical density (from either the unconditional or conditional estimation). However, the magnitudes of the higher moments of the risk-neutral density exceed those of the physical densities. Moreover, when comparing Panel B with Panel C, the first moment's time series statistics are fairly distinct from each other. The time series mean and standard deviation of the first moment in Panel B are much higher than those in Panel C. The mean value of the first moment in Panel B, 1.10%, fails to match the sample average of the recent 25-year (from 1996 to 2020) monthly market return, 0.89%. The unconditional estimate of the pricing kernel seems to 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 (9). Table 5 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 4, the time series average of the equity risk premium (ERP) in Panel A (0.90%) is estimated to be higher than its data counterpart, which is 0.70%, while the average of the conditional ERP in Panel B (0.68%) is closer to 0.70%. Meanwhile, the variance risk premium (VRP) looks reasonable, and the time series means are -11.09%^{2} and -12.27%^{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 close to the values in Panels A and B.

Figure 9 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 5, 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, the difference increases up to nearly 5% per month. This large gap during bad times is the main reason for the high average ERP in Panel A of Table 5. This implies that if conditional information is not adequately taken into account, the implied conditional ERP may be exaggerated during bad times. Thus, the time series of the ERPs also motivate the use of conditional information.

5.5 Implications: the Sources of the Conditional Equity Premium

It is also interesting to investigate the conditional ERP from a slightly different angle, namely, from the perspective of the sources of the ERP. Following Beason and Schreindorfer (2020), I evaluate how much each future 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 10 shows the contribution of market 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.[18] The low-VIX (high-VIX) period consists of the subsample where the VIX is lower (higher) than or equal to the 5th (95th) 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 becomes smaller during the low-VIX period. However, the gap is significantly larger during the high-VIX period. If no conditional information is taken into account, a 99% contribution is met at around 25% of the market return. In contrast, a 99% contribution is achieved at around -10% 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 11 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

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

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 25% 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 area. 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 10 and 11 confirm the message that different implications for the conditional ERPs between the unconditional and conditional empirical pricing kernels are magnified during economically bad times.

5.6 Implications: Out-of-Sample Market Return Forecasts

My conditional equity premium inferred from either the unconditional or conditional empirical pricing kernel is forward-looking. It enables me to analyze the out-of-sample forecast of the stock market return. For comparison, existing studies find return forecasts using the following predictive regression:

$$r_{t+\tau} - r_{f,t} = \gamma_0 + \gamma_1 x_t + u_{t+\tau}, \tag{10}$$

where x_t is a predictor variable, and $u_{t+\tau}$ is an error term. After estimating equation (10) using in-sample data of $t = 0, \cdots, \bar{t} - \tau$, the next-period forecast of the excess return is given as $E_{\bar{t}}[r_{\bar{t}+\tau} - r_{f,\bar{t}}] = \hat{\gamma}_0 + \hat{\gamma}_1 x_{\bar{t}}$, where $\hat{\gamma}_0$ and $\hat{\gamma}_1$ are the in-sample parameter estimates. My conditional equity premium in equation (9) at time $t = \bar{t}$ also measures a return forecast for the next period.

I follow Goyal and Welch (2008) and Elliott, Gargano, and Timmermann (2013) for testing the out-of-sample performance of return forecasts. For regression-based forecasts, I use 12 predictor variables (x_t): the variance risk premium (VRP_{BTZ}) of Bollerslev, Tauchen, and Zhou (2009), dividend-price ratio (DP), dividend yield (DY), earnings-price ratio (EP), book-to-market ratio (BM), net equity expansion (NTIS), 3-month Treasury bill rate (TBL), long-term rate of returns on US government bonds (LTR), Treasury term spread (TMS), default yield spread (DFY), default return spread (DFR), and inflation (INFL).[19] I additionally include the two forward-looking conditional equity premium measures inferred from the unconditional and conditional empirical pricing kernels (ERP_{unc} and ERP_{cond}, respectively).

I adopt an expanding-window estimation for out-of-sample monthly return forecasts based on predictive regressions and my empirical pricing kernels. The in-sample period starts in January 1996. The out-of-sample period is from January 2011 to December 2020 (120 months). For instance, the evaluation period of January 1996 to December 2010 is used to find return forecasts for January 2011, and the evaluation period of January 1996 to November 2020 is used to find return forecasts for December 2020. The out-of-sample performance is measured by the out-of-sample R^2, which is defined as follows:

$$R^2(\%) = \left(1 - \frac{\sum_{t=T_0}^{T-\tau} \left(r_{t+\tau} - r_{f,t} - E_t[r_{t+\tau} - r_{f,t}]\right)^2}{\sum_{t=T_0}^{T-\tau} \left(r_{t+\tau} - r_{f,t} - E_t^{bench}[r_{t+\tau} - r_{f,t}]\right)^2}\right) \times 100, \tag{11}$$

where T_0 (T) represents December 2010 (December 2020), and $E_t^{bench}[\cdot]$ denotes the time-t conditional expectation under a benchmark approach. For benchmark forecasts, I use the historical average of the excess return, ERP_{unc}, and ERP_{cond}.

Table 6 reports the out-of-sample R^2 values as well as their associated p-values. Out of 14 forecasts, ERP_{cond} and forecasts using VRP_{BTZ}, DP, DY, BM, and TBL perform better than the historical average at the 10% significance level. When ERP_{unc} is used as a benchmark, the

[19] The time series of VRP_{BTZ} is obtained from Hao Zhou's website. The rest of the data is obtained from Amit Goyal's website. See Goyal and Welch (2008) for a more detailed description of the variable construction.

out-of-sample R^2 of ERP_{cond} measures its relative performance compared to ERP_{unc}. The R^2 is 2.45 with its p-value of 0.10, indicating that the conditional equity premium inferred from the conditional pricing kernel estimate has significantly smaller out-of-sample forecasting errors than the premium inferred from the unconditional pricing kernel estimate. Lastly, the analysis with ERP_{cond} as a benchmark shows that none of the regression-based forecasts perform statistically better than ERP_{cond}.

6 Conclusion

Accurate estimation of the pricing kernel is crucial to understanding asset prices and investors' beliefs. Moreover, although the literature emphasizes the importance of conditional information, the conditional pricing kernel has not been extensively studied because it is more challenging than estimating the unconditional pricing kernel. Prior studies estimate the conditional pricing kernel by relying on historical data or specific model assumptions, but the results strongly depend on the specification. In this paper, I propose a novel approach to semi-nonparametrically estimate the conditional empirical pricing kernel measured on a forward-looking basis.

I use conditioning variables to capture the time variation of the pricing kernel. To estimate the conditional pricing kernel, I suggest a modified CDI method that adapts and improves upon the CDI method proposed by Linn, Shive, and Shumway (2018), which is designed 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, term spread, market return, and sentiment are state variables that are informative about the 1-month forward pricing kernel. Using these selected conditioning variables, I also estimate multivariate conditional pricing kernels. The estimated kernels are, on average, downward sloping, and they show significant time variation which is consistent with economic intuition.

Using the realized monthly market return data, I obtain the time series of the realized pricing kernel inferred from either the unconditional or conditional pricing kernel estimates. It shows significant reactions to major financial market crashes, in contrast with pricing kernels based on well-known cross-sectional factor models, such as the CAPM and Fama-French 3-factor model.

Lastly, I study the conditional risk premia derived from the pricing kernel estimates with/without conditional information, and I 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 implies a more reasonable time series path of the equity premium. Furthermore, the out-of-sample return forecast based on the conditional kernel estimate performs statistically better than the forecast based on the unconditional kernel estimate. In conclusion, considering relevant conditional information is critical in understanding investors' beliefs and financial asset prices.

