Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

A Tale of Two Option Markets: Pricing Kernels and Volatility Risk

Zhaogang Song and Dacheng Xiu

2014-58

A Tale of Two Option Markets: Pricing Kernels and Volatility Risk *

Zhaogang Song[†] Dacheng Xiu[‡]
Federal Reserve Board University of Chicago

This Version: January, 2014

Abstract

Using prices of both S&P 500 options and recently introduced VIX options, we study asset pricing implications of volatility risk. While pointing out the joint pricing kernel is not identified nonparametrically, we propose model-free estimates of marginal pricing kernels of the market return and volatility conditional on the VIX. We find that the pricing kernel of market return exhibits a decreasing pattern given either a high or low VIX level, whereas the unconditional estimates present a U-shape. Hence, stochastic volatility is the key state variable responsible for the U-shape puzzle documented in the literature. Finally, our estimates of the volatility pricing kernel feature a U-shape, implying that investors have high marginal utility in both high and low volatility states.

Key Words: Pricing Kernel, State-Price Density, VIX Option, Volatility Risk

JEL classification: G12,G13

*We benefited from discussions with Yacine Aït-Sahalia, Andrea Buraschi, Bjorn Eraker, Peter Carr, Peter Christoffersen, Fousseni Chabi-Yo, George Constantinides, Jianqing Fan, René Garcia, Kris Jacobs, Jakub Jurek, Ilze Kalnina, Ralph Koijen, Nicholas Polson, Eric Renault, Jeffrey Russell, Neil Shephard, George Tauchen (discussant), Viktor Todorov, Grigory Vilkov (discussant), Hao Zhou, as well as seminar and conference participants at the University of Chicago, Northwestern, Princeton, Toulouse School of Economics, Liverpool School of Management, the 2012 CICF, the 5th Annual SoFiE Conference, the Measuring Risk conference 2012, the 2012 Financial Engineering and Risk Management International Symposium, and the 2012 International Symposium on Risk Management and Derivatives. Xiu acknowledges research support by the Fama-Miller Center for Research in Finance at Chicago Booth. The views expressed herein do not reflect those of the Board of Governors of the Federal Reserve System.

[†]Board of Governors of the Federal Reserve System, Mail Stop 165, 20th Street and Constitution Avenue, Washington, DC, 20551. E-mail: Zhaogang.Song@frb.gov.

[‡]University of Chicago Booth School of Business, 5807 S. Woodlawn Avenue, Chicago, IL 60637. Email: dacheng.xiu@chicagobooth.edu.

1 Introduction

In addition to the uncertainty of market returns, volatility risk has been well documented as an essential component of time-varying investment opportunities. Together with the preferences of economic agents, a priced volatility factor leads to a pricing kernel (or stochastic discount factor) which depends on both market returns and volatility. Nevertheless, because volatility is neither tradable nor observable, existing studies on pricing kernels either impose strong parametric restrictions, or ignore the unobservable volatility factor in nonparametric analysis. The pricing kernel estimates produced by these studies exhibit a puzzling U-shape as a function of market return, in conflict with a standard expected utility theory.

The lack of tradable and observable volatility has changed substantially since the introduction of the Volatility Index (VIX) in 1993 by the Chicago Board of Options Exchange (CBOE),[1] and the introduction of VIX derivatives such as futures and options in 2004 and 2006, respectively. The VIX, derived from S&P 500 options as the square root of the expected average variance over the next 30 calendar days, provides investors with a direct measure of volatility; and VIX derivatives offer investors convenient instruments for trading on the volatility of S&P 500 index.[2] As a result, the VIX is constantly exposed in the media spotlight, and VIX options have achieved huge liquidity and become the third most active contracts at CBOE as of October 2011.

Taking advantage of the S&P 500 and VIX option markets, we nonparametrically identify and estimate the marginal pricing kernel of market returns and volatility, which equals the ratio of state-price density (or risk-neutral density) to physical density. We show that information in the two option prices is fully captured by the two marginal state-price densities of market returns and volatility separately, whereas the joint state-price density and hence joint

[1] The VIX, from its inception, was calculated from S&P 500 index options by inverting the Black-Scholes formula. In 2003, the CBOE amended this approach and adopted a model-free method to calculate the VIX.

[2] Previously, investors have to take positions in option portfolios, such as straddles or strangles, in order to trade volatility.

pricing kernel cannot be identified nonparametrically as a result of incomplete markets.[3] We then provide nonparametric estimates of pricing kernels with respect to return and volatility respectively. Our estimates not only shed light on the puzzling U-shaped pricing kernel, but also provide new empirical stylized facts on the pricing kernel of volatility. In particular, we make several important findings regarding asset pricing implications of volatility.

First, our estimates of pricing kernels with respect to the market return show that stochastic volatility is the key state variable responsible for the "pricing kernel puzzle." More specifically, we find that a pricing kernel of market return conditional on either a high or low VIX level presents a decreasing pattern, whereas an unconditional pricing kernel (i.e. the one that ignores volatility) may become U-shaped. In fact, marginal utility (the pricing kernel up to a scaling factor) conditional on high volatility is above that conditional on low volatility, as low volatility signals a good investment opportunity and hence is preferred by investors. As a mixture of pricing kernel estimates conditional on different volatility levels, unconditional estimates can exhibit an increasing pattern over the high return region (right tail), where high volatility is prevalent. Our finding echoes the conclusions of parametric models in Chabi-Yo et al. (2008) and Christoffersen et al. (2010), which show that missing state variables in the pricing kernel may result in a U-shape. Without restricting the specification of pricing kernels, however, we show that including volatility as a state variable is the solution to this puzzle.

Second, we provide nonparametric estimates of pricing kernels with respect to volatility, for the first time to the best of our knowledge. Our estimates exhibit a pronounced U-shape conditional on either a high or low VIX, indicating that investors attach high marginal utility to payoffs received in both high and low future volatility states, regardless of today's

[3] We emphasize that the joint pricing kernel, though not identifiable nonparametrically using the S&P 500 and VIX options, can be estimated with certain parametric correlation restriction on the two marginal pricing kernels. We do not explore this approach because our focus is to recover the pricing kernels without any parametric restrictions. A follow-up paper of our study, Jackwerth and Vilkov (2013), implemented such an exercise using the parametric Frank copula for the two marginal distributions.

volatility level. Bakshi et al. (2010) also document a U-shape for the unconditional volatility pricing kernel, but indirectly, by exploring the link between the monotonicity of pricing kernel and returns of VIX option portfolios. In contrast, we provide direct estimates of the conditional volatility pricing kernel by nonparametric methods, which provide further information about the shape and tail behavior of the pricing kernel. In particular, we find that the volatility pricing kernel is asymmetric, and the asymmetry conditional on a current high volatility is much stronger than that conditional on a low volatility. This finding implies that market investors price the volatility risk differently according to different scenarios of the economy, which presents new empirical regularities that need to be incorporated into models of volatility risk.

Finally, we evaluate the performance of our nonparametric estimator for in-sample fitting and out-of-sample forecasts against two alternative methods: the nonparametric approach of Aït-Sahalia and Lo (1998) without a volatility factor and a martingale approach commonly used by practitioners that simply predicts tomorrow's implied volatility by interpolating today's implied volatility surface. We find that our estimator outperforms both alternative methods for density and implied volatility forecasts, which again highlights the importance of conditioning on volatility.

Estimating pricing kernels from option prices is discussed in Aït-Sahalia and Lo (1998), Aït-Sahalia and Duarte (2003), Jackwerth (2000), and Rosenberg and Engle (2002), which ignore the volatility risk and discover a puzzling U-shape.[4] Hereafter, many studies have proposed different explanations for the U-shaped pricing kernel, including models with missing state variables in Chabi-Yo et al. (2008), Chabi-Yo (2011), and Christoffersen et al. (2010), and models with heterogeneous agents in Bakshi and Madan (2008) and Ziegler (2007). Our empirical study contributes to this literature by showing, without imposing any parametric restrictions, that volatility is the missing state variable responsible for the puzzle.

[4] A related study, Fan and Mancini (2009), proposes nonparametric methods for pricing derivatives based on state price distributions.

Our paper is also related to the large literature on models with volatility risk, including both reduced-form option pricing models, e.g. Bakshi et al. (1997), Bates (2000), Pan (2002), Eraker (2004), and Broadie et al. (2007), and equilibrium models, such as Bansal et al. (2012), Bollerslev et al. (2012), and Campbell et al. (2012). Unlike these studies, our framework does not depend on any parametric restrictions on volatility dynamics that may obscure the empirical characteristics of pricing kernels. Several recent studies have constructed model-free measures of risk-neutral volatility from S&P 500 options, e.g. Bakshi and Kapadia (2003), Bollerslev et al. (2009), Carr and Wu (2009), and Todorov (2010), and compared them with measures of realized volatility. Their focuses are on the sign, time variation, and return predictability of variance risk premium, which only relates to the conditional mean of variance distributions under different measures. In contrast, we recover the entire volatility pricing kernel.

Methodologically, our paper is also related to Boes et al. (2007) and Li and Zhao (2009) who estimate pricing kernels of stock market returns and interest rates, respectively, conditional on an ex-post volatility proxy filtered from historical time series. Our strategy differs from their approach by using the VIX, which possesses a monotonic functional relationship with the unobservable volatility for almost all state-of-the-art volatility models. Therefore, our method avoids estimation errors from the filtering stage, while making it possible to study volatility pricing kernels with the help of VIX options.

Furthermore, several recent studies document the importance of multiple volatility factors in capturing dynamics of option prices or the term structure of variance swaps, see e.g. Christoffersen et al. (2008), Egloff et al. (2010), Mencia and Sentana (2012), and Bates (2012). Our nonparametric framework can be extended to nest these models by including additional regressors such as a VIX future contract, or CBOE S&P 500 3-Month Volatility Index (VXV). Such an extension, though being interesting and important itself, is beyond the scope of the current focus, to which term structure of volatility is less relevant.

Section 2 discusses the nonparametric identification of pricing kernels of both the market return and volatility. Section 3 provides our nonparametric estimation framework and Monte Carlo simulations. Empirical estimates of pricing kernels are presented in Section 4. Section 5 concludes the paper.

2 Pricing Kernels with a Volatility Factor

The pricing kernel equals the ratio of risk-neutral density, also known as state-price density (SPD), to the density under the physical measure. To study pricing kernels, we first discuss the identification of state price densities, by exploring the underlying connection of S&P 500 options, VIX, and VIX options through the latent volatility factor.

2.1 Identification of State-Price Densities

To fix ideas, we denote the log price of the S&P 500 index as S_t, the VIX as Z_t, and the unobserved volatility as V_t. The information in the derivative markets is driven by the joint evolution of S_t and V_t, which determines Z_t endogenously. As V_t is not observable, there exist no Arrow-Debreu securities traded on V_t directly.

In fact, the payoffs of S&P 500 and VIX options depend on their own underlying indices at maturity T. Therefore, we focus on the marginal state-price densities with respect to S and Z separately. We show that the marginal densities together span the two option markets, and provide sufficient and necessary information about the dynamics of the market return and its volatility. The joint dynamics, nevertheless, cannot be identified nonparametrically unless certain options whose payoff depends on both S_T and Z_T are traded.

