An empirical test of pricing kernel monotonicity

Brendan K. Beare and Lawrence Schmidt*

Department of Economics, University of California, San Diego

July 21, 2011

Abstract

A recent literature in finance concerns a curious recurring feature of estimated pricing kernels. Classical theory dictates that the pricing kernel – defined loosely as the ratio of Arrow security prices to an objective probability measure – should be a decreasing function of aggregate resources. Yet a large number of recent empirical studies appear to contradict this prediction. The nonmonotonicity of empirical pricing kernel estimates has become known as the pricing kernel puzzle. In this paper we propose and apply a formal statistical test of pricing kernel monotonicity. The test involves assessing the concavity of the ordinal dominance curve associated with the risk neutral and physical return distributions. We apply the test using thirteen years of data from the market for European put and call options written on the S&P 500 index. Statistically significant violations of pricing kernel monotonicity occur in a substantial proportion of months.

*We thank seminar participants at UC Riverside, UC San Diego and the University of Chicago for helpful comments. We also thank Peter Hansen for providing the realized volatility data and related advice, Ana Monteiro for providing MatLab code for implementing the risk neutral density estimation procedure, and Allan Timmermann for helpful discussions.

1 Introduction

A recent literature in finance concerns a curious recurring feature of estimated pricing kernels. Classical theory dictates that the pricing kernel – defined loosely as the ratio of Arrow security prices to an objective probability measure – should be a decreasing function of aggregate resources. Yet a large number of recent empirical studies appear to contradict this prediction. The nonmonotonicity of empirical pricing kernel estimates has become known as the pricing kernel puzzle.

Jackwerth (2000) is the seminal paper in the pricing kernel puzzle literature. Jackwerth constructed nonparametric estimates of the physical and risk neutral distributions of monthly returns on the S&P 500 index using historical return data and option prices. He found that the pricing kernel – the ratio of the risk neutral and physical distributions – appeared to be monotone decreasing immediately prior to the 1987 stock market crash, but nonmonotone afterwards. In particular, the pricing kernel appeared to be increasing in the middle of the return distribution, and decreasing elsewhere. Other early studies identifying similar behavior in pricing kernels include Aït-Sahalia and Lo (2000) and Rosenberg and Engle (2002).

The bulk of the literature on the pricing kernel puzzle has been empirical. A recent working paper by Hens and Reichlin (2010) provides a useful discussion of the puzzle from a theoretical perspective. Hens and Reichlin identify three core assumptions under which the pricing kernel should be a decreasing function of aggregate resources. These are (1) risk averse investors, (2) complete markets, and (3) common and true beliefs. The violation of any of these conditions may lead to pricing kernel nonmonotonicity. As discussed by Constantinides et al. (2009), Beare (2011) and others, a nonmonotone pricing kernel may imply the existence of derivative securities that stochastically dominate the market portfolio. The subject is therefore of considerable interest from a practical perspective.

A crude theoretical explanation for pricing kernel monotonicity runs as follows. Consider a one period model with uncertainty in which a representative agent chooses a portfolio of Arrow securities to maximize the expected utility derived from the consumption of a single good. Normalizing the state space to the unit interval, and assuming an increasing, concave and differentiable utility function u, the agent's decision problem is to choose a consumption profile c that maximizes

$$E(u(c)) = \int_0^1 u(c(x))g(x)\mathrm{d}x$$

subject to the budget constraint

$$\int_0^1 c(x)f(x)\mathrm{d}x = w.$$

Here, w is the agent's wealth, f assigns Arrow prices to securities yielding a payoff of one unit of consumption in a given state, and g is the probability density function (pdf) over the state space. The first-order condition for a maximum in this problem is

$$u'(c(x))g(x) = \lambda f(x),$$

with λ a nonnegative Lagrange multiplier. Rearranging terms, we find that the pricing kernel, defined as the ratio $f(x)/g(x)$, is proportional to $u'(c(x))$, the marginal utility from consumption. Concavity of u therefore implies that the pricing kernel should be monotone decreasing in $c(x)$. If $c(x)$ is interpreted as the aggregate resource level in state x, we find that the pricing kernel should be a decreasing function of aggregate resources. This observation may be traced back to a fundamental contribution of Dybvig (1988).

Hens and Reichlin (2010) consider the extent to which pricing kernel monotonicity is preserved in more general models of price determination. They find that the assumption of prices being determined by a representative agent may be relaxed without affecting the monotonicity property. Risk seeking behavior, incomplete markets, and heterogeneous beliefs may all lead to violations of monotonicity. Hens and Reichlin dismiss the first of these explanations, noting that it depends critically on the state space being atomic, and argue that the latter two explanations are more plausible. Ziegler (2007) previously explored the possibility of heterogenous beliefs generating nonmonotone pricing kernels. Detlefsen et al. (2007), Chabi-Yo et al. (2008) and Härdle et al. (2009) have considered state-dependent preferences as an explanation for pricing kernel nonmonotonicity.

Recent empirical studies of pricing kernel monotonicity have led to mixed results. Bakshi et al. (2010) find that average returns on call options written on the S&P 500 are decreasing in the strike price. This suggests that the pricing kernel may in fact be U-shaped. Constantinides et al. (2009) develop a discretized model of monthly S&P 500 returns in which the existence of a monotone decreasing pricing kernel implies the solubility of a system of linear equations based on the return distribution and observed prices of stocks, bonds and options. Using a number of estimated models of market returns, they find that the system of equations is insoluble in a high proportion of months over a twenty year period. Barone-Adesi et al. (2008) and Barone-Adesi and Dall'O (2010) use an asymmetric GARCH model to obtain pricing kernel estimates for the S&P 500 index at a variety of dates and return horizons. The overall shape of their estimates is generally decreasing, though many of them are locally increasing over certain regions. No formal test of monotonicity is undertaken.

Perhaps the two papers that are closest in spirit to the present paper are Golubev et al. (2008) and Härdle et al. (2010). In both of these papers, an attempt is made to conduct formal statistical inference about the shape of the pricing kernel. Golubev et al. (2008) set up a likelihood ratio test in which the pricing kernel is nonincreasing under the null. They assume that the risk neutral distribution is known, while allowing the physical distribution to vary as a parameter. Applying their test to the German DAX in the years 2000, 2002 and 2004, they reject the null hypothesis of monotonicity in 2002 at the 5% significance level and in 2000 and 2002 at the 10% significance level. Härdle et al. (2010) provide a method for constructing uniform confidence bands for the pricing kernel, allowing for uncertainty about both the risk neutral and physical distributions. They report 95% confidence bands for the pricing kernel for DAX returns in 2006 at a variety of time horizons and specific dates. In four of the six cases considered, one can see that a decreasing pricing kernel is consistent with the reported confidence bands. In the other two cases, the confidence bands imply that the pricing kernel must be increasing in a range of negative returns.