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 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 from $\min\{0.75, \min_i\{K_i/S_t\}\}$ to $\max\{1.25, \max_i\{K_i/S_t\}\}$, where $\{K_i/S_t\}$ is the set of moneyness in the data at time t. Then, I interpolate the 30-day implied volatilities with respect to the moneyness and maturity over the moneyness grid points by following the second-degree polynomial method of Seo and Wachter (2019). For applying equation (2) to find the risk-neutral densities, I convert the interpolated 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 $v_t = \eta_0 + \eta_1 VIX_t$. 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. 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.

C Robustness to the Degree of the Polynomial Function

One may ask if the pricing kernel estimation through the modified CDI method is sensitive to the choice of the degree of the polynomial function. As an example, I show that the unrestricted unconditional pricing kernel estimate is robust to different degrees of the polynomial function in Figure A.2. I compare the results with the orders from three to six, and the shape of the pricing kernel changes little. Although I do not report the robustness result for the conditional pricing kernel estimates here, I confirm that the result is robust to using the 4th degree of the polynomial.[20]

[20] I do not test the robustness with a polynomial function of a degree higher than four because an enormous number of parameters must be estimated when conditioning variables are incorporated.

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.

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, Tyle, and David Schreindorfer, 2020, On sources of risk premia in representative agent models, Working paper.

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.

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, 2021, A decomposition of conditional risk premia and implications for representative agent models, Working paper.

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.

Clark, Todd E., and Kenneth D. West, 2007, Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics* 138, 291–311.

Cochrane, John H., 2001, *Asset Pricing.* (Princeton University Press Princeton, NJ).

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, 2019, Behavioral in the short-run and rational in the long-run? evidence from s&p 500 options, Working paper.

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

Elliott, Graham, Antonio Gargano, and Allan Timmermann, 2013, Complete subset regressions, *Journal of Econometrics* 177, 357–373.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3–56.

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.

Ghosh, Anisha, Christian Julliard, and Alex P. Taylor, 2017, What is the consumption-CAPM missing? An information-theoretic framework for the analysis of asset pricing models, *Review of Financial Studies* 30, 442–504.

Goyal, Amit, and Ivo Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

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.

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

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

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.

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, Working paper.

Santos, Tano, and Pietro Veronesi, 2006, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1–44.

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

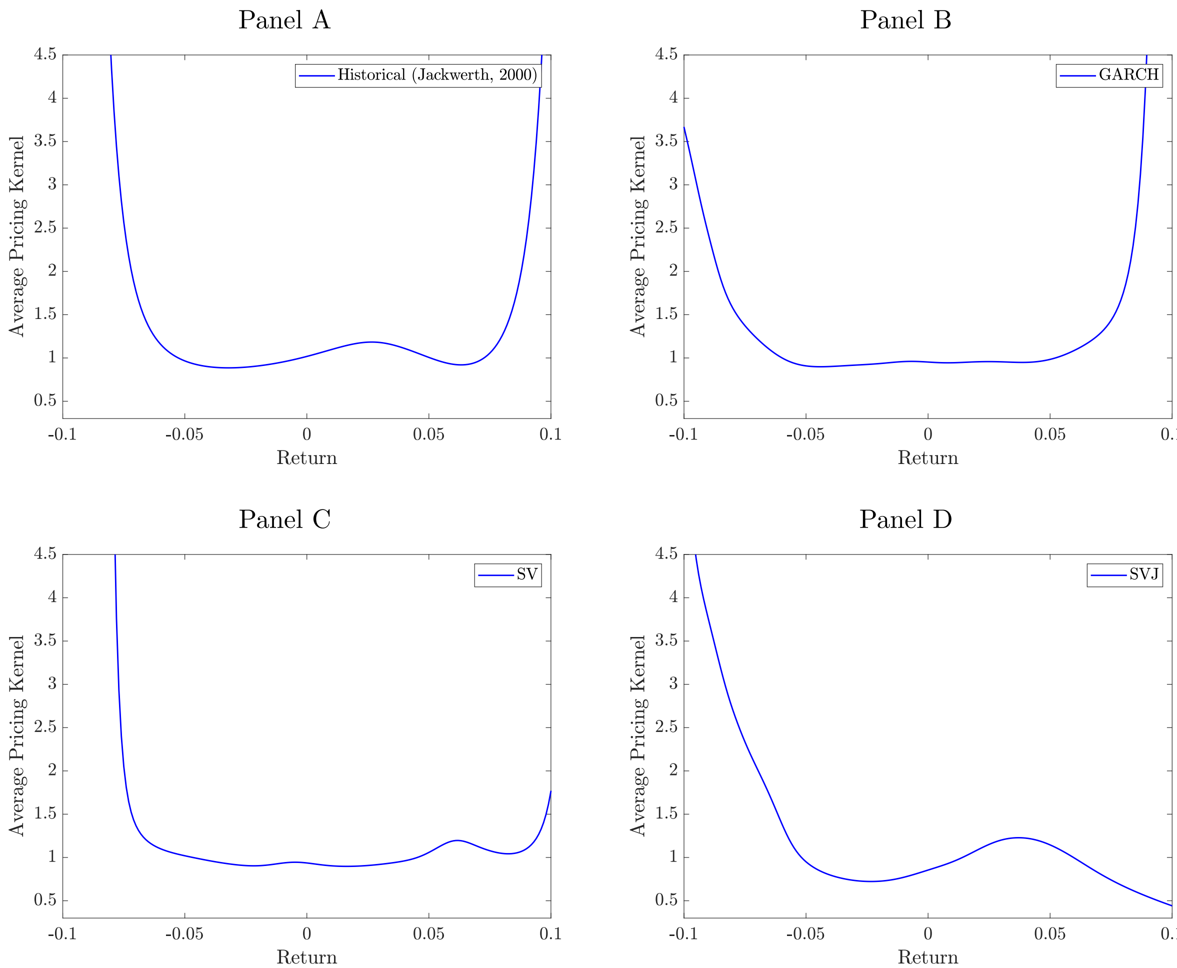
Shefrin, Hersh, 2001, On kernels and sentiment, Working paper.

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.

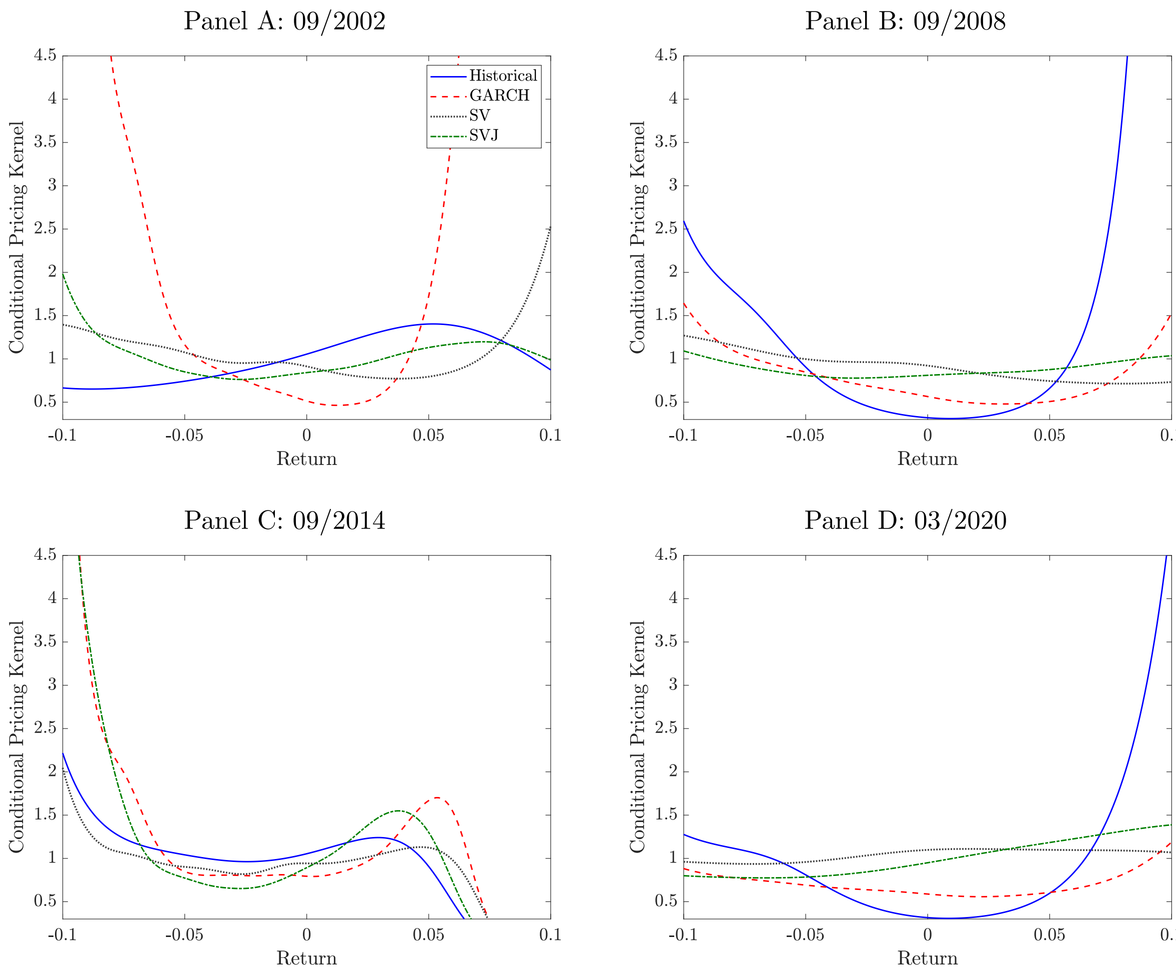
Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *Journal of Finance* 68, 987–1035.