We write the time-t price of a S&P 500 call option with maturity T and strike x as: [5]

$$\begin{aligned} C(\tau, f_{t,\tau}, v_t, x, r_{t,\tau}) =& e^{-r_{t,\tau}\tau} E^{\mathbb{Q}} \left[(e^{S_T} - x)^+ | F_{t,\tau} = f_{t,\tau}, V_t = v_t \right] \\ =& e^{-r_{t,\tau}\tau} \int_{\mathbb{R}} (e^{s_T} - x)^+ p^*(s_T | \tau, f_{t,\tau}, v_t) ds_T \end{aligned}$$

where $F_{t,\tau}$ denotes the log forward price of the S&P 500 index, $\tau = T - t$ is the time-to-maturity, and $r_{t,\tau}$ is the deterministic risk-free rate between t and T at time t. Similarly, the price of a VIX call option with strike y is given by:

$$\begin{aligned} H(\tau, f_{t,\tau}, v_t, y, r_{t,\tau}) =& e^{-r_{t,\tau}\tau} E^{\mathbb{Q}} \left[(e^{z_T} - y)^+ | F_{t,\tau} = f_{t,\tau}, V_t = v_t \right] \\ =& e^{-r_{t,\tau}\tau} \int_{\mathbb{R}} (e^{z_T} - y)^+ q^*(z_T | \tau, f_{t,\tau}, v_t) dz_T \end{aligned}$$

Observe that the two SPDs $p^*(s_T|\tau, f_{t,\tau}, v_t)$ and $q^*(z_T|\tau, f_{t,\tau}, v_t)$ completely determine these option prices. Building upon the insight of Breeden and Litzenberger (1978), they can be estimated as the second order derivative of option prices with respect to different strikes. In particular, we can recover

$$p^*(s_T|\tau, f_{t,\tau}, v_t) = e^{r_{t,\tau}\tau + s_T} \frac{\partial^2 C(\tau, f_{t,\tau}, v_t, x, r_{t,\tau})}{\partial x^2} \Big|_{x = e^{s_T}}, \tag{1}$$

from S&P 500 options and

$$q^*(z_T|\tau, f_{t,\tau}, v_t) = e^{r_{t,\tau}\tau} \frac{\partial^2 H(\tau, f_{t,\tau}, v_t, y, r_{t,\tau})}{\partial y^2} \Big|_{y = z_T}, \tag{2}$$

from VIX options. It is apparent that $p^*(s_T|\tau, f_{t,\tau}, v_t)$ and $q^*(z_T|\tau, f_{t,\tau}, v_t)$ summarize the entire information about these two option markets, hence the joint density of s_T and z_T cannot be identified from the data without additional parametric assumptions.

Nevertheless, these two densities $p^*(s_T|\tau, f_{t,\tau}, v_t)$ and $q^*(z_T|\tau, f_{t,\tau}, v_t)$ are not practically feasible to estimate as V_t is unobservable. Alternatively, with the observed VIX from the

[5] In our setting, the time-t information set $\mathcal{F}_t$ contains stock prices, instantaneous volatility, interest rates and dividends, which can be summarized by the log forward price $F_{t,\tau}$ and the volatility V_t.

market,[6] we may rewrite the option prices with z_t as a state variable, i.e., $C(\tau, f_{t,\tau}, z_t, x, r_{t,\tau})$ and $H(\tau, f_{t,\tau}, z_t, y, r_{t,\tau})$,[7] and take second order derivatives to obtain

$$\begin{aligned} p^*(s_T|\tau, f_{t,\tau}, z_t) =& e^{r_{t,\tau}\tau+s_T}\frac{\partial^2 C(\tau, f_{t,\tau}, z_t, x, r_{t,\tau})}{\partial x^2}\Big|_{x=e^{s_T}} \\ q^*(z_T|\tau, f_{t,\tau}, z_t) =& e^{r_{t,\tau}\tau}\frac{\partial^2 H(\tau, f_{t,\tau}, z_t, y, r_{t,\tau})}{\partial y^2}\Big|_{y=z_T}. \end{aligned} \tag{3}$$

In fact, writing options in terms of $f_{t,\tau}$ and z_t amounts to assuming that V_t can be determined from Z_t and $F_{t,\tau}$, which is rigorous under most models of volatility risk in the literature (see Section 2.3 below for details). With state variables being fully observable, $p^*(s_T|\tau, f_{t,\tau}, z_t)$ and $q^*(z_T|\tau, f_{t,\tau}, z_t)$ can be identified from the data.

In summary, state-price densities $p^*(s_T|\tau, f_{t,\tau}, z_t)$ and $q^*(z_T|\tau, f_{t,\tau}, z_t)$ encapsulate all the information in the two option markets. They complement each other to reveal an intact picture of the market return, its volatility dynamics and the interactions of the two markets.

2.2 From State-Price Densities to Pricing Kernels

We now discuss how to obtain the pricing kernels by combining the risk-neutral and physical densities of S_t and Z_t. We denote $\pi(s_T, z_T|\tau, f_{t,\tau}, z_t)$ as the pricing kernel and use π for short. Not surprisingly, for the same reason described in Section 2.1, the joint pricing kernel $\pi(s_T, z_T|\tau, f_{t,\tau}, z_t)$ cannot be identified nonparametrically. We therefore study the projections of pricing kernel π on S_T[8] and Z_T, denoted as $\pi(s_T|\tau, f_{t,\tau}, z_t)$ and $\pi(z_T|\tau, f_{t,\tau}, z_t)$, respectively. They are called the pricing kernel of the market return and the pricing kernel

[6] The CBOE constructs Z_t form a portfolio of options weighted by strikes according to the formula:

$$(Z_t/100)^2 = E^{\mathbb{Q}}(QV_{t,\tau}|\mathcal{F}_t) = \frac{2e^{r_{t,\tau}\tau}}{\tau}\Big(\int_0^{e^{f_{t,\tau}}} \frac{P(\tau,x)}{x^2}dx + \int_{e^{f_{t,\tau}}}^{\infty} \frac{C(\tau,x)}{x^2}dx\Big) + \epsilon$$

where $QV_{t,T}$ denotes the quadratic variation of the log return process from t to $t+\tau$, $P(\tau,x)$ and $C(\tau,x)$ are put and call options with time-to-maturity τ and strike x, and $f_{t,\tau}$ is the log price of forward contracts, see e.g. Britten-Jones and Neuberger (2000) and Carr and Wu (2009).

[7] Strictly speaking, the function $C(\cdot)$ here is a composite function, which is different from the previous call option pricing function. We recycle it to simplify our notations.

[8] The projection of π on S_T is defined as $E^{\mathbb{P}}(\pi|S_T = s_T, F_{t,\tau} = f_{t,\tau}, Z_t = z_t)$.

of the VIX in the following.

In fact, the price of a S&P 500 call option can be written as

$$\begin{aligned} C(\tau, f_{t,\tau}, z_t, x, r_{t,\tau}) =& e^{-r_{t,\tau}\tau} E^{\mathbb{P}} \left[\pi \cdot (e^{S_T} - x)^+ | F_{t,\tau} = f_{t,\tau}, Z_t = z_t \right] \\ =& e^{-r_{t,\tau}\tau} \int_{\mathbb{R}} \pi \left(s_T | \tau, f_{t,\tau}, z_t \right) \left(e^{s_T} - x \right)^+ p(s_T | \tau, f_{t,\tau}, z_t) ds_T, \end{aligned} \tag{4}$$

and the price of a VIX call option is

$$\begin{aligned} H(\tau, f_{t,\tau}, z_t, y, r_{t,\tau}) =& e^{-r_{t,\tau}\tau} E^{\mathbb{P}} \left[\pi \cdot (e^{z_T} - y)^+ | F_{t,\tau} = f_{t,\tau}, Z_t = z_t \right] \\ =& e^{-r_{t,\tau}\tau} \int_{\mathbb{R}} \pi \left(z_T | \tau, f_{t,\tau}, z_t \right) \left(e^{z_T} - y \right)^+ q(z_T | \tau, f_{t,\tau}, z_t) dz_T, \end{aligned} \tag{5}$$

where $p(s_T|\tau, f_{t,\tau}, z_t)$ and $q(z_T|\tau, f_{t,\tau}, z_t)$ are conditional densities of S_T and Z_T under the physical measure, respectively. Note that the law of iterated expectation is used in the second equality of both (4) and (5).

Similar to (3), equations (4) and (5) imply that the second order derivatives of the S&P 500 and VIX call prices with respect to their strikes are also equal to $\pi\left(s_T|\tau, f_{t,\tau}, z_t\right) p(s_T|\tau, f_{t,\tau}, z_t)$ and $\pi\left(z_T|\tau, f_{t,\tau}, z_t\right) q(z_T|\tau, f_{t,\tau}, z_t)$, respectively. This fact, combined with (3), further implies that

$$\begin{aligned} \pi\left(s_T|\tau, f_{t,\tau}, z_t\right) =& \frac{p^*(s_T|\tau, f_{t,\tau}, z_t)}{p(s_T|\tau, f_{t,\tau}, z_t)} \\ \pi\left(z_T|\tau, f_{t,\tau}, z_t\right) =& \frac{q^*(z_T|\tau, f_{t,\tau}, z_t)}{q(z_T|\tau, f_{t,\tau}, z_t)} \end{aligned}$$

That is, by combining the risk-neutral and physical densities of S_t and Z_t, we obtain the projections of π onto S_T and Z_T, respectively. These two pricing kernels contain rich information on how risks, especially those associated with volatility shocks, are priced in financial markets. In the equilibrium setup of Aït-Sahalia and Lo (2000) with a representative agent, these pricing kernels represent—up to a scaled factor—the marginal rate of substitution. While Aït-Sahalia and Lo (2000) and Jackwerth (2000) estimate the pricing kernels of S&P 500 returns, our $\pi\left(s_T|\tau, f_{t,\tau}, z_t\right)$ includes the VIX z_t in the conditional information set so that volatility becomes relevant to the price of risk regarding the expected returns. In addition, we are able to identify the pricing kernel of the VIX.

2.3 Nested Models

As discussed in Section 2.1, we employ the information set generated by $F_{t,\tau}$ and Z_t to replace the information generated by $F_{t,\tau}$ and V_t, because Z_t is directly observable. In fact, the information set of $F_{t,\tau}$ and Z_t is coarser than the set generated by $F_{t,\tau}$ and V_t, and equating these two effectively assumes that V_t is an invertible function of $F_{t,\tau}$ and Z_t. We now show that this assumption is satisfied in most parametric models proposed in the literature, including both reduced-form option pricing models and equilibrium models with a priced stochastic volatility factor. Unlike Boes et al. (2007) and Li and Zhao (2009) who use an ex-post volatility proxy filtered from historical time series, we use VIX instead, which bears no approximation errors in most cases.

We first consider the class of option pricing models that induce an affine relationship between the unobservable variance and squared VIX. This class of models has the following risk-neutral dynamics:[9]

$$\begin{aligned} dS_t &= (r - d - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^Q + dL_t^S \\ dV_t &= \kappa(\xi - V_t)dt + \sigma(V_t)dB_t^Q + dL_t^V \end{aligned} \tag{6}$$

where dL_t^S and dL_t^V may be driven by finite activity compound Poisson processes with correlated jump sizes J_t^S and J_t^V. Such models include those discussed in Bakshi et al. (1997), Bates (2000), Pan (2002), Chernov and Ghysels (2000), Eraker (2004), Carr et al. (2003), Eraker et al. (2003), and Broadie et al. (2007). Jumps can be driven by Lévy processes such as the CGMY process in Carr et al. (2003) and Bates (2012). Note that this class also includes non-Gaussian OU processes, as introduced in Barndorff-Nielsen and

[9] The discontinuous part of the quadratic variation of S_t is assumed to be linear in V.

Shephard (2001); see Shephard (2005) for a collection of similar models.

For models of this class, we have

$$Z_t^2 = aV_t + b,$$

where a and b are functions of model parameters (see Carr and Wu (2009) for details). That is, Z_t^2 is a linear function of V_t, hence Z_t and V_t deliver the same information set.

The second class of models introduces a non-affine structure between the squared VIX and variance, such as the exponential-OU-L models in Shephard (2005). In particular, under the risk-neutral measure, such models specify the volatility process as

$$\log V_t = \alpha + \beta F_t, \quad dF_t = \kappa F_t dt + dL_t,$$

The squared VIX, as calculated by Tauchen and Todorov (2011), is

$$Z_t^2 = \frac{1}{\tau}\int_0^\tau \gamma + (\eta+1)\exp\Big(\alpha + e^{\kappa u}(\log V_t - \alpha) + C(u)\Big)du, \tag{7}$$

where $C(u)$ is determined by the characteristic exponent of the Lévy process L_t, and γ and η are constants determined by the quadratic variation of L_t^S. Observe that the function $V_t \mapsto Z_t$ is invertible, so that the information sets generated by V_t and by Z_t are equivalent.