In this paper we propose and apply a new test of pricing kernel kernel monotonicity. Our procedure is an extension of a method proposed by Carolan and Tebbs (2005) for testing the monotonicity of a density ratio when we observe two independent samples drawn from the two distributions under consideration. Carolan and Tebbs' approach relies on the insight that a ratio of densities is monotone if and only if the corresponding ordinal dominance curve – defined in the following section – is concave. Suitable statistics for testing the null hypothesis of a monotone density ratio may therefore be constructed from a measure of the distance between the estimated ordinal dominance curve and its least concave majorant. We generalize the procedure of Carolan and Tebbs so that it may be applied when one or both distributions is estimated in a nonstandard fashion, as must be the case for the risk neutral distribution in the present context. In fact, we estimate the risk neutral density using a method proposed recently by Monteiro et al. (2008), which involves approximating the density using a smooth cubic spline. Under appropriate regularity conditions, the null asymptotic distributions of our test statistics are shown to be continuous functionals of a suitably defined continuous Gaussian process.

We apply our testing procedure using data on the prices of options written on the S&P 500 index between 1997 and 2008. Test statistics were calculated at a total of 128 dates during this period. In each case the return horizon considered was approximately 20 trading days. Using a test statistic based on the area between the estimated ordinal dominance curve and its least concave majorant, we reject the null hypothesis of monotonicity 16% of the time at the 5% significance level. Using another test statistic based on the maximum vertical distance between the two curves, the rejection rate rises to 30% at the 5% significance level. More detailed information about precisely when our tests detect violations of monotonicity, at a range of significance levels, may be found in Section 4.3.

A key difference between the results reported here, and previous empirical studies of pricing kernel monotonicity, is the way in which we deal with time variation in the volatility of the physical distribution. Golubev et al. (2008) and Härdle et al. (2010) do not account for time variation in volatility, instead treating the historical market returns as independent draws from the time invariant physical distribution. This changes the interpretation of their results – we must interpret the pricing kernel as the ratio of the risk neutral density, and the *unconditional* physical return density. In fact, it is the ratio of the risk neutral density to the physical density *conditional on current information* which classical theory posits to be monotone. In particular, we should condition on the present level of volatility. To this end, we employ the Realized GARCH model of Hansen et al. (2011), which combines the forward-looking GARCH structure with ex-post volatility measurements obtained from high-frequency intraday data using the realized kernel method developed by

Barndorff-Nielsen et al. (2008, 2009). We deflate excess market returns using the realized volatilities before estimating the physical return distribution, and then rescale the estimated physical distribution to the present level of volatility before calculating the pricing kernel.

The remainder of the paper is structured as follows. In Section 2 we introduce the statistical framework of our testing procedure, showing how the approach of Carolan and Tebbs (2005) may be applied when nonstandard estimators of the two distributions are employed. The null asymptotic distributions of our test statistics are obtained under high level conditions on the two estimated distribution functions. In Section 3 we explain precisely how we estimate the risk neutral and physical distributions in our application, and verify that these estimators satisfy the high level conditions given in Section 2. We also explain how critical values for our tests are calculated. Section 4 contains our empirical results. Detailed results are provided for the most recent date at which our tests are implemented, and a summary of the results obtained at all 128 dates in the sample period. In Section 5 we give some final thoughts and conclude. Appendix A.1 contains regularity conditions under which our estimator of the risk neutral density is well behaved, and Appendix A.2 contains proofs of mathematical results stated in the main part of the paper.

2 Statistical framework

The pricing kernel puzzle concerns the shape of the ratio of the risk neutral and physical densities governing the payoff of some base asset – typically, a market index – at a given future date. Let F and G be two cumulative distribution functions (cdfs) on the real line with $F(x) = G(x) = 0$ for all $x \leq 0$. The cdf G, referred to as the *physical distribution*, describes the value after one period of a \$1 investment in the base asset. It should be interpreted as being conditional on all information available at the time of investment. The cdf F, referred to as the *risk neutral distribution*, determines the price of derivative contracts delivering a payoff after one period that is determined by the value of the base asset at that time. Such contracts have price equal to their discounted expected payoff under F. For instance, a European call option written on the base asset at strike s, expiring after one period, will have price equal to $(1+r)^{-1}\int \max\{x-s,0\}\mathrm{d}F(x)$, with r the risk-free interest rate.

The following condition ensures that F and G admit suitably well behaved probability density functions (pdfs) f and g.

Assumption 2.1. The following statements are true.

(a) F and G have continuous derivatives f and g on $\mathbb{R}$.

(b) g is strictly positive on some interval (a,b) with $0 \leq a < b \leq \infty$, and zero elsewhere.

(c) $f(x) = 0$ for all x such that $g(x) = 0$.

Under Assumption 2.1 we may define the ratio $\pi(x) = f(x)/g(x)$ for each $x \in (a,b)$. The function $\pi : (a,b) \to [0,\infty)$ is referred to as the *pricing kernel*.

We wish to test whether π is nonincreasing over its domain, (a,b). Patton and Timmermann (2010) discuss a variety of ways to set up the null and alternative hypotheses in tests of monotonicity. We shall adopt the following formulation.

$$\mathrm{H}_0 : \pi \text{ is nonincreasing;} \quad \mathrm{H}_1 : \pi \text{ is not nonincreasing.}$$

The null hypothesis H_0 is composite, meaning that it admits multiple pricing kernels π. Consistent with Carolan and Tebbs (2005) and Golubev et al. (2008), we shall choose critical values for our test statistics that deliver the correct asymptotic size at a particular choice of nonincreasing pricing kernel: $\pi = 1$. This is the only choice of π that is constant, since f and g must integrate to one. When π is constant it is, in an intuitive sense, as close to violating H_0 as possible. Suitably constructed tests that deliver the correct size when $\pi = 1$ may be conservative for other choices of π in H_0.