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).

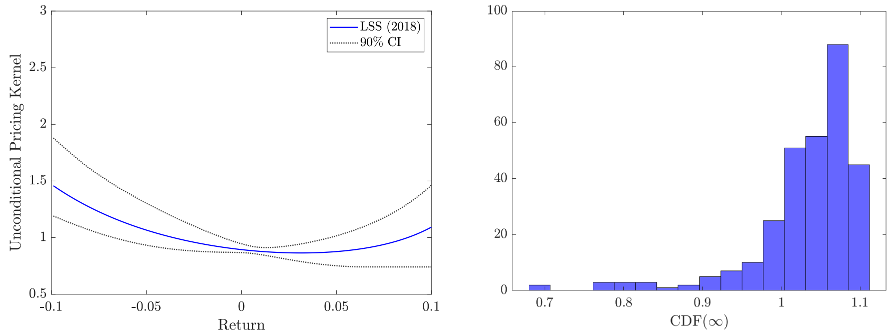
Figure 2: Conditional Empirical Pricing Kernels - Existing Approaches



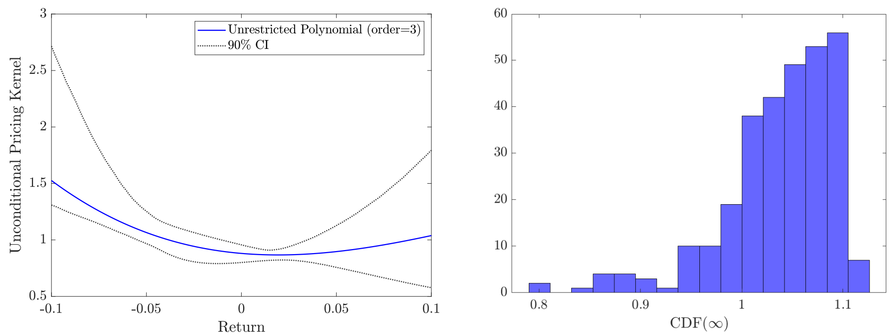
Notes: This figure shows the conditional empirical pricing kernels estimated using the historical data, GARCH model, stochastic volatility (SV) model, and stochastic volatility with jumps (SVJ) model on specific months: 09/2002 (Panel A), 09/2008 (Panel B), 09/2014 (Panel C), and 03/2020 (Panel D).

Figure 3: Unconditional Empirical Pricing Kernels

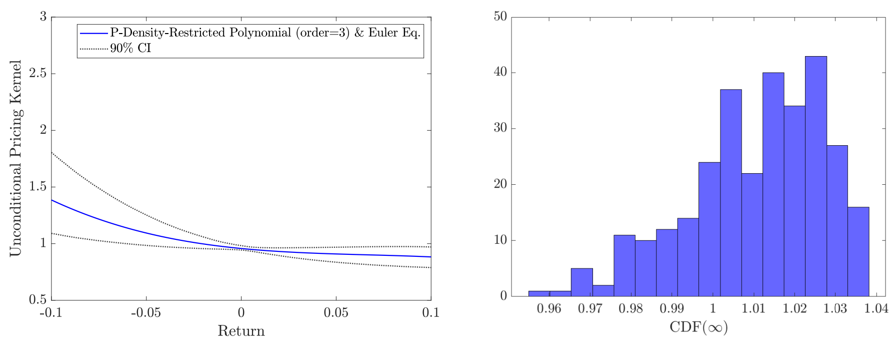
Panel A: Linn, Shive, and Shumway (2018)



Panel B: Unrestricted Polynomial EPK



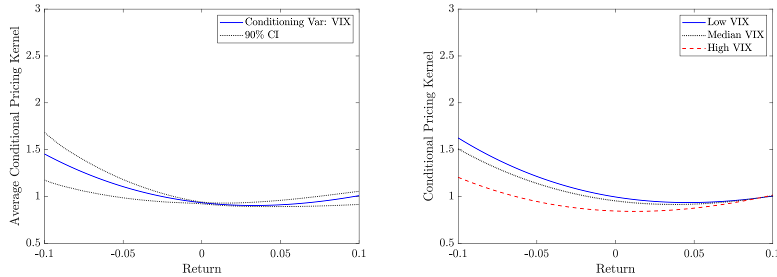
Panel C: Polynomial EPK with the P-Density Restriction and Euler Equations



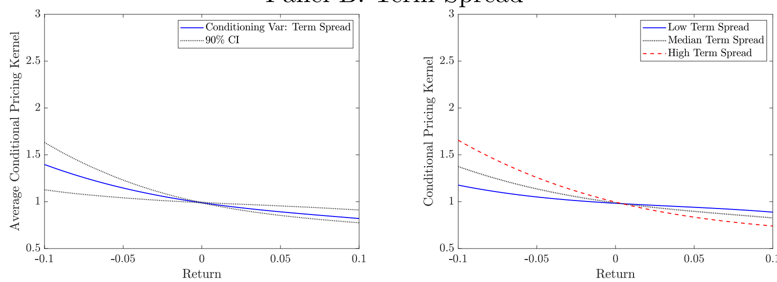
Notes: This figure shows the forward-looking unconditional empirical pricing kernels (EPKs). Panel A follows the CDI method of Linn, Shive, and Shumway (2018), Panel B describes the unrestricted polynomial pricing kernel estimation, and Panel C describes the polynomial pricing kernel estimation considering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