Finally, we consider a stylized general equilibrium model, which is a simplified version of Bollerslev et al. (2009) and Drechsler and Yaron (2011) that builds on the long-run risk framework of Bansal and Yaron (2004).[10] Specifically, the representative agent's preference over consumption is recursive (Epstein and Zin (1989)). Therefore, the log pricing kernel at time t+1 is

$$m_{t+1} = \theta \log\delta - \theta\psi^{-1}\Delta c_{t+1} + (\theta-1)r_{c,t+1}, \tag{8}$$

where $\theta = (1-\gamma)/(1-\psi^{-1})$, $0 < \delta < 1$ is the subjective discount factor, γ is the risk

[10] Other equilibrium models that satisfy the invertibility between Z_t and V_t include Bansal et al. (2012), Bollerslev et al. (2009), and Campbell et al. (2012). We choose to present the framework of Bollerslev et al. (2009) and Drechsler and Yaron (2011) for simplicity of illustration.

aversion coefficient, ψ is the intertemporal elasticity of substitution, Δc_{t+1} is the growth rate of log consumption, and $r_{c,t+1}$ is the time t to t+1 return on the aggregate wealth claim.[11] The state vector of the economy follows

$$\begin{aligned}\Delta c_{t+1} =& \mu_c + \sigma_{c,t} z_{c,t+1} + J_{c,t+1} \\ \sigma^2_{c,t+1} =& \mu_\sigma + \rho_\sigma \sigma^2_{c,t} + \sigma_{c,t} z_{\sigma,t+1} + J_{\sigma,t+1}\end{aligned} \tag{9}$$

where $\{z_{c,t}\}$ and $\{z_{\sigma,t}\}$ are independent i.i.d. N(0,1) processes, $J_{c,t+1}$ is a compound Poisson process with intensity $\lambda_{c,t}$ and i.i.d. jump size ζ_i^c, $J_{\sigma,t+1}$ is a compound Poisson process with intensity $\lambda_{\sigma,t}$ and i.i.d. jump size ζ_i^σ, and both jump processes are independent of each other and of the Gaussian shocks. Note that both the Gaussian process $z_{\sigma,t+1}$ and jump process $J_{\sigma,t+1}$ contribute to volatility shocks.

By the standard log-linearization approach following Campbell and Shiller (1988), we have

$$r_{c,t+1} = \kappa_0 + \kappa_1 w_{t+1} - w_t + \Delta c_{t+1}, \tag{10}$$

where the price-wealth ratio w_t is conjectured to be affine in the state vector:

$$w_t = A_0 + A_\sigma \sigma^2_{c,t} \tag{11}$$

with $A_0 > 0$ and $A_\sigma < 0$ as functions of the model parameters we suppress for notational brevity. With (10) and (11), we have

$$r_{c,t+1} = \Delta c_{t+1} + \kappa_1 A_\sigma \sigma^2_{c,t+1} - A_\sigma \sigma^2_{c,t} + \kappa_0 + \kappa_1 A_0 - A_0.$$

Therefore, the volatility factor $\sigma^2_{c,t+1}$ shows up in $r_{c,t+1}$, and hence in the pricing kernel m_{t+1} given in (8). Following the standard practice to proxy the aggregate wealth (consumption) by the aggregate stock market S_t (see Aït-Sahalia and Lo (2000), Bansal and Yaron (2004),

[11] The literature usually assumes that $\gamma > 1$ and $\psi > 1$, which implies $\theta < 0$. This assumption ensures that the representative agent has a preference for early resolution of uncertainty, which is the key for the price of volatility risk.

and Campbell et al. (2012)), the return $S_T - S_t$ corresponds to Δc_{t+1}, and Z_t corresponds to the square root of the risk-neutral expectation of the consumption growth variance $\sigma^2_{c,t}$. Although the state vector dynamics are specified in discrete time, the model (9) is actually a special case of the affine model in Bollerslev et al. (2012). Therefore, Z_t is an invertible function of $V_t = \sigma^2_{c,t+1}$, which represents the variance of consumption growth rate under this equilibrium model.

In summary, most parametric models with volatility risk proposed in the literature, whether reduced-form or structural, can be nested within our nonparametric framework. As a result, we do not lose any information about the dynamics of S_t and V_t when incorporating Z_t into the information set; instead, the implementation becomes feasible with the information set fully observable.

3 Estimation Strategy

3.1 Multivariate Local Linear Estimators for Densities

Here we introduce our nonparametric estimation strategies for SPDs. To fix ideas, we assume the observed prices, $\tilde{C}$ and $\tilde{H}$, are contaminated with observation errors, such that[12]

$$C(\tau, f_{t,\tau}, z_t, x) = E\left(\tilde{C}\middle|\tilde{\tau} = \tau, F_{t,\tau} = f_{t,\tau}, Z_t = z_t, X = x\right)$$
$$H(\tau, f_{t,\tau}, z_t, y) = E\left(\tilde{H}\middle|\tilde{\tau} = \tau, F_{t,\tau} = f_{t,\tau}, Z_t = z_t, Y = y\right).$$

We then construct nonparametric estimators of C and H, and take derivatives to estimate the SPDs. Different from the multivariate kernel regression approach adopted by Aït-Sahalia and Lo (1998), we prefer the local linear estimator (Fan and Gijbels (1996)) for two main reasons. First of all, the bias and variance of local polynomial estimators are of the same

[12]Hereafter, we multiply all option prices by the corresponding $e^{r_{t,\tau}\tau}$, so that we can omit $r_{t,\tau}$ in C and H, and reduce one state variable in the following regressions. Again, we recycle the notations C and H without ambiguity.

order of magnitude in the interior or near the boundary, whereas kernel estimators are notorious for the boundary effects. As our empirical studies focus on the tail of pricing kernels, it is advantageous to adopt more efficient estimators. Second, local polynomial regression provides estimates of derivatives, in addition to option prices, which makes it more convenient for our purpose.

Theoretically, it is better to use a local cubic estimator to obtain second-order derivatives. Since we have more than one state variable, including all cross-terms of cubic polynomials into the regression is cumbersome. We avoid this by applying the local linear estimator, so that estimators for SPDs can be obtained simply by a first-order differentiation with respect to the strike.

We write the option price C as a function of $\mathbf{u} = (\tau, f, z, x)'$, and consider the following minimization problem,

$$\min_{\alpha,\boldsymbol{\beta}} \sum_{i=1}^{n} \left\{C_i - \alpha - \boldsymbol{\beta}' \left(\mathbf{u}_i - \mathbf{u}\right)\right\}^2 K_h \left(\mathbf{u}_i - \mathbf{u}\right)$$

where $\mathbf{u}_i = (\tau_i, f_{t_i,\tau_i}, z_{t_i}, x_i)'$ and C_i are the characteristics and price respectively of the i-th option in the sample. K_h is a kernel function scaled by a bandwidth vector $\mathbf{h} = (h_\tau, h_f, h_z, h_x)'$:

$$K_h(\mathbf{u}_i - \mathbf{u}) = \frac{1}{h_\tau} k\left(\frac{\tau_i - \tau}{h_\tau}\right) \frac{1}{h_f} k\left(\frac{f_{t_i,\tau_i} - f}{h_f}\right) \frac{1}{h_z} k\left(\frac{z_{t_i} - z}{h_z}\right) \frac{1}{h_x} k\left(\frac{x_i - x}{h_x}\right) \tag{12}$$

where $k(\cdot)$ is, for example, the density a of standard normal distribution. The minimizer has a closed-form representation:

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\boldsymbol{\beta}} \end{bmatrix}_{(1+4)\times 1} = \left(\boldsymbol{\Omega}'\mathbf{K}\boldsymbol{\Omega}\right)^{-1} \boldsymbol{\Omega}'\mathbf{K}\mathbf{C} \tag{13}$$

where

$$\mathbf{\Omega} = \begin{bmatrix} 1 & (\mathbf{u}_1 - \mathbf{u})' \\ \vdots & \vdots \\ 1 & (\mathbf{u}_n - \mathbf{u})' \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} K_h(\mathbf{u}_1 - \mathbf{u}) & & \\ & \ddots & \\ & & K_h(\mathbf{u}_n - \mathbf{u}) \end{bmatrix}.$$

The nonparametric local linear estimator for the option pricing function $C(\tau, f, z, x)$ is

$$\widehat{C}(\tau, f, z, x) = \widehat{\alpha} = \mathbf{e}_1' \left(\mathbf{\Omega}'\mathbf{K}\mathbf{\Omega}\right)^{-1} \mathbf{\Omega}'\mathbf{K}\mathbf{C}, \tag{14}$$

with $\mathbf{e}_1 = (1, 0, 0, 0)'$ and the estimator $\widehat{p^*}(s'|\tau, f, z)$ for the SPD of S_t is

$$\widehat{p^*}(s'|\tau, f, z) = e^{s'} \frac{\partial \widehat{\boldsymbol{\beta}}_4}{\partial x}\Big|_{x=e^{s'}} = e^{s'} \frac{\partial \left(\mathbf{e}_5' \left(\mathbf{\Omega}'\mathbf{K}\mathbf{\Omega}\right)^{-1} \mathbf{\Omega}'\mathbf{K}\mathbf{C}\right)}{\partial x}\Big|_{x=e^{s'}}. \tag{15}$$

where $\mathbf{e}_5 = (0, 0, 0, 0, 1)'$. The nonparametric estimator $\widehat{H}(\cdot)$ and $\widehat{q^*}(z'|\tau, s, z)$ can be constructed similarly.

It may be worth pointing out that $\widehat{\boldsymbol{\beta}}$ in our local linear regression (13) provides estimates of option Greeks. Specifically, option Theta is given by by $\mathbf{e}_1'\boldsymbol{\beta} = \partial C/\partial \tau$, Delta by $\mathbf{e}_2'\boldsymbol{\beta} \cdot e^s$, and Vega by $\mathbf{e}_3'\boldsymbol{\beta}$.

3.2 Dimension Reduction

One of the major issues of nonparametric estimation is the curse of dimensionality. The rate of convergence decreases rapidly as the dimension of state variables increases. In the most general forms, the pricing functions $C(\cdot)$ and $H(\cdot)$ depend not only on time-to-maturity, strike, VIX, and the S&P 500 index, but also on interest rates and dividends. Instead of regressing on additional interest rate and dividend variables, we assume that option prices multiplied by $e^{r_{t,\tau}\tau}$ depend on these variables only through forward prices. As mentioned in Aït-Sahalia and Lo (1998), models that violate this assumption seem very remote empirically.

Furthermore, following many existing studies such as Aït-Sahalia and Lo (1998) and Li and Zhao (2009), we assume that the S&P 500 option price is homogeneous of degree one in the forward price level:

$$C(\tau, f, z, x) = e^f C(\tau, 0, z, x/e^f) = e^f \bar{C}(\tau, z, m) \tag{16}$$

where $m = x/e^f$ represents the moneyness of the option. Consequently, we obtain the estimate of $C(\tau, f, z, x)$ through multiplying the nonparametric estimate of $\bar{C}(\cdot)$ by e^f, and write the SPD of S_T as

$$p^*(s_T|\tau, f_{t,\tau}, z_t) = e^{s_T - f_{t,\tau}} \frac{\partial^2 \bar{C}(\tau, z_t, m)}{\partial m^2}\Big|_{m=e^{s_T}/e^{f_{t,\tau}}}.$$

As for VIX options, we assume that the information about $Z_{t'}$ in $F_{t,\tau}$ is fully incorporated into Z_t. In other words, conditional on Z_t, $Z_{t'}$ is independent of $F_{t,\tau}$, for any $t' > t$. This assumption further implies that the SPD of Z_T, obtained from VIX option prices, depends on $F_{t,\tau}$ only through Z_t, i.e., $q^*(z_T|\tau, f_{t,\tau}, z_t) = q^*(z_T|\tau, z_t)$. Thus, the number of state variables for the SPD of VIX is also decreased by one. We conduct robustness checks for these assumptions in Section 4.5, and find supportive evidence. Our dimension reduction strategy is motivated from the economic intuition, which is in sharp contrast to the statistical approach proposed by Yao and Hall (2005), who discuss an alternative method in the context of conditional density estimation.

3.3 Estimation of Pricing Kernels

Given the homogeneity assumption in (16), the risk neutral density of the return R_T can be estimated using the following formula:

$$p^*(R_T|\tau, z_t) = e^{R_T - r_{t,\tau}\tau} \frac{\partial^2 \bar{C}(\tau, z_t, m)}{\partial m^2}\Big|_{m=e^{R_T - r_{t,\tau}\tau}},$$

where $R_T = s_T - f_{t,\tau}$. Note that homogeneous of degree one in option prices is equivalent to that the conditional density of the log returns is independent of s_t, see, e.g. Joshi (2007) for more details. This property is satisfied by all parametric models discussed in Section 2.3.