The approach to monotonicity testing proposed by Carolan and Tebbs (2005) is based on an equivalence between the monotonicity of π and the concavity of a function called the *ordinal dominance curve*, or odc. Given our cdfs F and G, the corresponding odc is the map $\phi : [0,1] \to [0,1]$ given by

$$\phi(u) = F\left(G^{-1}(u)\right), \; u \in [0,1].$$

Here, G^{-1} is the quantile function for G, given by the usual expression

$$G^{-1}(u) = \inf\{x : G(x) \geq u\},\ u \in (0, 1],$$

or $G^{-1}(0) = \lim_{u \downarrow 0} G^{-1}(u)$. Under Assumption 2.1, ϕ is continuous and nondecreasing on $[0, 1]$ with $\phi(0) = 0$ and $\phi(1) = 1$. It is known (see e.g. Hseih and Turnbull, 1996) that the density ratio π is nonincreasing if and only if the odc ϕ is concave. In fact, the second derivative of ϕ and first derivative of π are of the same sign at any point where they both exist. One may therefore consider testing H_0 against H_1 by constructing a test statistic that is in some sense an empirical measure of the apparent nonconcavity of ϕ.

In Figure 2.1 we provide three examples of pairs of pdfs, their ratios, and the corresponding odcs. It may be helpful to think of the green pdfs as risk neutral densities and the blue pdfs as physical densities. In the first row, the two pdfs are normal with different means and the same variance, with the physical density shifted to the right of the risk neutral density. In this case, the pricing kernel is monotone decreasing, and the odc is concave. In the second row the two pdfs are normal with equal means, but the variance of the risk neutral density is greater than the variance of the physical density. In this case the pricing kernel is U-shaped, broadly consistent with the empirical findings of Bakshi et al. (2010), and the odc is not concave. In the third row the two pdfs are the risk neutral and physical pdfs estimated by Jackwerth (2000) for monthly S&P 500 returns on April 15, 1992. In this case the pricing kernel is decreasing at the extremes but nondecreasing around the center of the return distribution, and the odc again fails to be concave.

In our empirical application we shall estimate F using a cross-section of current option prices, and G using a time series of excess returns and realized volatilities. Let m denote the number of observed option prices, and let n denote the length of the time series. It will be convenient to treat m as a function of n, so that implicitly $m = m(n)$, and for the purposes of obtaining asymptotic approximations we shall assume that $m \to \infty$ as $n \to \infty$, with $n/m \to \lambda$ for some $\lambda \in [0, \infty)$. We shall index quantities depending on either sample size by n, not m. Let F_n and G_n denote our estimates of F and G. Carolan and Tebbs (2005) take F_n and G_n to be the empirical distribution functions (edfs) of independent and identically distributed (iid) samples of size m and n drawn from F and G. Here we allow for more general estimators. Details about the specific estimators used

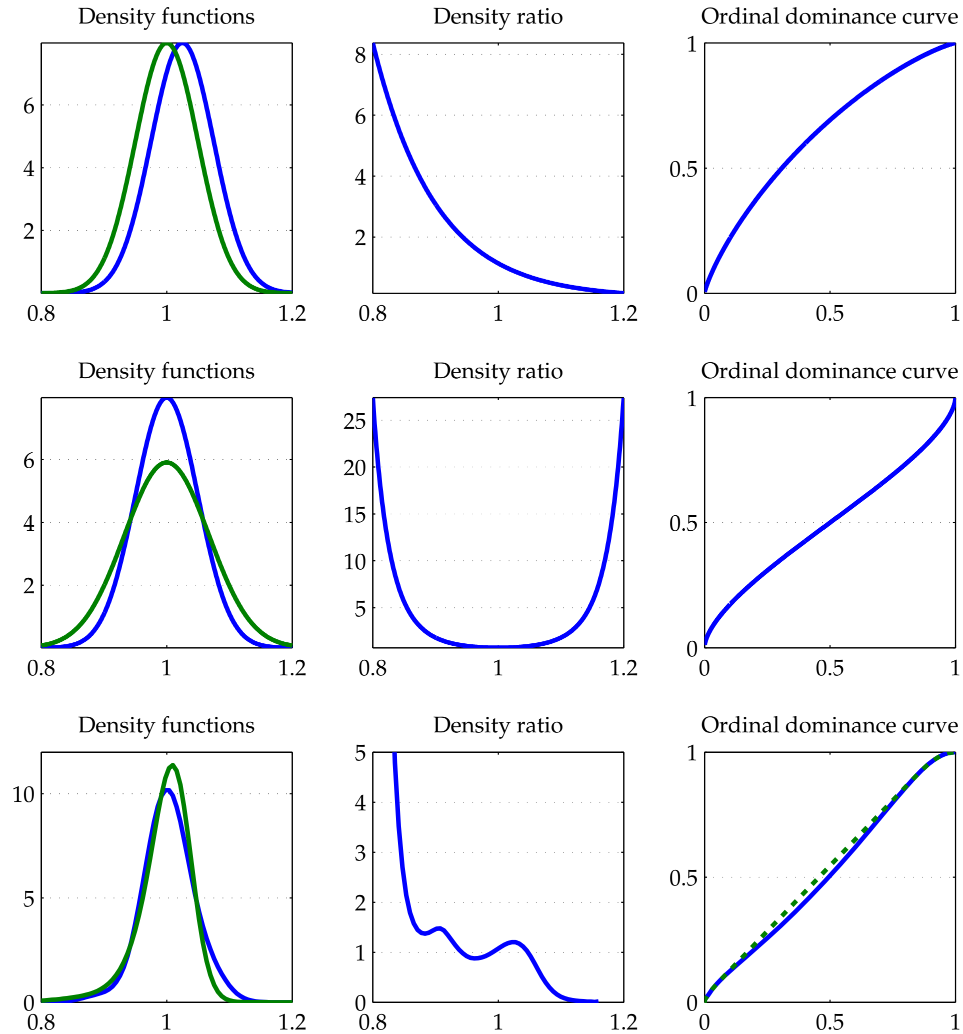

Figure 2.1: Density ratios and ordinal dominance curves.

in our application are provided in Section 3.

From F_n and G_n we may construct the estimated odc $\phi_n(u) = F_n(G_n^{-1}(u))$. The *least concave majorant*, or lcm of ϕ_n, denoted $\mathcal{M}\phi_n$, is the pointwise infimum of all concave functions on $[0,1]$ that lie above ϕ_n. Figure 2.2 provides an example of an estimated odc and its lcm. The estimated odc, in blue, was constructed using the edfs of two iid samples from the standard normal distribution, each with 50 observations. The green line is the lcm of the estimated odc.