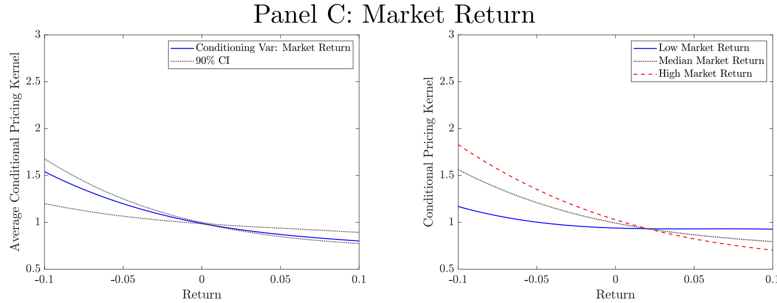
Panel A: VIX



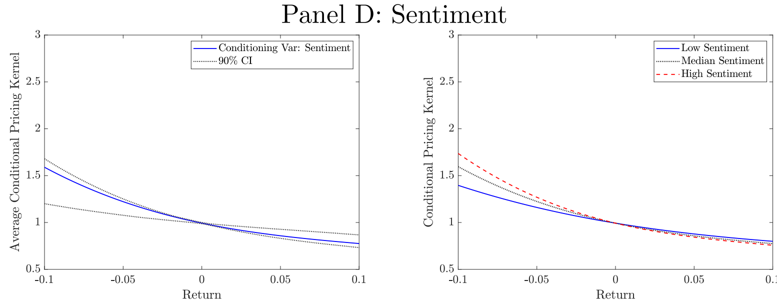
Panel B: Term Spread



Panel C: Market Return

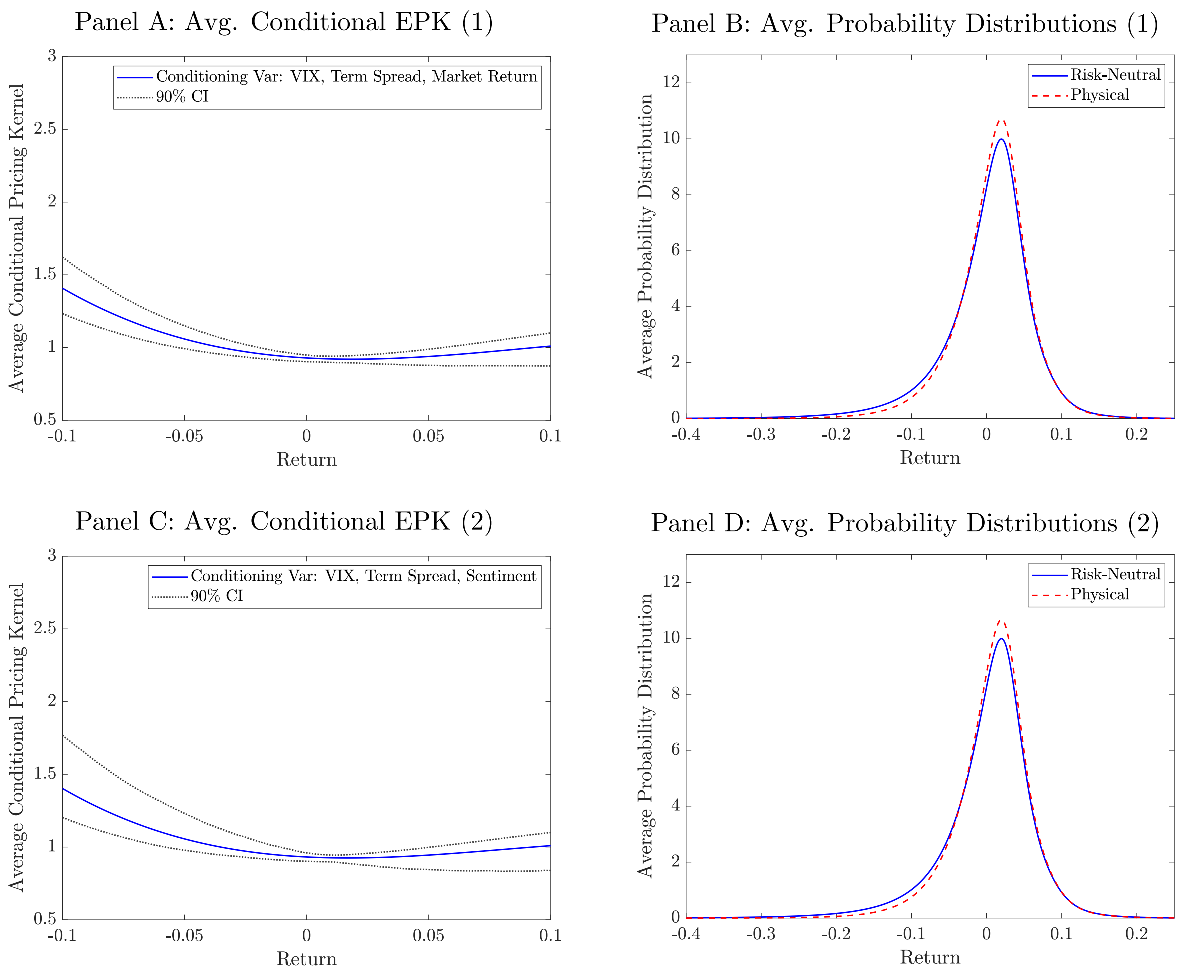


Panel D: Sentiment



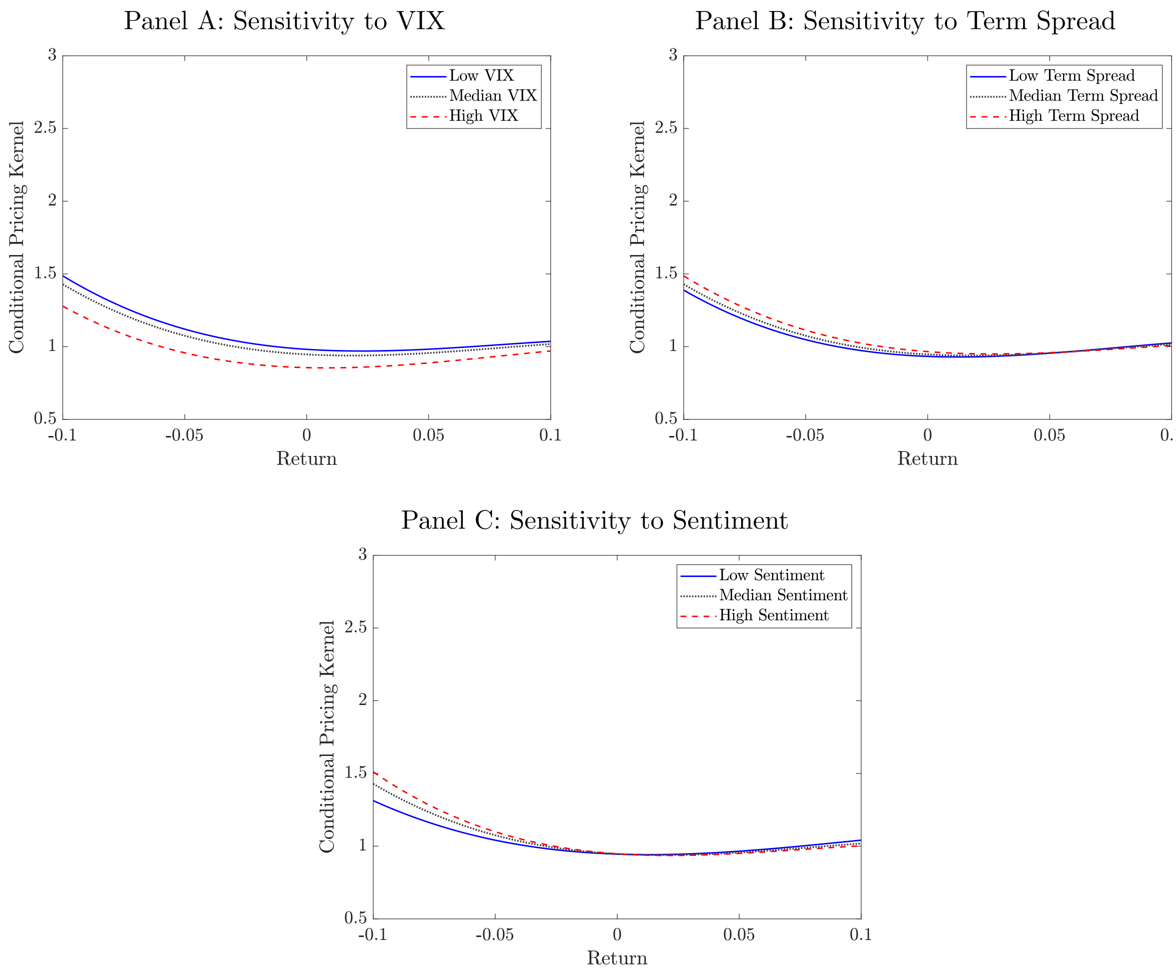
Notes: This figure shows the forward-looking conditional empirical pricing kernels with one of four conditioning variables: the VIX (Panel A), term spread (Panel B), market return (Panel C), and sentiment (Panel D). The left side of each panel is the average empirical conditional pricing kernel with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