While estimating the risk neutral density from option prices, we estimate the physical density $p(R_T|\tau, z_t)$ using the time series of the S&P 500 index and VIX based on the local linear method. A similar strategy has been adopted by Aït-Sahalia et al. (2009). We collect time series of (R_{T_i}, z_{t_i}), $i = 1, \ldots, n$, with $\tau = T_i - t_i$ fixed. We then construct the local linear estimator of the conditional density of returns $p(R|\tau, z)$ by minimizing:

$$\min_{\gamma,\eta} \sum_{i=1}^{n} \left\{K_{b_R}(R_{T_i} - R) - \gamma - \eta' \left(z_{t_i} - z\right)\right\}^2 W_{b_z}\left(z_{t_i} - z\right)$$

where b_R and b_z are the bandwidths to be selected, and $K_{b_R}(\cdot) = 1/b_R \cdot k(\cdot/b_R)$ and $W_{b_z}(\cdot) = 1/b_z \cdot w(\cdot/b_z)$ are kernels. Therefore, our density estimator is,

$$\widehat{p}(R|\tau, z) = \widehat{\gamma}. \tag{17}$$

Consequently, our pricing kernel estimator can be constructed as

$$\widehat{\pi}(R|\tau, z) = \frac{\widehat{p}^*(R|\tau, z)}{\widehat{p}(R|\tau, z)}.$$

Similarly, we can construct the estimator for the pricing kernel of VIX.

3.4 Asymptotic Theory

To provide theoretical guidance for our approach, we derive the asymptotic distribution of the option price and density for S&P 500 options as an example. Suppose the sample size of the S&P 500 options is n. Using the equivalent kernels introduced in Fan and Gijbels (1996) and following the derivation in Aït-Sahalia and Lo (1998), we obtain:[13]

$$n^{1/2}\left(h_\tau h_f h_z h_x\right)^{1/2}\left(\widehat{C}(\tau,f,z,x)-C(\tau,f,z,x)\right) \tag{18}$$
$$\xrightarrow{d} N\left(0,\left[\int k^2(c)\,dc\right]^3 s^2(\tau,f,z,x)/\pi(\tau,f,z,x)\right),\quad \text{as}\quad nh_\tau h_f h_z h_x \longrightarrow \infty;$$
$$n^{1/2}h_x^2\left(h_\tau h_f h_z h_x\right)^{1/2}\left(\widehat{p}(s'|\tau,f,z)-p(s'|\tau,f,z)\right) \tag{19}$$
$$\xrightarrow{d} N\left(0,\left[\int k^2(c)\,dc\right]^3\left[\int\left(c\dot{k}(c)+k(c)\right)^2 dc\right]/\left[\int k(c)\,c^2dc\right]^2 s^2(\tau,f,z,s')/\pi(\tau,f,z,s')\right),$$
$$\text{as}\quad nh_\tau h_f h_z h_x^5 \longrightarrow \infty,$$

where $s^2(\tau,f,z,x)$ is the conditional variance for the local linear regression of C on the state variables, and $\pi(\tau,f,z,x)$ is the joint density of these variables. The estimator for $s^2(\cdot)$ can be constructed using similar nonparametric regressions of squared fitting errors on these state variables. The same asymptotic distributions apply to estimators for VIX option prices and their SPDs. Similar technique has been adopted in Ruppert and Wand (1994).

In addition to estimating the option price and its first and second order derivatives, we

[13] We sketch a proof here for the asymptotic theory as part of it is non-standard. Notice from (13) that

$$\begin{bmatrix}\widehat{\alpha}\\ \widehat{\boldsymbol{\beta}}\end{bmatrix}_{(1+4)\times 1} = \left(\boldsymbol{\Omega}'\mathbf{K}\boldsymbol{\Omega}\right)^{-1}\boldsymbol{\Omega}'\mathbf{K}\mathbf{C}.$$

Using the properties of Gaussian kernel, we have

$$\frac{1}{n}\left(\boldsymbol{\Omega}'\mathbf{K}\boldsymbol{\Omega}\right)^{-1} = \begin{bmatrix}\frac{1}{n}\sum_{i=1}^n K_h(\mathbf{u}_i-\mathbf{u}) & \frac{1}{n}\sum_{i=1}^n K_h(\mathbf{u}_i-\mathbf{u})(\mathbf{u}_i-\mathbf{u})' \\ \frac{1}{n}\sum_{i=1}^n K_h(\mathbf{u}_i-\mathbf{u})(\mathbf{u}_i-\mathbf{u}) & \frac{1}{n}\sum_{i=1}^n K_h(\mathbf{u}_i-\mathbf{u})(\mathbf{u}_i-\mathbf{u})(\mathbf{u}_i-\mathbf{u})'\end{bmatrix}$$
$$\xrightarrow{P}\begin{bmatrix} f(\mathbf{u}) & \mathbf{0}' \\ \mathbf{0} & f(\mathbf{u})\cdot\int c^2k^2(c)dc\cdot\text{diag}(h_\tau^2,h_f^2,h_z^2,h_x^2)\end{bmatrix},\quad \text{as } h\to 0, n\to 0.$$

Therefore, we can write the estimators in their equivalent kernel forms:

$$\widehat{\alpha}\approx\frac{1}{nf(\mathbf{u})}\sum_{i=1}^n K_h(\mathbf{u_i}-\mathbf{u})\cdot C_i$$
$$\widehat{\beta}_4\approx\frac{1}{nh_x^2 f(\mathbf{u})\int c^2k^2(c)dc}\sum_{i=1}^n K_h(\mathbf{u_i}-\mathbf{u})(x_i-x)\cdot C_i$$

Using the standard kernel asymptotic results, we can obtain the above asymptotic theory.

apply a local linear method to estimate the conditional density in (17). Its asymptotic theory is given by (see, e.g. Fan et al. (1996)):

$$n^{1/2}(b_r b_z)^{1/2}\Big(\widehat{\widetilde{p}}(r'|\tau,z)-\widetilde{p}(r'|\tau,z)\Big)\xrightarrow{d} N\left(0,\left[\int k^2(c)dc\right]\left[\int w^2(c)dc\right]\widetilde{p}(r'|\tau,z)/\pi(z)\right).$$

The asymptotic theories provided here are applied to construct confidence bands in our empirical studies.

3.5 Bandwidth Selection

Bandwidth selection is important especially for multivariate nonparametric regressions. In theory, the optimal rate of bandwidth for estimating the option price is $n^{-1/(4+d)}$, whereas to estimate densities, we need to adopt a bandwidth with rate $n^{-1/(6+d)}$ due to the curse of differentiation. These bandwidth choices ensure that the nonparametric pricing function achieves the optimal rate of convergence in the mean-squared sense. Empirically, we can choose a bandwidth h_j ($j = \tau, z$, and m for S&P 500 options, or y for VIX options) as $h_j = c_j\sigma_j n^{-1/(4+d+2\nu)}$, where σ_j is the unconditional standard deviation of the regressor j, d is the number of regressors, and $\nu = 0$ and 1 for option prices and SPDs, respectively. The constant c_j is chosen by minimizing the mean-squared error of option prices via cross-validation. The cross-validation objective function for regression (14) is given by the weighted mean squared errors:

$$\min_h \frac{1}{n}\sum_{i=1}^{n}\Big(C_i-\widehat{C}_{h,-i}(\tau_i,f_i,z_i,x_i)\Big)^2\omega(\tau_i,f_i,z_i,x_i)$$

where $-i$ means leaving the ith observation out, and $\omega(\cdot)$ is the weighting function. To further accelerate the cross-validation, we adopt the popular K-fold cross-validation, which is faster compared with this leave-one-out method.

The bandwidths of our nonparametric conditional density estimator (17) under the physical measure are chosen by the cross-validation following Fan and Yim (2004):

$$\min_b \frac{1}{n}\sum_{i=1}^{n}\omega(s_{t_i}, z_{t_i})\int(\widehat{p}_b(s'|\tau, s_{t_i}, z_{t_i}))^2 ds' - \frac{2}{n}\sum_{i=1}^{n}\widehat{p}_{b,-i}(s_{T_i}|\tau, s_{t_i}, z_{t_i})\omega(s_{t_i}, z_{t_i}).$$

where the first integral can be calculated in closed-form from (17). Alternative choices of bandwidths have been discussed in Yao and Tong (1998) and Ruppert et al. (1995).

3.6 Monte Carlo Simulations

Here we provide simulation studies of our local linear estimators. The Monte Carlo experiments are designed to match our empirical studies. First, we select the same option characteristics as those traded on CBOE in our sample. Second, we select a sample path generated from the following stochastic volatility models with both jumps in volatility and prices:

$$\begin{aligned} dS_t &= (r - d - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^Q + J_S^Q dN_t - \mu\lambda_t dt \\ dV_t &= \kappa(\xi - V_t)dt + \sigma\sqrt{V_t}dB_t^Q + J_V^Q dN_t \end{aligned}$$

where W_t^Q and B_t^Q are standard Brownian motions satisfying $E(dW_t^Q dB_t^Q) = \rho dt$, J_S^Q and J_V^Q are random jump sizes, dN_t is a pure-jump process with intensity $\lambda_t = \lambda_0 + \lambda_1 V_t$, and $\mu = E(e^{J_S^Q} - 1)$. The jump sizes follow:

$$J_V^Q \sim exp(\beta_V), \quad J_S^Q \sim \begin{cases} \quad exp(\beta_+) & \text{with probability } q \\ - \quad exp(\beta_-) & \text{with probability } 1 - q \end{cases}$$

The parameters are taken from Amengual and Xiu (2012), where $\kappa = 2$, $\sigma = 0.3$, $\rho = -0.8$, $\xi = 0.04$, $\beta_+ = 0.01$, $\beta_- = 0.03$, $q = 0.3$, $\beta_V = 0.02$, $\lambda_0 = 2$, and $\lambda_1 = 30$. We then calculate S&P 500 and VIX option prices, according to the closed-form formulae given in Amengual and Xiu (2012). Finally, we pollute the prices with multiplicative measurement errors following log-normal distribution with a 5% standard deviation.

Based on the generated sample, we evaluate our nonparametric estimators of option prices on the grid of time-to-maturity and current index level, with the VIX Z_t and strike X fixed at their sample median. We also calculate the index densities on the grid of τ and S_T, with S_t fixed at the sample mean, to evaluate our density estimators. The nonparametric estimators of VIX option prices and densities are evaluated similarly. All of these quantities and their percentage errors are reported in Figure 1, averaged over 1000 replications. We observe that the nonparametric estimates are within 5% and 10% of their theoretical Black-Scholes implied volatilities for S&P 500 and VIX options, respectively. The errors for densities are slightly larger, due to the fact that derivatives are estimated with slower rates of convergence, i.e., the so-called curse of differentiation.