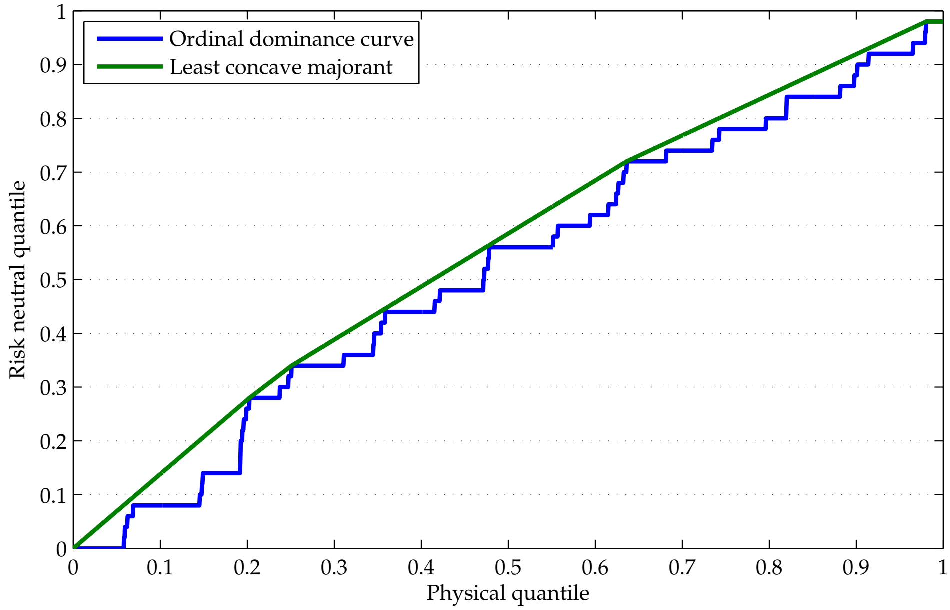

Figure 2.2: The least concave majorant of an estimated ordinal dominance curve.

The difference between ϕ_n and its lcm is a nonnegative function on $[0, 1]$, and shall be denoted $\mathcal{D}\phi_n = \mathcal{M}\phi_n - \phi_n$. Carolan and Tebbs (2005) propose using the following statistics T_n and T_n' for testing H_0 against H_1:

$$T_n = \sqrt{n} \int_0^1 \mathcal{D}\phi_n(u)\mathrm{d}u \tag{2.1}$$

$$T_n' = \sqrt{n} \sup_{u \in [0,1]} \mathcal{D}\phi_n(u). \tag{2.2}$$

In Figure 2.1, T_n is $\sqrt{n}$ times the area between the blue and green lines, while T_n' is $\sqrt{n}$ times the maximum vertical distance between the blue and green lines.

Before we can say anything about the statistical properties of T_n and T_n', we must introduce conditions on the behavior of F_n and G_n. We shall impose the following high level condition. The symbol $\rightsquigarrow$ denotes weak convergence (see e.g. Definition 1.3.3 in van der Vaart and Wellner, 1996), while $\circ$ denotes composition. Given an arbitrary set D, $\ell^\infty(D)$ denotes the space of uniformly bounded real valued functions on D equipped with the topology of uniform convergence.

Assumption 2.2. For each $n \in \mathbb{N}$, F_n and G_n are random cdfs on $\mathbb{R}$. As $n \to \infty$ they satisfy

$$\begin{pmatrix} \sqrt{n}(F_n - F) \\ \sqrt{n}(G_n - G) \end{pmatrix} \rightsquigarrow \begin{pmatrix} \xi \circ F \\ \zeta \circ G \end{pmatrix}$$

in the product topology on $\ell^\infty(\mathbb{R})^2$, where ξ and ζ are independent continuous random elements of $\ell^\infty[0,1]$.

If F_n and G_n are the edfs of independent iid samples of size m and n drawn from F and G, as they are in the case considered by Carolan and Tebbs (2005), then Assumption 2.2 is a simple consequence of Donsker's theorem. In this case we have $\xi = \lambda^{1/2} B_1$ and $\zeta = B_2$, where B_1 and B_2 are independent standard Brownian bridges on $[0,1]$. Conditions under which Assumption 2.2 is satisfied by the estimators used in our application are provided in Section 3.

Under Assumptions 2.1 and 2.2, we are able to establish the following result, Theorem 2.1. Part (a) of Theorem 2.1 provides the limiting distributions of T_n and T_n' when π is constant, which we treat as the "least favorable point" in H_0. Part (b) of Theorem 2.1 shows that tests formed by comparing T_n and T_n' to critical values taken from the limiting distributions in part (a) are consistent against all points in H_1.

Theorem 2.1. *Suppose Assumptions 2.1 and 2.2 are true, and let T_n and T_n' be given as in (2.1) and (2.2) above.*

(*a*) *If π is constant, then as $n \to \infty$ we have*

$$T_n \to_d \int_0^1 \mathcal{D}\mathscr{G}(u)\,du \quad \text{and} \quad T_n' \to_d \sup_{u \in [0,1]} \mathcal{D}\mathscr{G}(u),$$

where $\mathscr{G}(u) = \xi(u) - \zeta(u)$.

(*b*) *If π is not nonincreasing, then for any $c \in \mathbb{R}$ we have $T_n > c$ and $T_n' > c$ with probability approaching one as $n \to \infty$.*

The proof of Theorem 2.1 may be found in Appendix A.2. As noted earlier, in the framework considered by Carolan and Tebbs (2005) we have $\xi = \lambda B_1$ and $\zeta = B_2$. When

π is constant, we therefore have $\mathscr{G} = \lambda^{1/2}B_1 - B_2$. The linear combination of independent Brownian bridges $\lambda^{1/2}B_1 - B_2$ has the same distribution as $(1+\lambda)^{1/2}B$, where B is another standard Brownian bridge. Thus we obtain the null limiting distributions given by Carolan and Tebbs (2005) as a special case of Theorem 2.1. There, the random process $\mathscr{G}$ does not depend on any nuisance parameters other than the limiting sample size ratio λ. This turns out not to be the case in general, and in the application in this paper we will find that $\mathscr{G}$ depends on unknown nuisance parameters. Consequently, the limiting distributions in Theorem 2.1(a) are unknown. We obtain critical values by computing the desired tail quantiles of uniformly consistent estimates of the relevant limiting distributions. This will be discussed in more detail in Section 3.

3 Construction of test statistics and critical values

In Section 2 we outlined the statistical framework we shall use for testing pricing kernel monotonicity. Assumption 2.2 provided high level conditions that our estimators F_n and G_n of the risk neutral and physical distributions must satisfy in order for the limiting distributions given in Theorem 2.1 to be valid. We have yet to explain how F_n and G_n are constructed from data. This will be done in Sections 3.1 and 3.2. We give sufficient conditions under which F_n and G_n satisfy Assumption 2.2. In a Section 3.3, we explain how critical values are calculated for our tests.