Figure 5: Multivariate Conditional Empirical Pricing Kernels



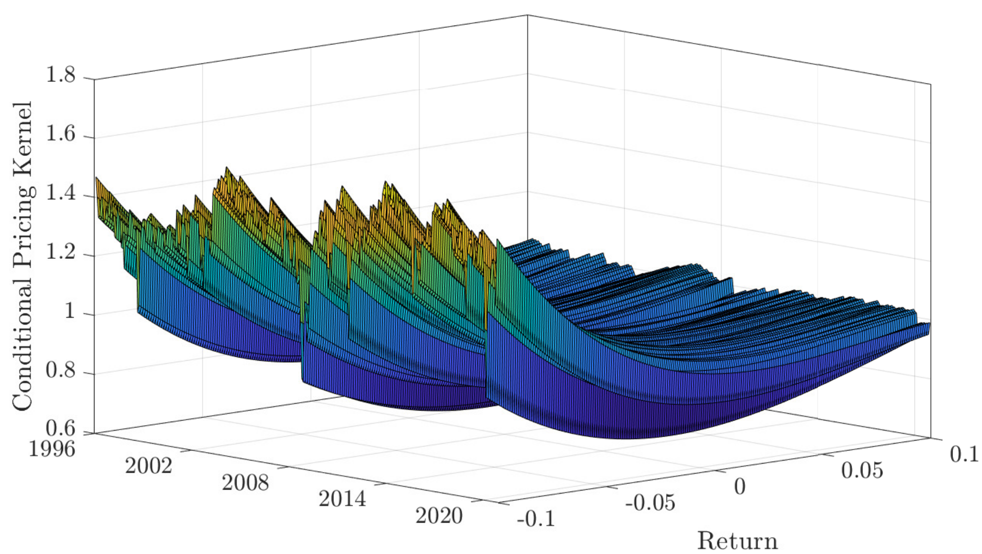
Notes: The left side of this figure shows the average forward-looking conditional empirical pricing kernels with two sets of three conditioning variables: (1) the VIX, term spread, and market return (Panel A) or (2) the VIX, term spread, and sentiment (Panel C). Panels B and D present their corresponding average risk-neutral (solid blue line) and physical (dashed red line) probability densities. In Panels A and C, the dotted black lines are the point-wise bootstrapped 90% confidence interval.

Figure 6: Sensitivity of the Multivariate Conditional Empirical Pricing Kernel



Notes: This figure shows how the multivariate conditional empirical pricing kernel changes when the VIX (Panel A), term spread (Panel B), or sentiment (Panel C) varies from the 5th percentile to the 95th percentile of its data while the other two variables are fixed at their median values.

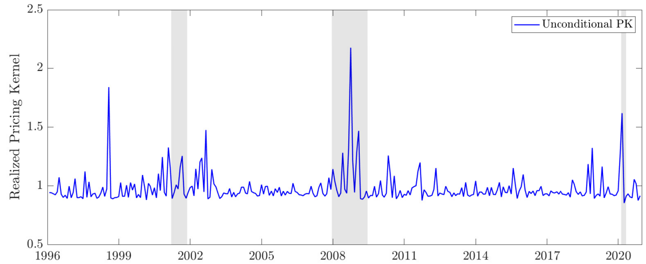
Figure 7: Time Series of the Multivariate Conditional Empirical Pricing Kernel



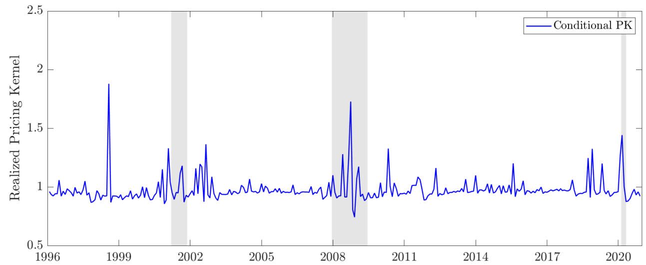
Notes: This figure describes the time series of the multivariate conditional empirical pricing kernel when conditioning variables are the VIX, term spread, and sentiment. The sample period is from January 1996 to December 2020.