4 Empirical Results

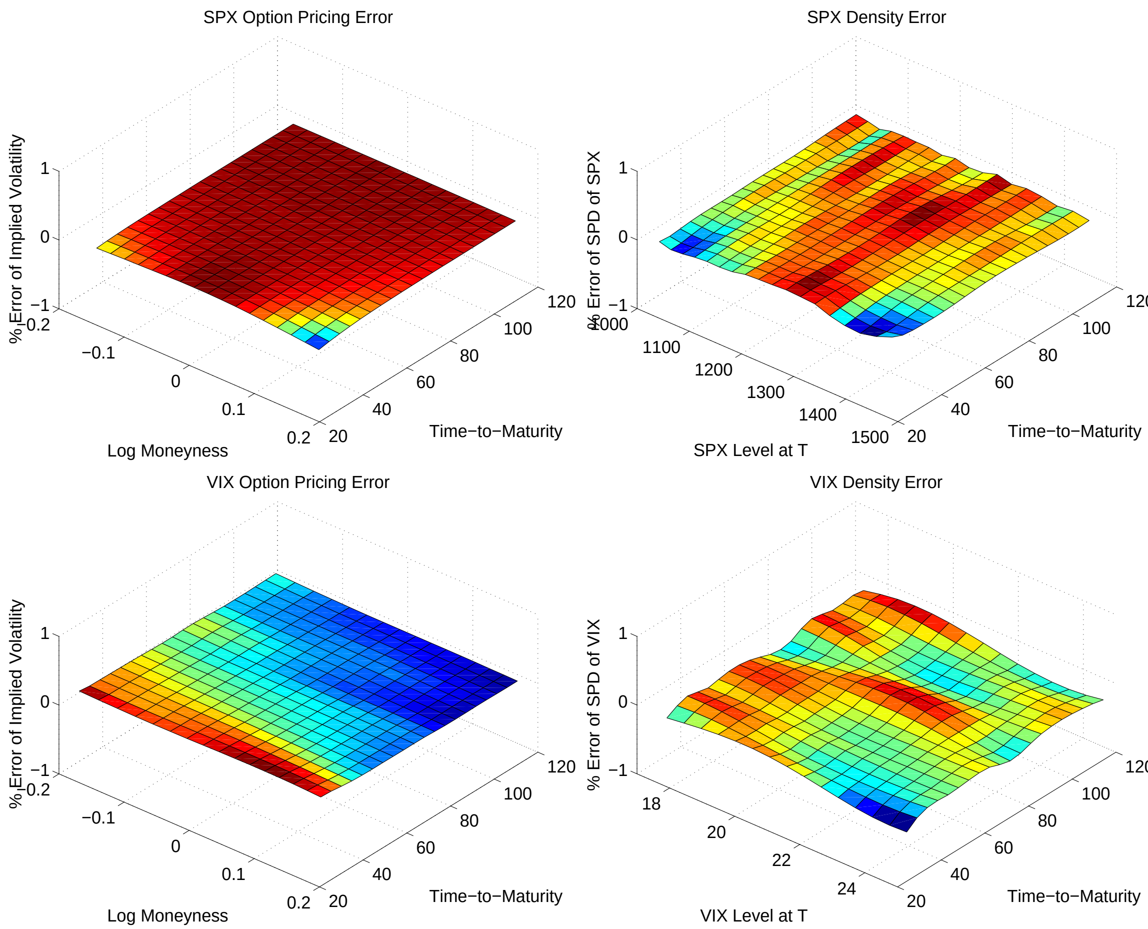
In this section, we estimate nonparametric SPDs and pricing kernels using both S&P 500 and VIX options, and present our empirical findings. Before delving into the details, we introduce the dataset.

4.1 Data

We obtain daily bid and offer prices of S&P 500 and VIX options, quoted between 3:59 p.m. and 4:00 p.m. EST from the OptionMetrics. Our sample period is chosen as June 1, 2009–May 31, 2011, during which the liquidity of VIX options is satisfactory. We plot the daily open interests of VIX options in Figure 2, along with those of S&P 500 options for comparison. It is obvious from the figure that the liquidity of VIX options has improved dramatically since introduced in 2006, and their open interests have achieved roughly 1/4 of those of S&P 500 options. As a result, our choice of sample ensures that our empirical results are not subject to liquidity issues.

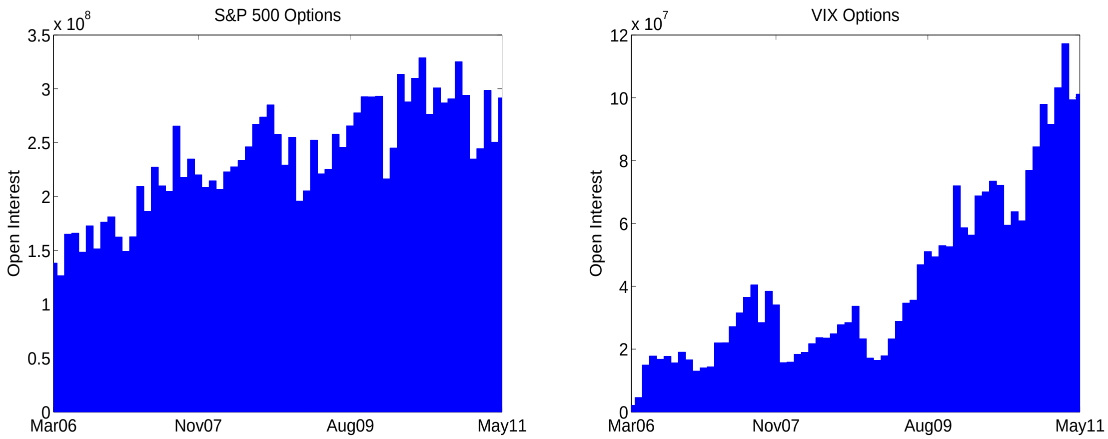
Figure 3 plots the joint time series of the S&P 500 index and VIX over the sample period,

Figure 1: Monte Carlo Simulations



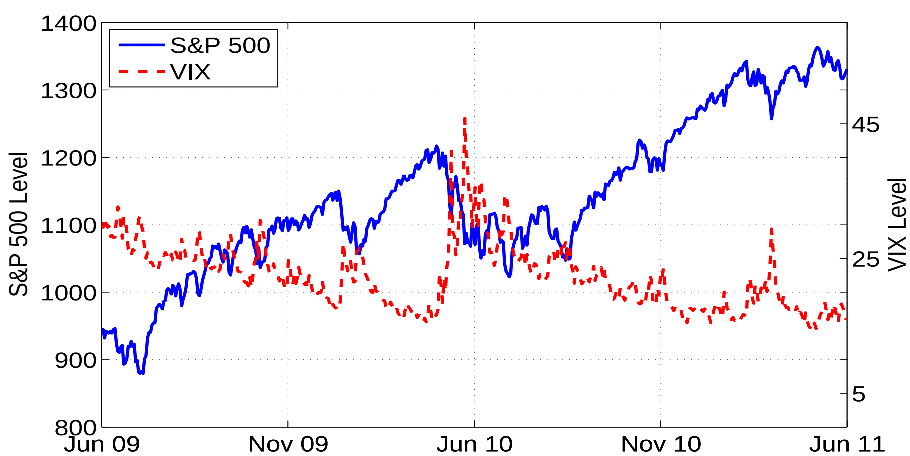
Note: This figure plots the nonparametric estimation error in the Monte Carlo simulations. The left panel plots the pricing error measured in terms of difference in implied volatility, whereas the right panel plots the percentage error in density estimates. The number of Monte Carlo samples is 1000.

Figure 2: Open Interests of the S&P 500 and VIX Options



Note: This figure plots the monthly time series of the open interests of S&P 500 and VIX options from March 1, 2006 to May 31, 2011.

Figure 3: Time Series of the S&P 500 Index and the VIX



Note: This figure plots the time series of the S&P 500 index and VIX from Jun 1, 2009 to May 31, 2011.

while Table 1 provides their summary statistics. We observe that the VIX ranges between 14.62 and 45.79, which is large enough to have both relatively low and high volatility levels. Moreover, Table 2 also presents summary statistics of option prices. It is worth pointing out that the differences between adjacent strikes of VIX options range from $0.50 to $5 for smaller strikes and from $1 to $10 for large strikes, which are significantly larger percentage-wise than their counterparts for S&P 500 options. Therefore, the impact of price discreteness on the nonparametric estimation of VIX densities could be more severe than on the density estimation of the S&P 500 index, as discussed in Section 3.6.

Table 1: Summary Statistics of the S&P 500 Index and VIX

	Mean	Std	Skew	Kurt	Min	25%	75%	Max
Index	1141.670	115.806	0.065	2.414	879.130	1070.453	1221.178	1363.610
Return	0.001	0.011	-0.330	4.822	-0.040	-0.004	0.006	0.043
VIX	22.307	4.900	0.891	4.173	14.620	18.000	25.153	45.790

Note: This table reports the summary statistics of the time series of S&P 500 index, return, and VIX from June 1, 2009 to May 31, 2011.

We follow the data-cleaning routine commonly used in the literature; see, e.g., Aït-Sahalia and Lo (1998). First, observations with bid or ask prices smaller than $0.025 are eliminated to mitigate the effect of pricing errors. For each option, we take the midquote as the observed option price. Due to liquidity concerns, we eliminate any options with zero open interests or trading volumes as well as options with time-to-maturity of less than 5 days. In addition, we only consider options with maturity of less than 136 days, because only VIX option contracts with maturities shorter than 6 months are offered by the CBOE after 2009. It is well known that in-the-money S&P 500 options are less liquid than out-of-the-money options. Therefore, we delete in-the-money options, and use the put-call parity to construct prices of in-the-money call options from out-of-the-money put options. There is no such pattern

Table 2: Summary Statistics of S&P 500 and VIX Options

		SPO			VXO		
Moneyness		ITM	ATM	OTM	ITM	ATM	OTM
# of Records		72161	27144	33245	5573	16823	17726
Volume		102.12	82.73	35.76	2.43	37.55	32.86
Open Interest		1665.98	616.39	513.05	44.57	356.76	471.44
	min	41.07	0.13	0.05	2.03	0.03	0.03
	25%	120.19	18.15	0.35	6.40	1.83	0.18
Derivative Prices	50%	186.63	33.36	1.73	8.55	3.05	0.50
	75%	284.04	49.00	6.40	11.70	4.55	1.03
	max	1128.50	110.67	60.00	25.90	11.70	4.75
	min	100	845	915	10	14	22.5
	25%	825	1075	1160	15	21	35
Strike Prices	50%	940	1135	1225	17	25	42.5
	75%	1040	1250	1325	20	30	50
	max	1305	1415	3000	40	65	100
	min	5	5	5	5	5	5
	25%	18	17	21	21	27	28
Time-to-Maturity	50%	32	32	35	44	50	50
	75%	53	51	56	77	82	78
	max	136	136	136	128	128	126
	min	16.05	9.39	9.48			
	25%	27.56	16.67	15.31			
Implied Volatility	50%	33.85	19.84	18.13			
	75%	43.41	23.37	20.84			
	max	143.48	45.81	49.44			

Note: This table reports the summary statistics (minimum, quantiles, and maximum) for selected S&P 500 and VIX option quotes from June 1, 2009 to May 31, 2011, including the number of records, trading volume, open interest, option price, strike price, time-to-maturity, and implied volatility. In total, there are 132,550 trading records for S&P 500 options, and 40,122 records for VIX options. All options are call options. The prices of S&P 500 ITM call options are computed from OTM put options using the put-call parity for liquidity concerns. For S&P 500 options, ATM is defined as $K/F \in [0.96, 1.04]$, whereas for VIX options, it is defined as $K/\text{VIX} \in [0.9, 1.5]$, with F as the forward price and K as the generic strike.

for VIX options and hence we only consider VIX call options. The last step is to eliminate option contracts that violate no-arbitrage conditions. The resulting sample covers a broad cross section of options, including 420,711 S&P 500 call options, and 53,530 VIX call options, which account for 50.84% and 54.18% of their total number of records, respectively.