3.1 Risk neutral distribution estimation

Nonparametric methods for estimating the risk neutral density governing the prices of state contingent claims were first introduced by Jackwerth and Rubinstein (1996) and Aït-Sahalia and Lo (1998). In this paper we shall apply a method proposed more recently by Monteiro et al. (2008) which involves approximating the risk neutral density with a cubic spline. The data used for estimating f at a single point in time consist of m quadruplets (d_i, s_i, p_i, v_i), $i = 1, \ldots, m$, describing the observed characteristics of European put and call options written on our base asset, expiring after one period. d_i is equal to zero if the ith option is a call and one if the ith option is a put, s_i is the strike at which the ith option is written, p_i is the price of the ith option, and v_i is the trading volume of options with strike s_i and of type d_i on the date in question. Additionally, we assume that we observe the one period risk-free interest rate r that applies at the present date. Further details about how our data are constructed are provided in Section 4.1.

Following Monteiro et al. (2008), we assume that f is a cubic spline (i.e., a smooth piecewise cubic polynomial) with fixed knots $x_0 < \cdots < x_k$. The choice of knots is arbitrary; in our application we employ an *ad hoc* data dependent method to choose the knot locations. Given the knots, f is specified up to a vector of parameters $\theta \in \mathbb{R}^{4k}$:

$$f(x;\theta) = \begin{cases} \theta_{4j-3} + \theta_{4j-2}x + \theta_{4j-1}x^2 + \theta_{4j}x^3 & \text{if } x \in (x_{j-1}, x_j] \\ 0 & \text{if } x \notin (x_0, x_k]\,. \end{cases} \tag{3.1}$$

The parameter vector θ is restricted in such a way that f integrates to one, is twice continuously differentiable on (x_0, x_k), and is continuous and equal to zero at x_0 and x_k. These requirements amount to $3k$ linear equality restrictions on θ, which we write as $R\theta = 0$ using a suitably chosen $3k \times 4k$ matrix R. In addition to these linear equality restrictions, we require that f is nonnegative. This condition imposes a family of linear inequality restrictions on θ. We require

$$\theta_{4j-3} + \theta_{4j-2}x + \theta_{4j-1}x^2 + \theta_{4j}x^3 \geq 0 \tag{3.2}$$

for all $x \in (x_{j-1}, x_j]$ and all $j = 1, \ldots, k$. Let Θ denote the collection of $\theta \in \mathbb{R}^{4k}$ satisfying $R\theta = 0$ and the inequality restrictions in (3.2).

Monteiro et al. (2008) propose choosing $\theta \in \Theta$ to minimize a weighted sum of squared differences between the observed prices p_i and the prices implied by $f(\cdot, \theta)$. The trading volumes v_i are used to weight the different squared pricing errors. This approach is consistent with the idea that highly traded assets are more likely to be accurately priced. The estimator for the "true" value of θ, which we denote θ^*, can be written as

$$\theta_n = \operatorname*{argmin}_{\theta \in \Theta} \sum_{i=1}^{m} v_i \left(p_i - h(d_i, s_i; \theta)\right)^2 ,$$

where

$$h(d, s; \theta) = \frac{1}{1+r} \int_0^\infty \max\{(-1)^d (x - s), 0\} f(x; \theta) \mathrm{d}x.$$

Note that, in view of the spline model (3.1), $h(d, s; \theta)$ is a linear function of θ for each (d, s). For each i, we may therefore write $h(d_i, s_i; \theta) = z_i' \theta$ for all θ, with z_i an element of $\mathbb{R}^{4k}$ determined by s_i, d_i, r, and the spline knots $x_0, \ldots, x_k$. Estimation of θ^* thus amounts to weighted linear regression subject to linear equality and inequality restrictions. Monteiro et al. (2008) provide a fast algorithm to compute θ_n based on the method of semidefinite programming. We shall not repeat the details here.

Given the estimated risk neutral density $f(\cdot; \theta_n)$, our estimate for the risk neutral distribution F_n is simply $F_n(x) = \int_0^x f(y; \theta_n) \mathrm{d}y$. We can deduce a uniform limit theory for $\sqrt{n}(F_n - F)$ consistent with Assumption 2.2 by considering the limiting distribution of the scaled parameter estimation error $\sqrt{n}(\theta_n - \theta^*)$. Monteiro et al. (2008) do not consider the stochastic properties of θ_n. Nevertheless, under suitable regularity conditions, we can obtain the limiting distribution of $\sqrt{n}(\theta_n - \theta^*)$ using standard fixed regressor asymptotics for weighted linear regression subject to linear restrictions. The regularity conditions we employ, though fairly standard, are technical. We relegate them to Appendix A.1, where they are given as Assumption A.1.

Under Assumption A.1, it is straightforward to obtain the following result, which provides the limiting distributions of $\sqrt{n}(\theta_n - \theta^*)$ and $\sqrt{n}(F_n - F)$. Let I denote the $4k \times 4k$ identity matrix, and let $M = I - \Xi^{-1} R' \left(R \Xi^{-1} R'\right)^{-1} R$. The matrices Ξ and Σ are defined in Assumption A.1.

Theorem 3.1. *Suppose Assumptions 2.1 and A.1 are satisfied. Suppose also that $m \to \infty$ as $n \to \infty$, with $n/m \to \lambda \in [0, \infty)$. Let N denote a vector of $4k$ independent standard normal random variables. Then, as $n \to \infty$, the following limiting statements are valid.*

(a) $\sqrt{n}\,(\theta_n - \theta^*) \to_d \Psi^{1/2} N$, *where* $\Psi = \lambda M \Xi^{-1} \Sigma \Xi^{-1} M'$.

(b) $\sqrt{n}\,(F_n - F) \rightsquigarrow \xi \circ F$ *in* $\ell^\infty(\mathbb{R})$, *where* $\xi(u) = \left(\frac{\partial}{\partial \theta} F(F^{-1}(u; \theta^*); \theta) \big|_{\theta = \theta^*} \right)' \Psi^{1/2} N$.

Theorem 3.1(b) gives the form of the limiting process ξ that appears in Assumption 2.2 when F_n is constructed as discussed in this subsection. ξ is a finite dimensional continuous

Gaussian process, having dimension of at most k. Finite dimensionality is a consequence of the parametric nature of the cubic spline model for f.

3.2 Physical distribution estimation

The data we use to estimate the physical distribution G consist of n triples (q_t, r_t, σ_t), $t = 1, \ldots, n$. q_t is the rate of return on an investment in the base asset – in our case, the S&P 500 index – from time $t-1$ to time t. It is calculated as the natural logarithm of the price ratio between the two dates. r_t is the risk-free rate of interest from time $t-1$ to time t. σ_t^2 is a measure of the conditional volatility of excess returns on the base asset from time $t-1$ to time t. It is based on a realized volatility estimate constructed from high frequency intraday return data. Additional details are provided in Section 4.1.