Figure 8: Time Series of the Realized Monthly Pricing Kernel Estimate

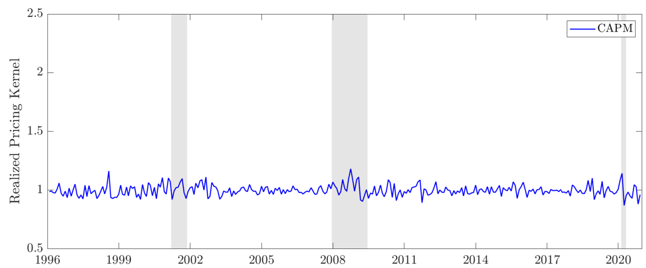
Panel A: Unconditional Pricing Kernel



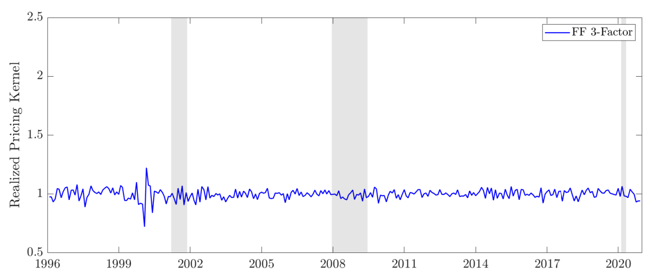
Panel B: Conditional Pricing Kernel



Panel C: CAPM-Implied Pricing Kernel

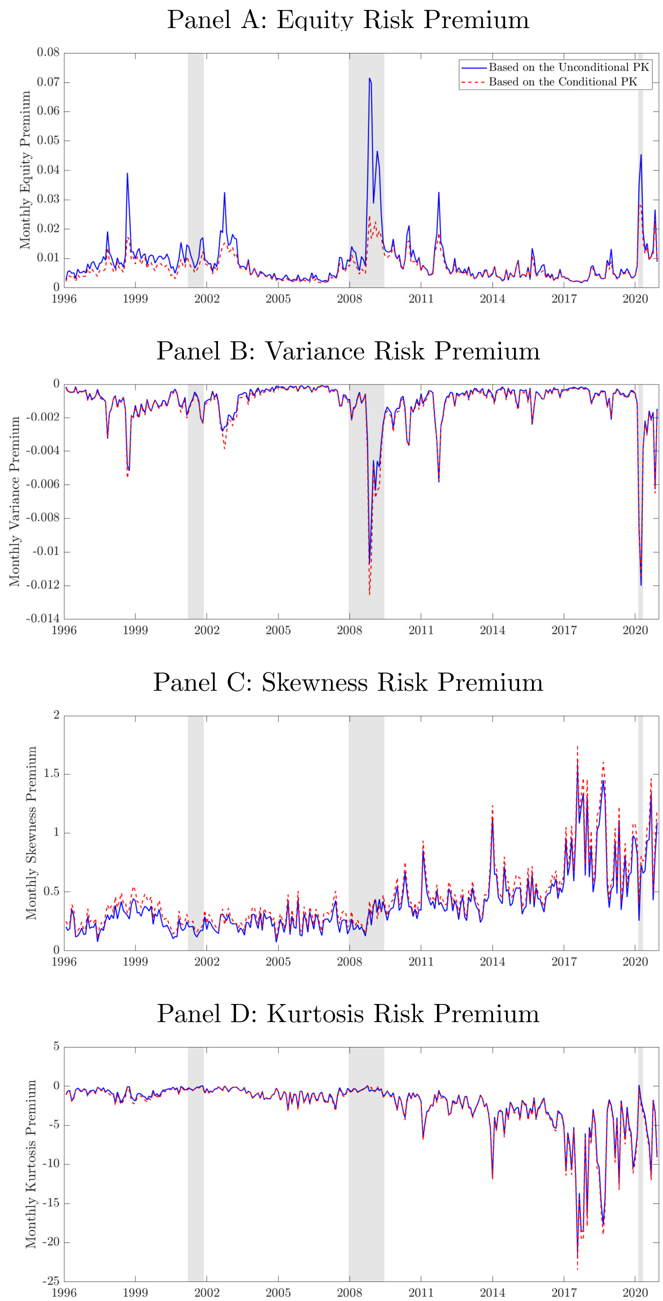


Panel D: FF3-Implied Pricing Kernel



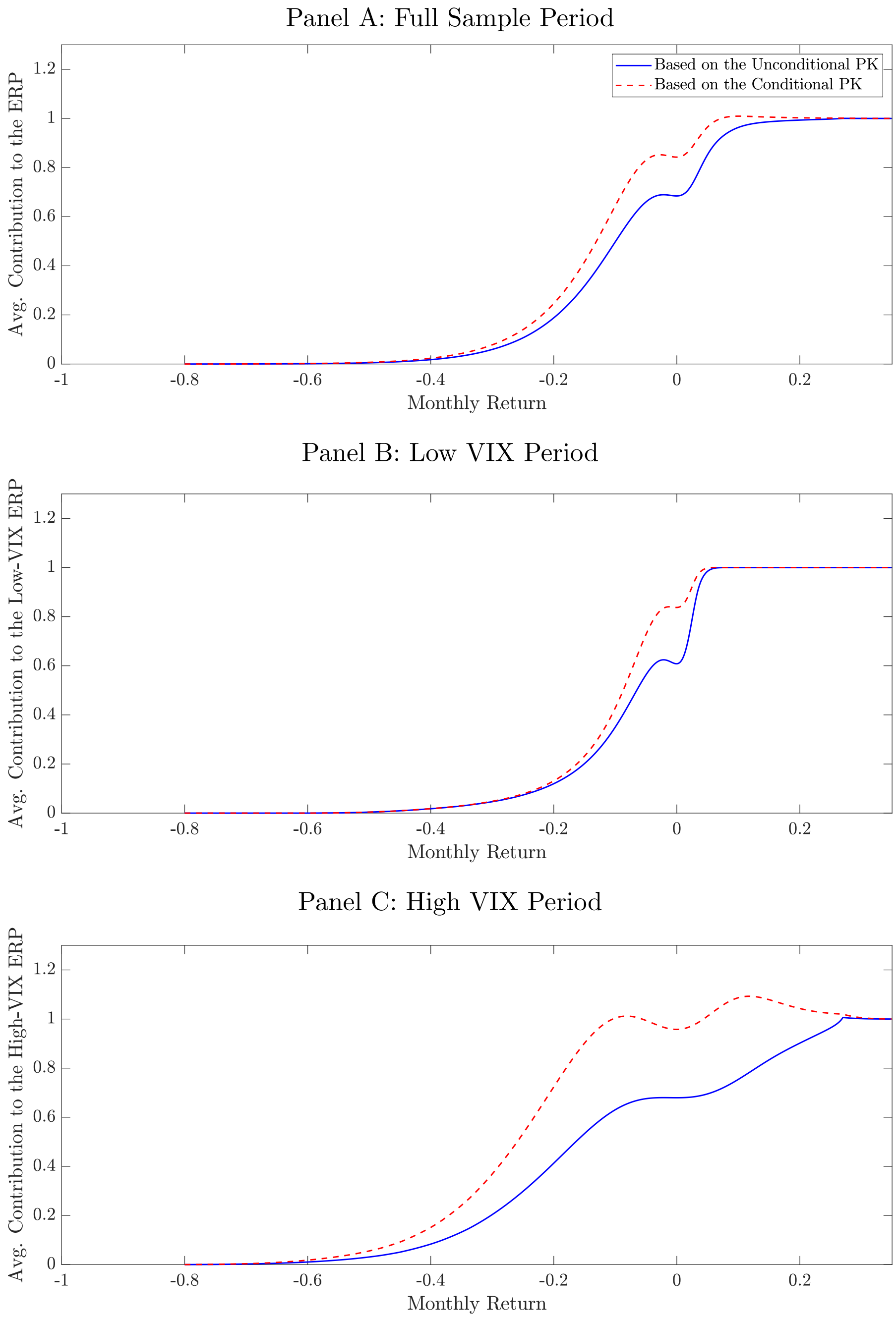
Notes: This figure describes the time series of the realized monthly pricing kernel inferred from the unconditional pricing kernel estimate (Panel A), the conditional pricing kernel estimate (Panel B), the CAPM (Panel C), and the Fama-French 3-factor model (Panel D). The sample period is from January 1996 to December 2020.

Figure 9: 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.

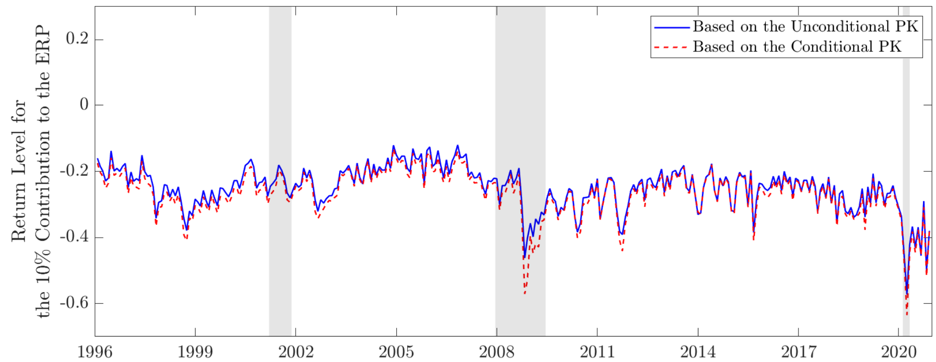
Figure 10: Sources of the Conditional Equity Risk Premium



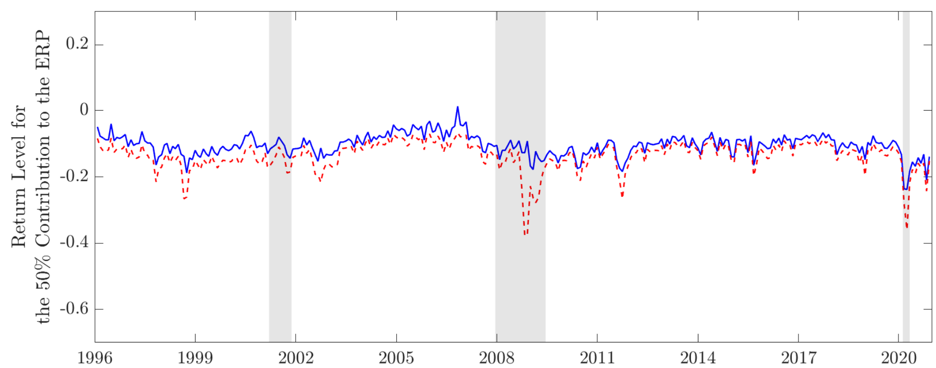
Notes: This figure shows the sources of the conditional equity risk premium during the full sample period (Panel A), low-VIX period (Panel B), and high-VIX period (Panel C). In each panel, the average contributions to the ERP based on the unconditional (solid blue line) and conditional (dashed red line) pricing kernel estimates are plotted.