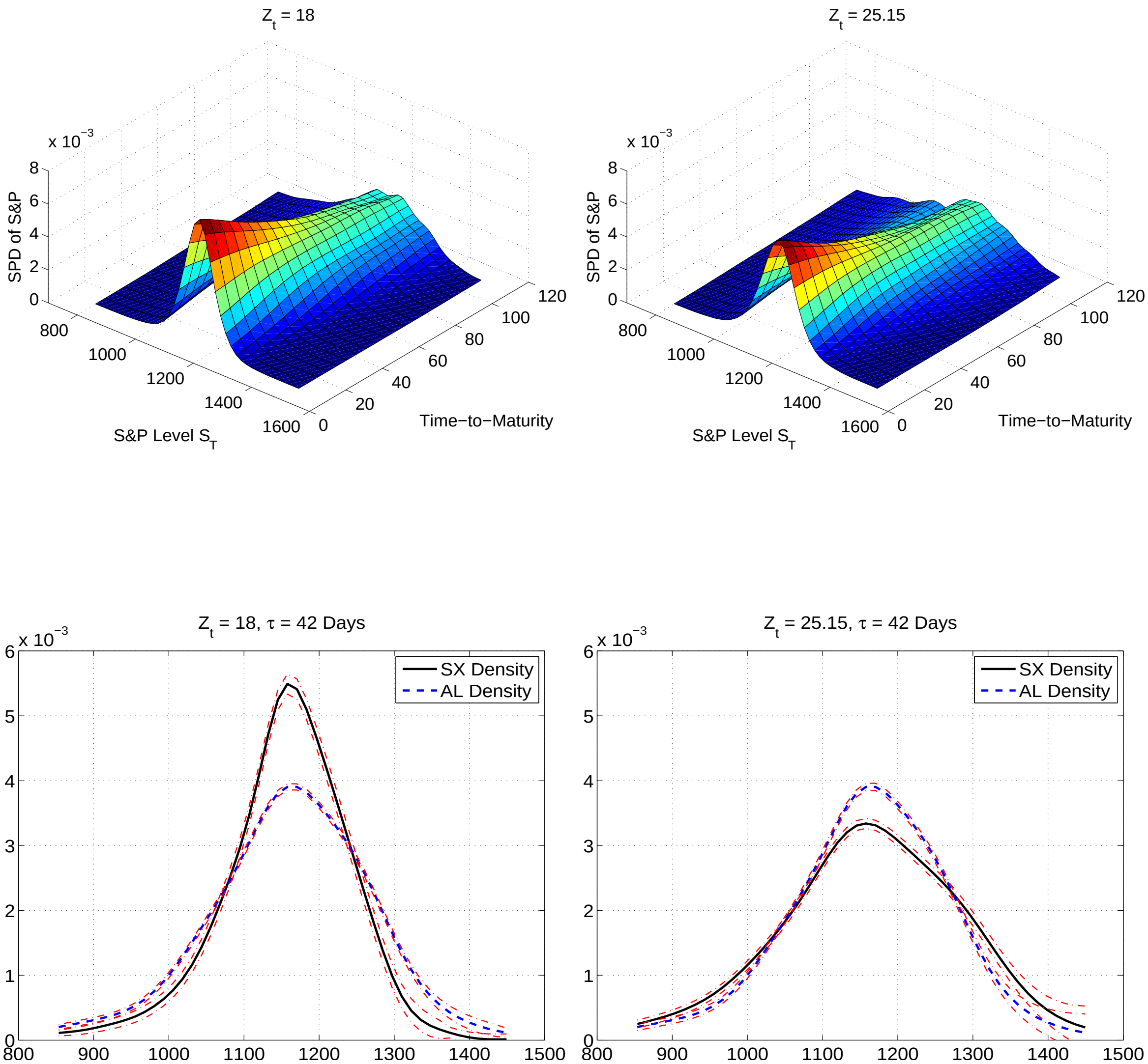
4.2 Pricing Kernels of the Market Return

The upper panels of Figure 4 provide nonparametric SPD estimates of the S&P 500 index for both low and high levels of VIX, fixed at 18.00 and 25.15 that correspond to the 25% and 75% quantiles of the VIX time series in our sample, respectively. We observe that index densities strongly depend on the VIX level Z_t. Conditional on a low Z_t, $p(s_T|\tau, s_t, z_t)$ has pronounced spikes, while the density becomes more dispersed when Z_t rises to a high level, suggesting that volatility is a key state variable that should be included in the SPDs.

To further demonstrate the importance of volatility in studying the SPDs of market return, the bottom panels of Figure 4 compare the nonparametric SPD estimates proposed by Aït-Sahalia and Lo (1998) (AL) who neglect the volatility variable, with $p(s_T|\tau, s_t, z_t)$ conditional on the two different VIX levels of 18.00 and 25.15. We choose the time-to-maturity as 42 days, and compute the 95% confidence intervals by the asymptotic theory given in (19). Observe that our SPDs differ from the AL densities substantially, with the former more compact and showing higher spikes for a low Z_t, which confirms the importance of incorporating volatility into the SPDs.

Given the importance of volatility in SPDs of the S&P 500 index documented above, we now study whether and how volatility affects the shape of the pricing kernel $p(R_T|\tau, z_t)$. According to Section 3.3, we further estimate the physical densities of the S&P 500 return conditional on VIX and obtain the pricing kernel estimates. The top two panels of Figure 5 report the pricing kernel estimates with Z_t equal to 18.00 (left) and 25.15 (right) and a maturity of 42 days. We observe that the pricing kernels conditional on either a low or high

Figure 4: State-Price Densities of the S&P 500 Index



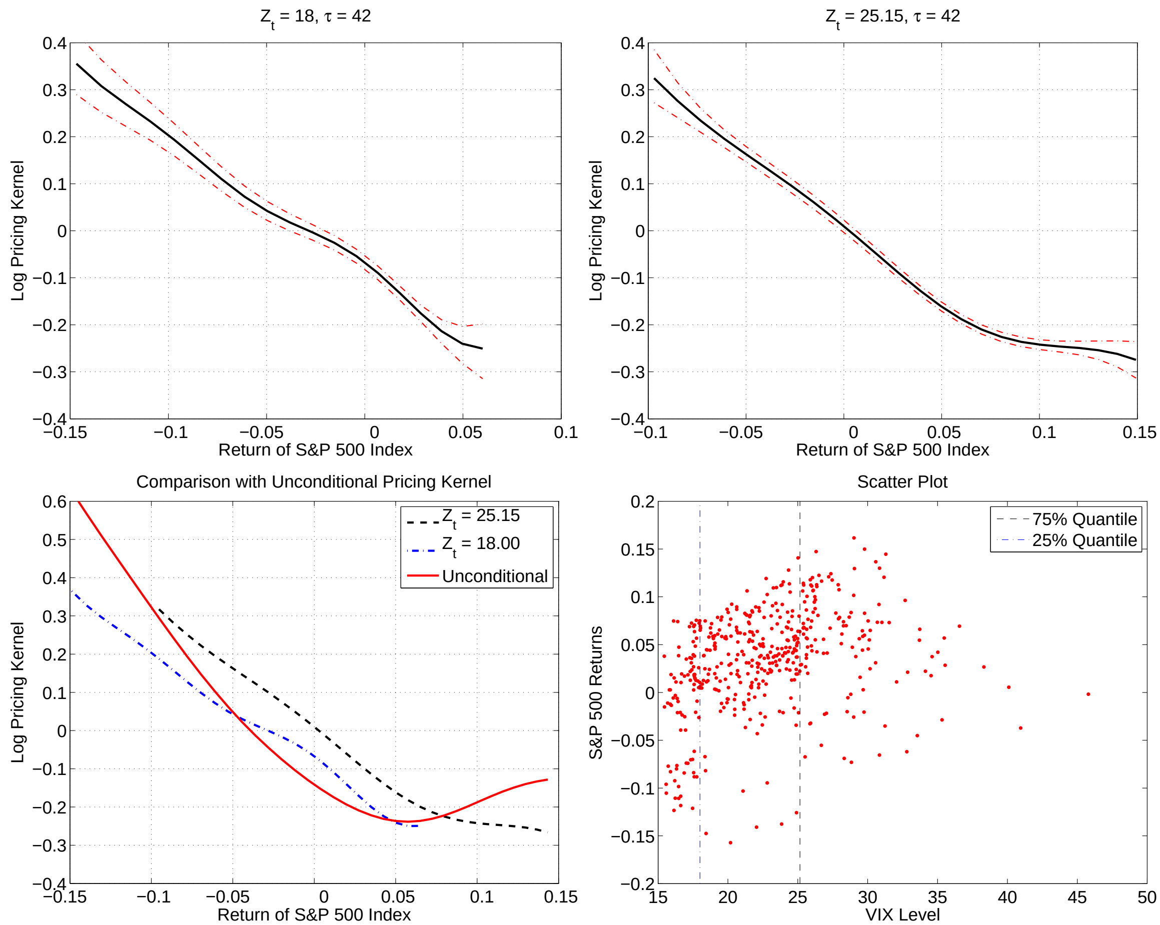
Note: The top panels provide our nonparametric estimates of SPDs of the S&P 500 index at various time-to-maturities, with volatility levels at 18.00 (left) and 25.15 (right) that correspond to the 25% and 75% quantiles of the VIX time series in our sample, respectively. The bottom panels compare our estimates (SX) of index SPDs (black, solid) with those using the Aït-Sahalia and Lo (1998) (AL) method (blue, dashed) for the maturity of 42 days, and two current VIX levels at 18.00 and 25.15. Dotted lines around each SPD estimate are the 95% confidence intervals constructed by the asymptotic distribution theory in (19). The interest rate and dividend are fixed at their averages, 2.15% and 2.06%, respectively.

VIX level exhibit a decreasing shape, consistent with a standard expected utility theory, which prescribes that the pricing kernel decrease when expected returns are increasing. In contrast, the bottom left panel of Figure 5 shows that the unconditional estimator of the pricing kernel shows a pronounced U-shape, consistent with what have been found in the literature (Jackwerth (2000) and Bakshi et al. (2010)). Therefore, it is the volatility factor, missing in the unconditional estimates, that may lead to the puzzling U-shape.

Specifically, high volatility signals bad future investment opportunities, and investors should have high marginal utility in such a state. Hence, the pricing kernel of market return conditional on a high volatility, which equals the marginal utility up to a re-scaling, is higher than that conditional on a low volatility, as shown in the bottom left panel of Figure 5. The unconditional pricing kernel, however, is a mixture of pricing kernels conditional on different volatility levels, and could exhibit a U-shape when volatility switches from low to high levels. Our finding echoes the conclusions of parametric models in Jackwerth and Brown (2001), Chabi-Yo et al. (2008), Chabi-Yo (2011), and Christoffersen et al. (2010), that missing state variables in the pricing kernel may result in the U-shape. Without restricting the specification of pricing kernels, however, our result shows that stochastic volatility is the key but missing state variable of pricing kernels estimated in the literature.

The pricing kernels conditional on low and high values of Z_t have different supporting regions on the left and right tail. For instance, over the interval (0.08, 0.15), we only have the pricing kernel estimates conditional on a high Z_t. The reason is that the realized return R_T never exceeds 8% given $Z_t = 18.00$, as can be seen from the scatter plot of (R_T, Z_t) on the bottom right panel of Figure 5. In fact, this observation implies that the unconditional pricing kernel estimates around high levels of market return R_T are dominated by high level volatility, which shifts the unconditional estimates upwards, and explains why they present a U-shape. In other words, large market returns R_T are accompanied by high current volatility Z_t, because of which investors have a high marginal utility that leads to the increasing portion

Figure 5: Pricing Kernels of the S&P 500



Note: The top panels plot the nonparametric estimates of pricing kernels of the S&P 500 index return (black, solid) for the maturity of 42 days, with two current VIX levels at 18.00 and 25.15 that correspond to the 25% and 75% quantiles of the VIX time series in our sample, respectively. Dotted lines are the 95% confidence intervals. The bottom left figure compares the unconditional pricing kernel (red, solid) with the previous two conditional pricing kernels. The bottom right panel presents the scatter plot of S&P 500 returns R_T against the current VIX level Z_t.

of the unconditional pricing kernel on the right tail.