The physical distribution G we seek to estimate describes the distribution of the dollar value at time τ of an investment of \$1 in the market index at time $\tau - 1$, conditional on r_τ and σ_τ. Here, τ is an arbitrary integer between 1 and n. The option prices used to estimate the observed risk neutral distribution F are quoted at time $\tau - 1$, with the options expiring at time τ. In our application, we implement our test separately for each of 128 dates between 1997 and 2008. Taking τ as fixed, our estimator G_n of G is given by

$$G_n(x) = \frac{1}{n} \sum_{t=1}^{n} 1\left(\frac{\sigma_\tau}{\sigma_t} (q_t - r_t) + r_\tau \leq \ln x \right)$$

if $x > 0$, or $G_n(x) = 0$ if $x \leq 0$. We shall obtain a uniform limit theory for $\sqrt{n}(G_n - G)$ under the following technical conditions.

Assumption 3.1. The following statements are true.

(a) $(q_t - r_t)/\sigma_t$, $t \in \mathbb{N}$, is an iid sequence of random variables.

(b) G is the cdf of $\exp(q_\tau)$, treating r_τ and σ_τ as known.

Assumption 3.1 states that market returns q_t are iid once we subtract the risk-free rate of interest r_t and deflate by our realized volatility measure σ_t. Clearly, this is an imperfect approximation to reality, but some form of stationarity condition must be employed in order to estimate G from historical data. The recent empirical studies of pricing kernel monotonicity by Golubev et al. (2008) and Härdle et al. (2010) treat the sequence of market returns q_t as iid.

We are now in a position to state the following result, which gives the form of the limiting process ζ that appears in Assumption 2.2 when G_n is constructed in the manner just described.

Theorem 3.2. *Suppose Assumptions 2.1 and 3.1 are satisfied. Then, as $n \to \infty$, we have $\sqrt{n}\,(G_n - G) \rightsquigarrow \zeta \circ G$ in $\ell^\infty(\mathbb{R})$, where $\zeta = B$, a standard Brownian bridge on $[0,1]$.*

The proof of Theorem 3.2 is a simple application of Donsker's theorem.

3.3 Calculating critical values

In Theorem 2.1(a), we expressed the limiting distributions of our test statistics T_n and T_n' when π is constant as functionals of the random process $\mathscr{G}$, where $\mathscr{G}(u) = \xi(u) - \zeta(u)$, and ξ and ζ are random processes characterizing the estimation uncertainty associated with F_n and G_n. Theorems 3.1 and 3.2 deliver us, under suitable regularity conditions, explicit characterizations of ξ and ζ: we have

$$\xi(u) = \left(\frac{\partial}{\partial \theta} F(F^{-1}(u;\theta^*);\theta) \Big|_{\theta=\theta^*} \right)' \Psi^{1/2} N \quad \text{and} \quad \zeta(u) = B(u), \tag{3.3}$$

where N is a $4k$-vector of independent standard normal random variables and B is a standard Brownian bridge. In Assumption 2.2, we required that ξ and ζ are independent random processes. With ξ and ζ given by (3.3), this amounts to assuming that N and B are mutually independent. A sufficient condition for independence of ξ and ζ is that our estimated distributions F_n and G_n are independent of one another. This seems a reasonable supposition, at least as an approximation. Variation in F_n is driven by the random pricing errors occurring at a particular date, while variation in G_n is driven by the history of volatility-adjusted excess returns.

As explained in Section 2, we follow Carolan and Tebbs (2005) by choosing the critical values for our test statistics to deliver the correct asymptotic size when π is constant. Under this condition, the distribution of the random process $\mathscr{G}$ may be parametrized in terms of the unknown quantities θ^* and Ψ. We shall write

$$\mathscr{G}(u;\theta^*,\Psi) = \left(\frac{\partial}{\partial\theta} F\left(F^{-1}(u;\theta^*);\theta\right)\bigg|_{\theta=\theta^*} \right)' \Psi^{1/2} N - B(u). \tag{3.4}$$

Though the distribution of $\mathscr{G}(\cdot;\theta^*,\Psi)$ is unknown, we may approximate it by substituting consistent estimators (θ_n,Ψ_n) for (θ^*,Ψ) in (3.4), and then simulating the process $\mathscr{G}(\cdot;\theta_n,\Psi_n)$ by drawing a large number of realizations of (N,B). The estimator θ_n introduced in Section 3.1 is known to be consistent from Theorem 3.1(a). We may estimate Ψ using

$$\Psi_n = \frac{n}{m} M_n \Xi_n^{-1} \Sigma_n \Xi_n^{-1} M_n',$$

where the sample analogues to Σ, Ξ and M are given by

$$\Sigma_n = \frac{1}{m}\sum_{i=1}^{m} v_i^2 z_i z_i' \left(p_i - z_i'\theta_n\right)^2, \quad \Xi_n = \frac{1}{m}\sum_{i=1}^{m} v_i z_i z_i', \quad M_n = I - \Xi_n^{-1} R' \left(R\Xi_n^{-1}R'\right)^{-1} R.$$

Under Assumption A.1 it is straightforward to show that Ψ_n provides a consistent estimate of Ψ. We obtain critical values for T_n and T_n' with an asymptotic rejection rate of α when π is constant by drawing a large number of independent realizations of N and B, and using them to compute the $1-\alpha$ quantiles of the distributions of

$$\int_0^1 \mathcal{D}\mathscr{G}(u;\theta_n,\Psi_n)\mathrm{d}u \quad \text{and} \quad \sup_{u\in[0,1]} \mathcal{D}\mathscr{G}(u;\theta_n,\Psi_n).$$

One may show that this approach generates consistent estimates of the $1-\alpha$ quantiles of $\int_0^1 \mathcal{D}\mathcal{G}(u)\mathrm{d}u$ and $\sup_{u\in[0,1]}\mathcal{D}\mathscr{G}(u)$. We omit the details in the interests of brevity.