Figure 11: Time Series of the Sources of the Conditional Equity Risk Premium

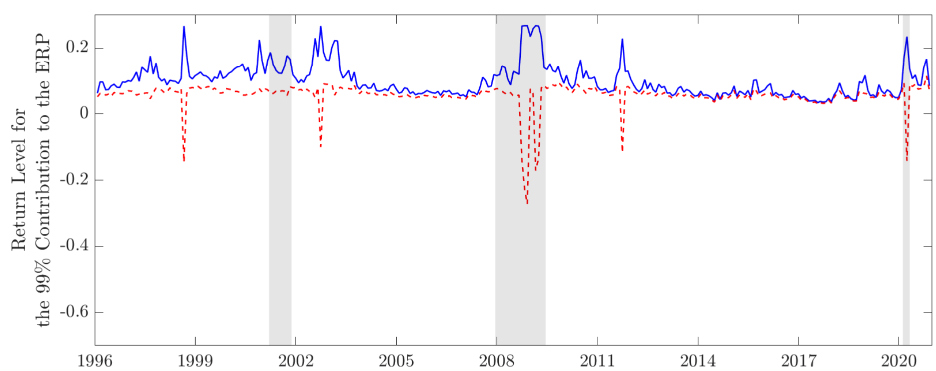
Panel A: 10% Contribution



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: Descriptive Statistics for the Conditioning Variables

	VIX	Skewness	Kurtosis	LeftTailRisk	Sentiment	TermSpread	CFNAI	ADS	IndusProd	ConsGrowth	MktReturn
Mean	0.204	-2.470	17.485	6.862	0.223	0.011	-0.093	-0.260	0.001	0.001	0.009
StdDev	0.080	1.037	14.377	2.814	0.139	0.009	1.259	2.089	0.011	0.010	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			
IndusProd	-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	
MktReturn	-0.394	-0.178	0.191	-0.262	0.436	-0.020	-0.014	0.191	-0.033	0.011	1.000

Notes: This table summarizes the time series statistics (mean, standard deviation, and correlations) of 11 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, real consumption growth, and S&P 500 market return.

Table 2: Parameter Estimates with One Conditioning Variable

	Parameter Estimates for the Time-Varying Coefficients									
	Intercept		$r_{t+\tau}$		$r^2_{t+\tau}$		$r^3_{t+\tau}$			
Cond. Var.	$\beta_{0,0}$	$\beta_{0,1}$	$\beta_{1,0}$	$\beta_{1,1}$	$\beta_{2,0}$	$\beta_{2,1}$	$\beta_{3,0}$	$\beta_{3,1}$	Hansen's J	p-value
VIX	1.07	-0.68	-3.74	9.29	34.88	-25.04	-47.85	48.56	3.64	0.056
	(56.98)	(-3.60)	(3.13)	(1.94)	(0.61)	(-0.19)	(-0.03)	(0.01)		
Skewness	0.98	-0.01	-1.60	1.06	7.06	-8.30	-39.14	18.64	12.25	0.000
	(38.89)	(-0.31)	(-0.97)	(1.62)	(0.09)	(-0.29)	(-0.01)	(0.02)		
Kurtosis	0.99	0.00	-3.82	0.10	13.63	0.14	-49.53	0.76	14.15	0.000
	(18.01)	(-0.10)	(-8.10)	(4.45)	(0.76)	(0.09)	(-0.08)	(0.02)		
LeftTailRisk	1.07	-0.01	-5.16	0.38	9.17	0.50	17.00	-7.74	7.73	0.005
	(7.83)	(-0.80)	(-8.74)	(5.12)	(0.11)	(0.07)	(0.03)	(-0.22)		
Sentiment	0.99	0.00	-2.89	-2.72	11.81	31.96	-26.81	-137.52	6.58	0.010
	(40.57)	(-0.00)	(5.64)	(-1.66)	(0.41)	(0.18)	(-0.07)	(-0.05)		
TermSpread	0.98	0.48	-1.10	-118.40	5.50	577.19	-45.64	-1.37	5.71	0.017
	(26.35)	(0.17)	(-0.90)	(-1.73)	(0.13)	(0.11)	(-0.03)	(0.00)		
CFNAI	1.00	0.01	-3.62	-0.12	15.15	-1.24	-25.16	7.28	11.44	0.001
	(19.05)	(0.08)	(-16.26)	(-0.45)	(0.42)	(-0.02)	(-0.06)	(0.02)		
ADS	1.00	0.01	-2.94	-0.12	11.55	-0.39	-27.60	3.47	23.10	0.000
	(24.18)	(0.35)	(-9.62)	(-0.60)	(0.35)	(-0.06)	(-0.06)	(0.02)		
IndusProd	1.00	-1.08	-2.58	-1.59	9.71	118.70	-2.89	-6.25	4.90	0.027
	(27.98)	(-0.12)	(-14.40)	(-0.15)	(0.65)	(0.07)	(-0.03)	(-0.03)		
ConsGrowth	0.99	1.86	-2.82	-2.57	11.79	-237.24	-8.95	93.63	6.43	0.011
	(19.38)	(0.14)	(-14.10)	(-0.07)	(0.45)	(-0.07)	(-0.05)	(0.00)		
MktReturn	0.98	0.57	-2.88	-29.63	17.44	83.97	-56.67	116.60	6.49	0.011
	(22.46)	(0.26)	(-6.11)	(-2.29)	(0.47)	(0.11)	(-0.13)	(0.01)		

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. Hansen's J statistic for the over-identification test, as well as its p-value, are also reported.

Table 3: Parameter Estimates with Multiple Conditioning Variables

	(1)				(2)			
Cond. Var.	Intercept	$r_{t+\tau}$	$r^2_{t+\tau}$	$r^3_{t+\tau}$	Intercept	$r_{t+\tau}$	$r^2_{t+\tau}$	$r^3_{t+\tau}$
Constant ($\beta_{i,0}$)	1.05	-1.50	25.71	-109.29	1.03	-0.91	24.53	-90.13
	(8.81)	(-2.24)	(0.41)	(-0.04)	(14.35)	(-2.30)	(0.18)	(-0.02)
VIX ($\beta_{i,1}$)	-0.68	6.75	5.04	69.86	-0.57	3.08	-5.30	14.70
	(-1.55)	(1.85)	(0.01)	(0.01)	(-0.81)	(1.69)	(-0.02)	(0.00)
TermSpread ($\beta_{i,2}$)	1.44	-71.59	47.49	-581.64	1.24	-23.90	27.08	205.10
	(0.09)	(-2.38)	(0.00)	(-0.00)	(0.17)	(-1.96)	(0.00)	(0.00)
MktReturn ($\beta_{i,3}$)	-0.19	-2.31	88.00	-405.39				
	(-0.05)	(-0.37)	(0.03)	(-0.01)				
Sentiment ($\beta_{i,3}$)					0.00	-1.24	16.52	-129.58
					(0.01)	(-1.55)	(0.06)	(-0.02)
Hansen's J	5.16				5.89			
p-value	0.023				0.015			
Wald Statistic	72.74				39.54			
p-value	0.000				0.000			

Notes: This table reports the parameter estimates of the conditional pricing kernel with three conditioning variables. Specification (1) includes the conditioning variables of the VIX, term spread, and market return. Specification (2) includes the conditioning variables of the VIX, term spread, and sentiment. The GMM t-statistics are displayed in parentheses. Hansen's J statistic for testing the over-identification and the Wald statistic for testing the restriction of $\beta_{i,1} = \beta_{i,2} = \beta_{i,3} = 0$, as well as their p-values, are also reported.