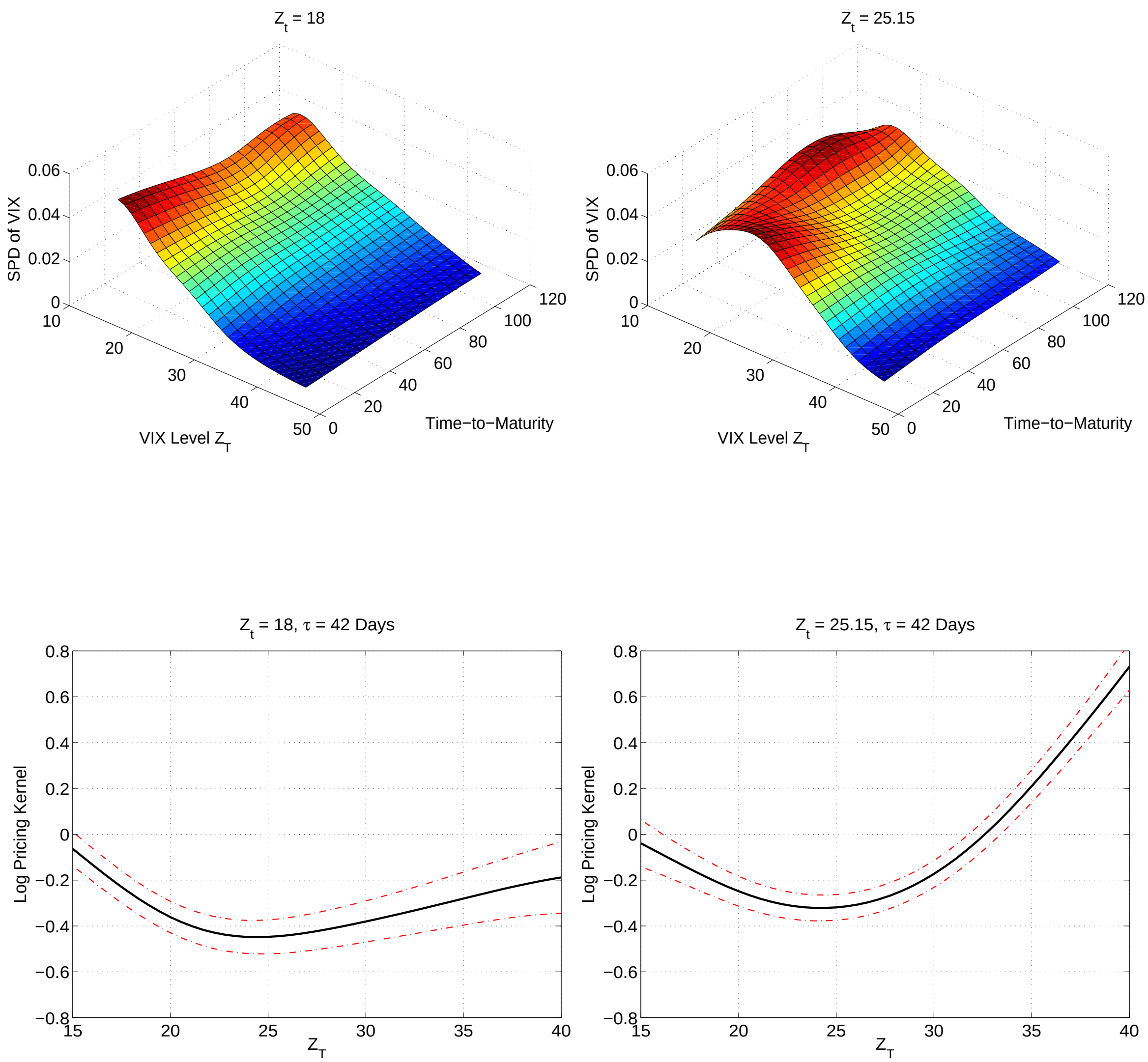
Overall, our nonparametric state-price density estimates differ significantly from those without conditioning on volatility, and confirms that volatility is a key state variable that should be included in the pricing kernel. More importantly, without imposing any restrictions on the dynamics of the market return and volatility, our pricing kernel estimates conditional on VIX show that stochastic volatility is the key variable responsible for the "pricing kernel puzzle."

4.3 Pricing Kernels of the VIX

We now present nonparametric estimates of SPDs and pricing kernels of the VIX and investigate their implications for the pricing of volatility risk. The top panels of Figure 6 present the VIX SPDs at various maturities conditional on two different levels of Z_t equal to 18.00 and 25.15. We find first that the VIX SPDs are all positively skewed, with the probability of achieving higher VIX levels decreasing given a low time-t VIX level. Second, the SPD of VIX conditional on a high Z_t (right panel) has a spike around median volatility levels, consistent with the conventional wisdom that volatility reverts to its long-run mean.

Furthermore, we estimate the pricing kernel $\pi(Z_T|\tau, z_t)$ by combining estimates of both risk-neutral and physical densities of the VIX. The bottom panels of Figure 6 provide nonparametric estimates of $\pi(Z_T|\tau, z_t)$ for a maturity of 42 days and two different levels of Z_t at 18.00 and 25.15. We observe that the pricing kernel exhibits a pronounced U-shape as a function of future VIX levels. Therefore, volatility risk is priced, and the price of volatility risk increases when volatility deviates from its median level. In other words, investors attach high marginal utility to payoffs received when the future volatility is either extremely high or low. Bakshi et al. (2010) document the U-shape for the volatility pricing kernel indirectly, by exploring the link between the monotonicity of the pricing kernel and returns of VIX option portfolios. They further provide a model with heterogeneity in beliefs to account for

Figure 6: State-Price Densities and Pricing Kernels of the VIX



Note: The top panels provide the nonparametric estimates of SPDs of the VIX at various time-to-maturities, with volatility level at 18.00 (left) and 25.15 (right) that correspond to the 25% and 75% quantiles of the VIX time series in our sample, respectively. The bottom panels plot the nonparametric estimates of VIX pricing kernels (black, solid) for the maturity of 42 days, and two current VIX levels at 18.00 and 25.15. Dotted lines are the the 95% confidence intervals. The interest rate and dividend are fixed at their averages, 2.15% and 2.06%, respectively.

the U-shape, in which the volatility market is dominated by investors with zero market risk. In contrast, we provide direct estimates of the volatility pricing kernel by nonparametric methods, which provide more robust information about the shape. In particular, we find that the volatility pricing kernel is asymmetric, and the asymmetry conditional on a high time-t volatility is much stronger than that conditional on a low volatility. This finding implies that investors price the volatility risk differently according to different scenarios of the economy, which presents new empirical regularities that need to be incorporated into models of volatility risk.

In summary, our SPD estimates of VIX document empirical features of risk-neutral dynamics of volatility such as positive skewness and mean reversion. Although the volatility process under the physical measure is well documented as displaying a mean-reverting pattern using historical time series, its risk-neutral behavior is not crystal clear. Our findings uncover the risk-neutral dynamics of volatility without any parametric restrictions. More importantly, our estimates of the volatility pricing kernel show that investors have high marginal utility even in low volatility states, which supports the model with heterogeneity in beliefs.

4.4 In-Sample Fitting and Out-of-Sample Forecasts

We evaluate the performance of our nonparametric estimator (SX) by comparing in-sample fitting and out-of-sample forecasts with two alternative methods discussed in Aït-Sahalia and Lo (1998): the nonparametric approach without volatility factor (AL) in terms of both density and option implied volatility forecasts, and the martingale approach (MKT) for option implied volatility forecasts only. As it is widely used by practitioners, the MKT approach simply forecasts tomorrow's implied volatility by interpolation using today's implied volatility surface.

Intuitively, a potential advantage of our estimator over the MKT method lies in the inclusion of historical options with similar characteristics. As opposed to the MKT approach that relies exclusively on the cross section of options on the previous day, the SX estimator is able to capture a more stable pricing function over time, and hence is expected to outperform the MKT approach in out-of-sample forecasting, although not surprisingly, the SX estimator may fit the cross section of option prices worse on certain days, but better on other days. With historical option prices incorporated, the AL estimator is also capable of capturing certain stability in the historical data, which helps make predictions. However, it misses an important volatility factor that is incorporated into the SX approach.

Panel A of Table 3 reports the forecasting performance of the SX, AL, and MKT methods for option prices (quoted in implied volatility). For each date t, we adopt a preceding 16-month window, within which the SX, AL, and MKT estimators for the target options are obtained. The selected target options have a maturity of 42 days with moneyness ranging between -0.15 and 0.15. We forecast such options on day $t+\gamma$, for $\gamma = 0$ (in-sample), and γ =7, 14, 21, 28, 35, 42, 63 and 84 days (out-of-sample) progressively. We repeat the procedure for each day t in the last 8 months of our sample period and average across days to obtain the root-mean-squared percentage difference between the predictions and the realized option prices.

We observe first that the MKT approach outperforms the AL approach uniformly in forecasting option prices for the sample period we consider, which is in contrast with findings of Aït-Sahalia and Lo (1998). However, this is not surprising as the AL estimator does not include volatility as a conditioning variable which changed substantially over the sample period we consider, i.e., June 1, 2009 – May 31, 2011. In contrast, the SX estimator outperforms both the AL and MKT methods especially for longer horizons. The superior performance of the SX estimator highlights the benefit of predicting by capturing certain stable price patterns in the historical data and incorporating the volatility factor. Not surprisingly, the MKT approach has a better in-sample performance given its implementation.

Table 3: In-Sample Fitting and Out-of-Sample Forecasts

Panel A: Implied Volatility Forecast Error (%)

γ	0	7	14	21	28	35	42	63	84
SX	17.32	14.21	16.68	15.82	17.07	16.92	18.88	17.56	17.56
AL	28.25	28.93	32.10	32.14	33.27	32.52	34.15	35.55	36.61
MKT	13.74	15.94	16.18	15.77	17.26	18.22	18.80	20.63	22.30

Panel B: Density Forecast Error (%)

γ	0	7	14	21	28	35	42	63	84
SX	0.00	0.28	0.52	0.73	0.96	1.15	1.34	1.76	2.20
AL	5.00	5.05	5.22	5.47	5.69	6.07	6.25	6.75	7.35

Note: Panel A reports average forecast errors of implied volatility produced by the Aït-Sahalia and Lo (1998) estimator (AL), our estimator conditional on VIX (SX), and a martingale interpolation (MKT) method, while Panel B reports those of risk-neutral densities using the AL and SX methods. The nonparametric option implied volatility and their corresponding SPDs are estimated with a rolling-window of 16 months, and out-of-sample forecasts are generated for various forecast horizons γ on a daily rolling basis from June 1 2009 to May 31, 2011. The time-to-maturity for both SPDs and option prices is chosen as 42 days.

Panel B of Table 3 reports the forecasting performance of the SX and AL estimators for state-price densities. The empirical design is similar to the forecasting exercise of option implied volatility, with a 16-month window, a target maturity of 42 days, and horizons of τ =7, 14, 21, 28, 35, 42, 63, and 84 days for the out-of-sample performance. We compute the average forecast error (root-mean-squared percentage difference) as a percentage of the mode value of the realized density over the last 8 months of our sample period. The realized density is computed by the SX approach using a 16-month window including the target day. Results in Panel B show that the SX estimator outperforms the AL density substantially over all horizons, due to the missing volatility factor in AL densities. For example, the forecast error for $\tau = 84$ is 2.2% and 7.4% for the SX and AL density estimators, respectively.

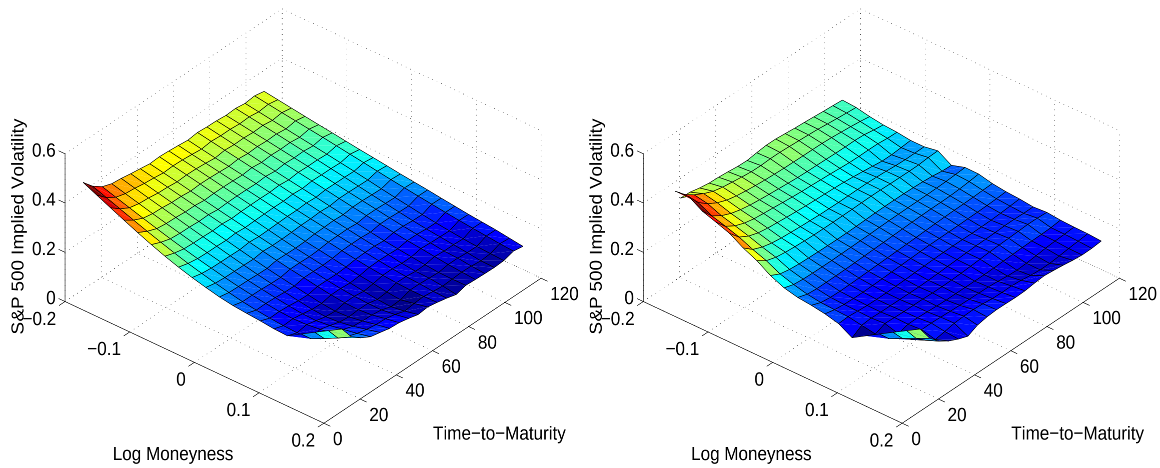
4.5 Robustness Checks

As robustness checks, we verify the two dimension-reduction assumptions employed in our nonparametric procedure: the homogeneity of degree one for S&P 500 options, and the conditional independence of state-price densities of the VIX with respect to S_t.