It is clear from (3.3) and (3.4) that the dependence of the limiting distributions of T_n and T_n' on θ^* and Ψ is due entirely to the estimation uncertainty associated with the risk neutral distribution F. Uncertainty about the physical distribution G contributes the term B appearing in (3.4), which is free of nuisance parameters. If the risk neutral distribution were known with certainty, we would have $\xi = 0$ and therefore $\mathscr{G} = -B$, and asymptotic critical values for T_n and T_n' could be obtained from the quantiles of

$\int_0^1 \mathcal{D}B(u)\mathrm{d}u$ and $\sup_{u\in[0,1]} \mathcal{D}B(u)$. Carolan and Tebbs (2005) report that these critical values are 0.66 and 1.474 at the 5% significance level. In the robustness checks for our empirical application reported in Section 4.4, we find that the results of our monotonicity tests change very little when pivotal critical values obtained under the assumption $\xi = 0$ are used in place of simulated critical values accounting for uncertainty about F. This is because the estimated covariance matrix Ψ_n is usually very close to zero, reflecting the fact that the pricing errors associated with the fitted risk neutral density are small.

4 Empirical results

The presentation of our empirical results is divided into four subsections. In Section 4.1 we describe our data set, summarizing some of the key features. In Section 4.2 we provide a detailed description of our results for December 16, 2009 – the most recent date at which we apply our test. In Section 4.3 we summarize our results for the full sample. In Section 4.4 we report the outcome of a number of robustness checks used to confirm the validity of our results.

4.1 Data

Our primary dataset consists of prices for European call and put options written on the S&P 500 from January 1997 to December 2009. Daily option price data were purchased from DeltaNeutral[1], which collects prices for options reported by the Options Price Reporting Authority[2] (OPRA). OPRA compiles information from a number of different exchanges in order to find a national option price. Our dataset consists of bid-ask averages of OPRA's reported prices for different options at the close of the market on each trading day, along with the corresponding trading volumes.

We exclude all options that do not have between 18 and 22 trading days to maturity, and nonzero trading volume. Tables 4.1 and 4.2 summarize some key features of the remaining prices. In Table 4.1 we report the maximum, minimum and average number of put and call

[1] `http://www.deltaneutral.com`.

[2] `http://www.opradata.com`.

	Number of calls			Number of puts		
Year	Avg	Min	Max	Avg	Min	Max
1997	29	24	34	35	26	52
1998	27	21	38	36	27	43
1999	26	21	37	33	24	47
2000	25	15	38	28	19	42
2001	19	12	25	28	20	43
2002	22	16	39	29	18	42
2003	25	19	31	31	24	40
2004	30	27	35	36	31	49
2005	32	28	40	40	35	46
2006	37	28	45	54	43	67
2007	48	34	80	73	56	93
2008	81	58	133	85	64	111
2009	71	60	85	98	79	114

Table 4.1: Option count statistics.

option prices observed in the different months of each year. The average number of option prices increased from 64 in 1997 to 169 in 2009. Most prices were roughly at-the-money or out-of-the-money. Towards the end of our sample, a larger percentage of observed option prices were far out-of-the-money, particularly for puts. In Table 4.2 we report average implied volatilities by option moneyness, computed using the Black-Scholes formula. For all moneyness categories, we find that implied volatilities vary substantially over time. Implied volatilities were moderately high during the the late 1990s and early 2000s, at their lowest in the mid-2000s, and reached their highest levels during the financial crisis years at the end of our sample. Average implied volatilies exhibit the familiar volatility smile, where at-the-money options tend to have lower implied volatilities than their very in-the-money or out-of-the-money counterparts. Moreover, we observe some asymmetry in the smile, in that options with moneyness lower than -10% tend to have higher implied volatilities than options with moneyness greater than 10%. This difference appears larger in periods with generally higher implied volatilities.

For the benchmark results reported in Sections 4.2 and 4.3, we exclude options with moneyness outside $\pm 15\%$ from the risk neutral estimation procedure. In the robustness

	Average implied volatility by moneyness category									
Year	<-15	-15:-10	-10:-6	-6:-3	-3:0	0:3	3:6	6:10	10:15	>15
1997	46	30	25	22	21	20	19	21	25	36
1998	54	32	28	23	21	19	18	18	21	28
1999	47	32	28	24	22	21	19	17	19	25
2000	43	30	26	23	21	20	19	20	22	35
2001	45	35	29	25	24	23	22	22	24	44
2002	45	35	31	27	24	23	23	23	27	47
2003	38	28	25	21	20	18	19	21	23	40
2004	35	24	19	16	14	13	13	15	19	41
2005	34	22	17	14	11	10	10	12	17	29
2006	35	23	18	14	12	10	10	12	17	23
2007	36	26	22	18	16	14	13	14	17	23
2008	62	39	35	31	30	28	27	27	27	42
2009	55	36	31	29	27	26	25	25	25	33

Table 4.2: Option implied volatilities.

checks reported in Section 4.4, we find that including these far away-from-the-money options leads to a small increase in the rejection rate of our tests. The effect is minimal because of the volume weighting incorporated in the estimation procedure.

Our physical volatility measure is derived from a daily series of realized volatilities kindly provided by Peter R. Hansen. These realized volatilities were constructed using high frequency intraday data on the SPY, an exchange traded fund which tracks the S&P 500 index, using the realized kernel method described by Barndorff-Nielsen et al. (2008, 2009). Following Hansen et al. (2011), we used a log-linear RealGARCH(1,1) model to produce forward-looking volatilities from the historical SPY returns and realized volatilities.[3] In Figure 4.1 we graph the realized volatility forecasts alongside implied volatilities extracted from options with moneyness in the range $\pm 3\%$ using the Black-Scholes formula. The level of aggregation is 20 days. We can see that the realized volatility forecasts track the implied volatilities fairly closely.

[3] Code for estimating the RealGARCH(1,1) model is available at `http://qed.econ.queensu.ca/jae/datasets/hansen003`.