Table 4: Descriptive Statistics for the Monthly Risk-Neutral and Physical Distributions

	Mean (%)	StdDev (%)	Skewness	Kurtosis
Panel A. Risk-Neutral Distribution				
Time-Series Mean	0.20	5.72	-1.51	8.32
Time-Series StdDev	0.19	2.19	0.64	5.46
Quantile (5%)	0.00	3.26	-2.86	3.75
Quantile (50%)	0.15	5.28	-1.41	6.68
Quantile (95%)	0.53	9.94	-0.68	20.53
Panel B. Physical Distribution (from the Unconditional EPK)				
Time-Series Mean	1.10	4.87	-1.11	5.78
Time-Series StdDev	0.89	1.64	0.41	2.21
Quantile (5%)	0.35	2.91	-1.86	3.53
Quantile (50%)	0.90	4.56	-1.09	5.26
Quantile (95%)	2.50	7.70	-0.48	10.31
Panel C. Physical Distribution (from the Conditional EPK)				
Time-Series Mean	0.88	4.76	-1.05	5.57
Time-Series StdDev	0.47	1.60	0.39	2.02
Quantile (5%)	0.35	2.80	-1.75	3.52
Quantile (50%)	0.80	4.49	-1.03	5.07
Quantile (95%)	1.78	7.55	-0.45	9.59

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 5: Descriptive Statistics for Monthly Risk Premia

	ERP (%)	VRP (%^{2})	SRP	KRP
Panel A. Unconditional Empirical Pricing Kernel				
Time-Series Mean	0.90	-11.09	0.40	-2.54
Time-Series StdDev	0.87	14.57	0.26	3.36
Quantile (5%)	0.26	-34.51	0.14	-10.44
Quantile (50%)	0.67	-6.57	0.33	-1.40
Quantile (95%)	2.33	-1.52	0.98	-0.17
AR1	0.76	0.70	0.74	0.74
Panel B. Conditional Empirical Pricing Kernel				
Time-Series Mean	0.68	-12.27	0.46	-2.75
Time-Series StdDev	0.44	15.77	0.28	3.57
Quantile (5%)	0.22	-38.65	0.18	-11.16
Quantile (50%)	0.56	-7.59	0.39	-1.59
Quantile (95%)	1.58	-1.98	1.08	-0.19
AR1	0.77	0.72	0.75	0.74

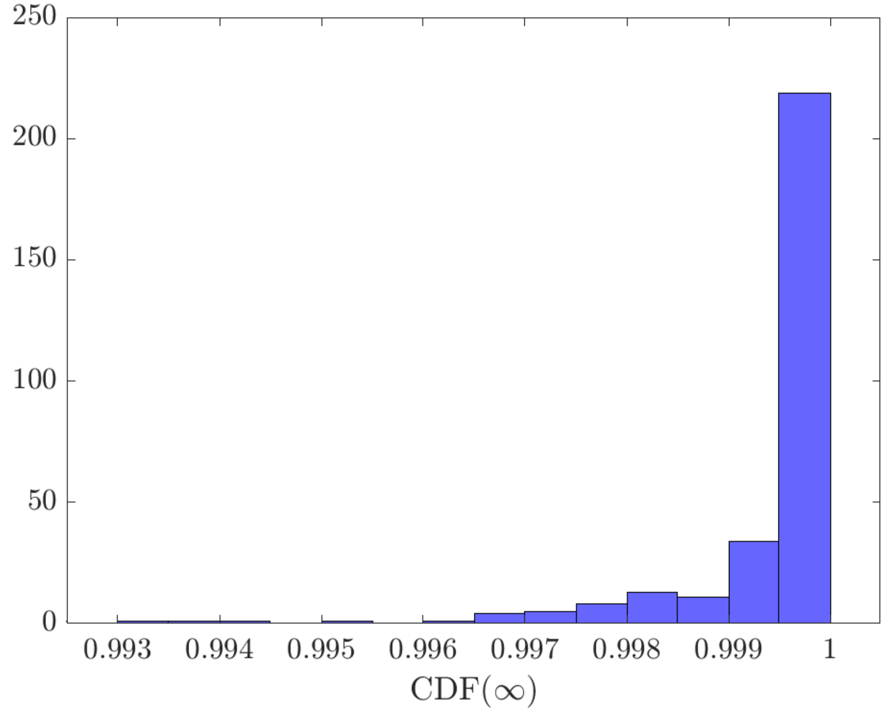
Notes: This table reports the descriptive statistics for the conditional risk premia. Panel A is based on the unconditional pricing kernel estimate, and Panel B is based on the conditional pricing kernel estimate. The conditional equity, variance, skewness, and kurtosis risk premia are first calculated,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

Table 6: Out-Of-Sample R^2 of the Market Return Forecasts

			Regression-Based Predictor Variables											
Benchmark	ERP_{unc}	ERP_{cond}	VRP_{BTZ}	DP	DY	EP	BM	NTIS	TBL	LTR	TMS	DFY	DFR	INFL
Hist. Avg.	0.69	3.12	2.95	3.34	3.42	0.26	1.43	-4.06	0.85	-0.61	-0.27	-1.14	-1.43	-3.47
	0.28	0.08	0.06	0.01	0.00	0.26	0.01	0.92	0.06	0.69	0.86	0.93	0.53	0.94
ERP_{unc}		2.45												
		0.10												
ERP_{cond}			-0.17	0.23	0.31	-2.95	-1.75	-7.41	-2.34	-3.85	-3.50	-4.40	-4.70	-6.80
			0.25	0.26	0.27	0.50	0.43	0.64	0.47	0.55	0.53	0.57	0.50	0.65

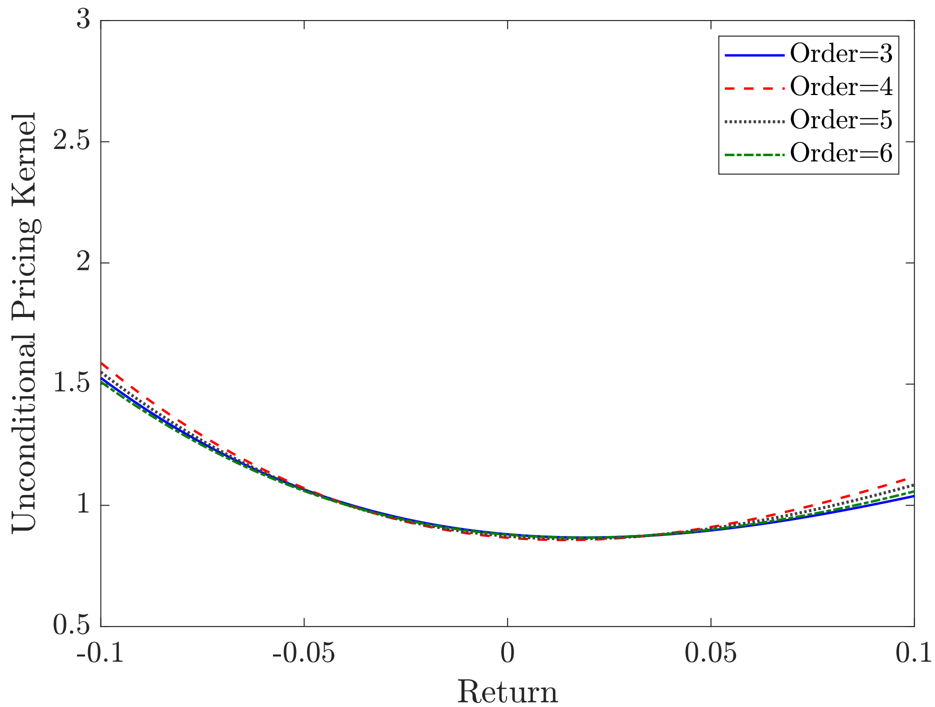
Notes: This table reports the out-of-sample performance of the monthly market return forecasts. For each benchmark, the first row displays the out-of-sample R^2 (%) calculated by equation (11), and the second row displays their associated p-values following Clark and West (2007). All monthly forecasts are computed with expanding windows. The in-sample period begins from January 1996, and the out-of-sample period is from January 2011 to December 2020.

Figure A.1: Validity of 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.

Figure A.2: Unrestricted Unconditional EPK with Various Polynomial Orders



Notes: This figure shows the unrestricted unconditional empirical pricing kernels with various degrees of the polynomial function for the empirical pricing kernel approximation. The pricing kernel estimates with a highest order of three (solid blue line), four (dashed red line), five (dotted black line), and six (dashed-dotted green line) are plotted.