Figure 7 plots nonparametric estimators of the implied volatility surface of S&P 500 options across both log-moneyness and time-to-maturities: one with the assumption of homogeneity of degree one (left panel) and the other without using it (right panel). We observe that the shape of the two surfaces match each other well in general, although there are slight differences around the boundaries where nonparametric estimators usually incur relatively large biases. Moreover, the estimator without dimension reduction is noiseier as its convergence rate is lower due to the "curse of dimensionality."

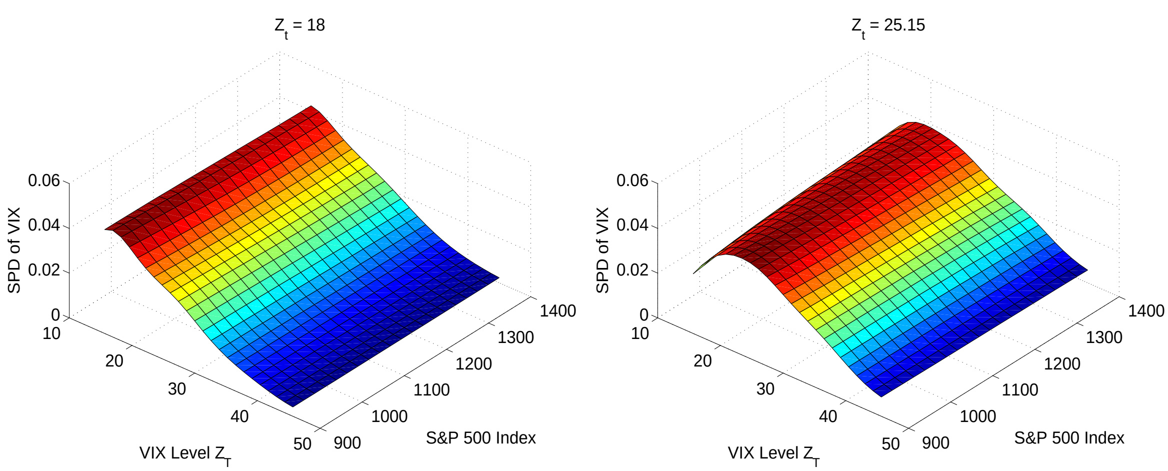
Figure 8 plots estimates of VIX SPDs against the S&P 500 index S_t and VIX Z_T for $\tau = 42$, and for $Z_t = 18.00$ and 25.15 respectively. We observe that conditional densities do not vary much with S_t conditional on either the low or high level of Z_t, especially for the part away from the boundary. Overall, the dimension-reduction assumption for VIX options, i.e.,

Figure 7: Robustness Check I



Note: This figure plots the nonparametric estimates for the implied volatility surface of S&P 500 option prices. The left panel plots the estimates based on dimension reduction techniques, whereas the right panel plots the estimates without such techniques.

Figure 8: Robustness Check II



Note: This figure plots the nonparametric estimates of VIX state-price densities with both Z_t and S_t as conditioning variables. The time-to-maturity is $\tau = 42$, and Z_t is fixed at 18.00 and 25.15 respectively.

the dependence of VIX SPD on S_t mainly through Z_t, seems valid for the sample period we consider.

5 Conclusion

Volatility has been well documented as a priced risk factor, and hence an essential component of pricing kernels. Taking advantage of the rapidly developed volatility derivative markets, we provide nonparametric estimates of both SPDs and pricing kernels with volatility. We show that volatility is the key but missing state variable in the unconditional pricing kernel estimates that exhibit the puzzling U-shape. Moreover, we document a U-shaped pricing kernel of volatility, which cannot be captured by standard models with volatility risk, such as Bollerslev et al. (2009) and Drechsler and Yaron (2011). Therefore, it remains important to develop extensions of these models that are in compliance with our empirical findings.

In addition, our framework extends the nonparametric option pricing method to allow for stochastic volatility, by exploring additional information from the VIX. Existing parametric stochastic volatility models face an unfortunate compromise between model flexibility and tractability. In contrast, our method enjoys several advantages, such as being model-free, robust to model misspecification and pricing measures, and computationally efficient. Hence, our nonparametric option pricing approach with VIX alleviates the compromise to a great extent.

References

Aït-Sahalia, Y. and Duarte, J. (2003), “Nonparametric Option Pricing Under Shape Restrictions,” *Journal of Econometrics*, 116, 9–47.

Aït-Sahalia, Y., Fan, J., and Peng, H. (2009), “Nonparametric Transition-Based Tests for Jump-Diffusions,” *Journal of the American Statistical Association*, 104, 1102–1116.

Aït-Sahalia, Y. and Lo, A. (1998), “Nonparametric Estimation of State-Price-Densities Implicit in Financial Asset Prices,” *Journal of Finance*, 53, 499–547.

— (2000), “Nonparametric Risk Management and Implied Risk Aversion,” *Journal of Econometrics*, 94, 9–51.

Amengual, D. and Xiu, D. (2012), “Delving into Risk Premia: Reconciling Evidence from the S&P 500 and VIX Derivatives,” Tech. rep., CEMFI and University of Chicago Booth School of Business.

Bakshi, G., Cao, C., and Chen, Z. (1997), “Empirical Performance of Alternative Option Pricing Models,” *Journal of Finance*, 52, 2003–2049.

Bakshi, G. and Kapadia, N. (2003), “Delta-Hedged Gains and the Negative Market Volatility Risk Premium,” *Review of Financial Studies*, 16, 527–566.

Bakshi, G. and Madan, D. (2008), “Investor Heterogeneity and the Non-Monotonicity of the Aggregate Marginal Rate of Substitution in the Market Index,” *working paper, University of Maryland*.

Bakshi, G., Madan, D., and Panayotov, G. (2010), “Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels,” *Journal of Financial Economics*, 97, 130–154.

Bansal, R., Kiku, D., Shaliastovich, I., and Yaron, A. (2012), “Volatility, the Macroeconomy, and Asset Prices,” Tech. rep., University of Pennsylvania.

Bansal, R. and Yaron, A. (2004), “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles.” *Journal of Finance*, 59.

Barndorff-Nielsen, O. E. and Shephard, N. (2001), “Non-Gaussian Ornstein-Uhlenbeck-Based Models And Some Of Their Uses In Financial Economics,” *Journal of the Royal Statistical Society, B*, 63, 167–241.

Bates, D. S. (2000), “Post-’87 Crash Fears in the S&P 500 Futures Option Market,” *Journal of Econometrics*, 94, 181–238.

— (2012), “U.S. Stock Market Crash Risk, 1926-2010.” *Journal of Financial Economics*, 105, 229–259.

Boes, M., Drost, F., and Werker, B. J. (2007), “Nonparametric Risk-Neutral Return and Volatility Distributions,” Tech. rep., Tilburg University.

Bollerslev, T., Sizova, N., and Tauchen, G. (2012), “Volatility in Equilibrium: Asymmetries and Dynamic Dependencies,” *Review of Finance*, 16, 31–80.

Bollerslev, T., Tauchen, G. E., and Zhou, H. (2009), “Expected Stock Returns and Variance Risk Premia,” *Review of Financial Studies*, 22, 4463–4492.

Breeden, D. and Litzenberger, R. H. (1978), “Prices of State-Contingent Claims Implicit in Option Prices,” *Journal of Business*, 51, 621–651.

Britten-Jones, M. and Neuberger, A. (2000), “Option Prices, Implied Price Processes, and Stochastic Volatility,” *Journal of Finance*, 55, 839–866.

Broadie, M., Chernov, M., and Johannes, M. S. (2007), “Model Specification and Risk Premia: Evidence from Futures Options,” *Journal of Finance*, 62.

Campbell, J., Christopher, P., Turley, B., and Giglio, S. (2012), “An Intertemporal CAPM with Stochastic Volatility,” Tech. rep., Harvard University.

Campbell, J. Y. and Shiller, R. J. (1988), "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661–676.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2003), "Stochastic Volatility for Lévy Processes," *Mathematical Finance*, 13, 345–342.

Carr, P. and Wu, L. (2009), "Variance Risk Premiums," *Review of Financial Studies*, 22, 1311–1341.

Chabi-Yo, F. (2011), "Pricing Kernels with Stochastic Skewness and Volatility Risk," *Management Science*.

Chabi-Yo, F., Garcia, R., and Renault, E. (2008), "State Dependence can Explain the Risk Aversion Puzzle," *Review of Financial Studies*, 21, 973–1011.

Chernov, M. and Ghysels, E. (2000), "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 57, 407–458.

Christoffersen, P., Heston, S., and Jacobs, K. (2010), "Option Anomalies and the Pricing Kernel," Tech. rep., McGill University.

Christoffersen, P., Jacobs, K., Ornthanalai, C., and Wang, Y. (2008), "Option Valuation with Long-Run and Short-Run Volatility Components," *Journal of Financial Economics*, 90, 272–297.

Drechsler, I. and Yaron, A. (2011), "What's Vol Got to Do with It?" *Review of Financial Studies*, 24, 1–45.

Egloff, D., Leippold, M., , and Wu, L. (2010), "The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments," *Journal of Financial and Quantitative Analysis*, 45, 1279–1310.

Epstein, L. and Zin, S. (1989), "Substitution, Risk aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Eraker, B. (2004), "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," *Journal of Finance*, 59.

Eraker, B., Johannes, M. S., and Polson, N. (2003), "The Impact of Jumps in Equity Index Volatility and Returns," *Journal of Finance*, 58, 1269–1300.

Fan, J. and Gijbels, I. (1996), *Local Polynomial Modelling and Its Applications*, London, U.K.: Chapman & Hall.

Fan, J. and Mancini, L. (2009), "Option Pricing with Aggregation of Physical Models and Nonparametric Statistical Learning," *Journal of American Statistical Association*, 104, 1351–1372.

Fan, J., Yao, Q., and Tong, H. (1996), "Estimation of Conditional Densities and Sensitivity Measures in Nonlinear Dynamical Systems," *Biometrika*, 83, 189–206.

Fan, J. and Yim, T. H. (2004), "A Crossvalidation Method for Estimating Conditional Densities," *Biometrika*, 91, 819–834.

Jackwerth, J. (2000), "Recovering Risk Aversion from Option Prices and Realized Returns," *Review of Financial Studies*, 13, 433–451.

Jackwerth, J. and Brown, P. (2001), "The Pricing Kernel Puzzle: Reconciling Index Option Data and Economic Theory," Tech. rep., University of Konstanz.

Jackwerth, J. and Vilkov, G. (2013), "Asymmetric Volatility Risk: Evidence from Option Markets," *working paper*.

Joshi, M. (2007), "Log-type models, Homogeneity of Option Prices and Convexity," Tech. rep., Melbourne University.

Li, H. and Zhao, F. (2009), "Nonparametric Estimation of State-Price-Densities Implicit in Interest Rate Cap Prices," *Review of Financial Studies*, 22, 4335–4376.

Mencia, J. and Sentana, E. (2012), "Valuation of VIX Derivatives," *Journal of Financial Economics, forthcoming.*

Pan, J. (2002), "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3–50.

Rosenberg, J. V. and Engle, R. F. (2002), "Empirical Pricing Kernels," *Journal of Financial Economics*, 64, 341–372.

Ruppert, D., Sheather, S., and Wand, M. P. (1995), "An Effective Bandwidth Selector for Local Least Squares Kernel Regression," *Journal of American Statistical Association*, 90.

Ruppert, D. and Wand, M. (1994), "Multivariate Locally Weighted Least Squares Regression," *Annals of Statistics*, 22, 1346–1370.

Shephard, N. (2005), *Stochastic Volatility*, Oxford University Press.

Tauchen, G. E. and Todorov, V. (2011), "Volatility Jumps," *Journal of Business and Economic Statistics*, 29.

Todorov, V. (2010), "Variance Risk Premium Dynamics: The Role of Jumps," *Review of Financial Studies*, 23, 345–383.

Yao, Q. and Hall, P. (2005), "Estimation for Conditional Distribution Functions via Dimension Reduction," *Annals of Statistics*, 33, 1404–1421.

Yao, Q. and Tong, H. (1998), "Cross-Validatory Bandwidth Selections for Regression Estimation Based on Dependent Data," *Journal of Statistical Planning and Inference*, 68, 387–415.

Ziegler, A. (2007), "Why does Implied Risk Aversion Smile?" *Review of Financial Studies*, 20, 859–904.