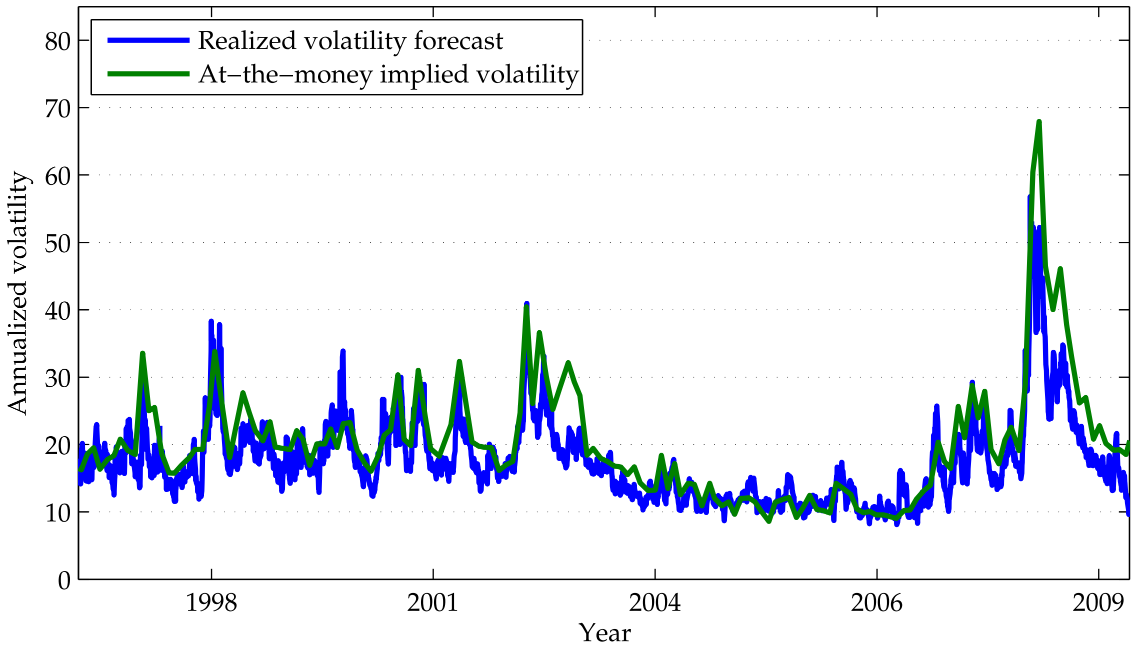

Figure 4.1: Forward looking realized and implied volatilities.

Data for the risk-free interest rate were obtained from Kenneth French's website[4], which reports the daily return on the one-month Treasury bill rate (from Ibbotson Associates).

4.2 Representative example

In this subsection we present detailed results for the date December 16, 2009. This is the most recent date at which we apply our tests. In Figure 4.2 we graph the estimated risk neutral and physical densities for this date, as well as their ratio, the implied pricing kernel. The physical density was obtained by applying a standard kernel smoother to our estimated physical cdf; this is purely for graphical purposes, and not used in our testing procedure. We can see that the physical density has a sharper peak than the risk neutral density, while the risk neutral density has fatter tails than the physical density. This results in a U-shaped pricing kernel, similar to the findings of Bakshi et al. (2010). Care should be exercised when judging the shape of the pricing kernel, as sampling uncertainty

[4] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

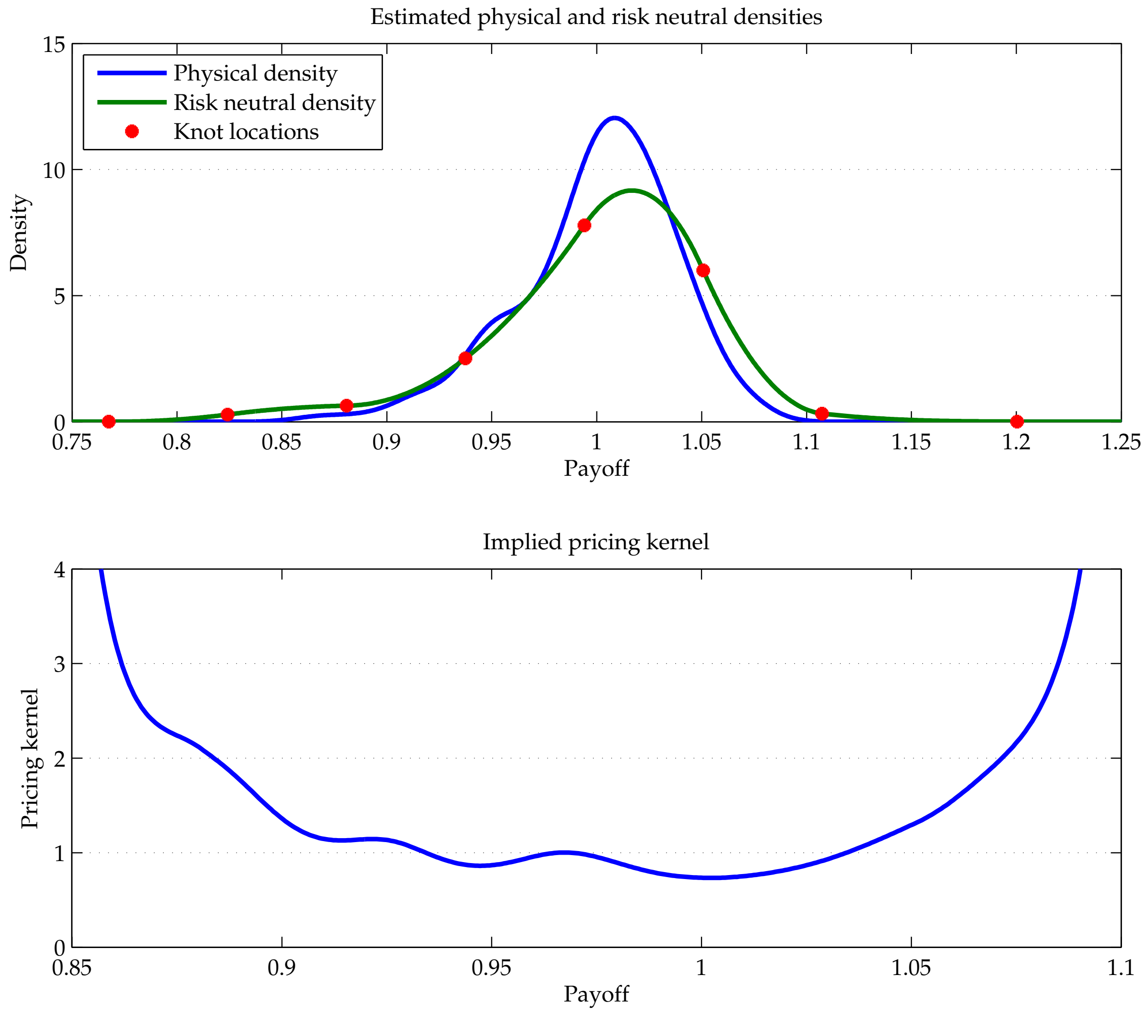

Figure 4.2: Estimated densities and pricing kernel for December 16, 2009.

may be large in the tails of either distribution.

We used $k = 11$ spline knots in the risk neutral estimation procedure. Their locations are represented by red dots in Figure 4.2; 3 of the 11 knots lie outside the range of the horizontal axis. The knot locations were chosen using a fairly complicated *ad hoc* procedure exploiting information about strike prices, trading volumes, and the estimated physical distribution, in order to achieve a sensible distribution of knots. It appears to work well in practice. Details are available on request. We use the same knot selection procedure for all 128 dates at which we apply our tests.

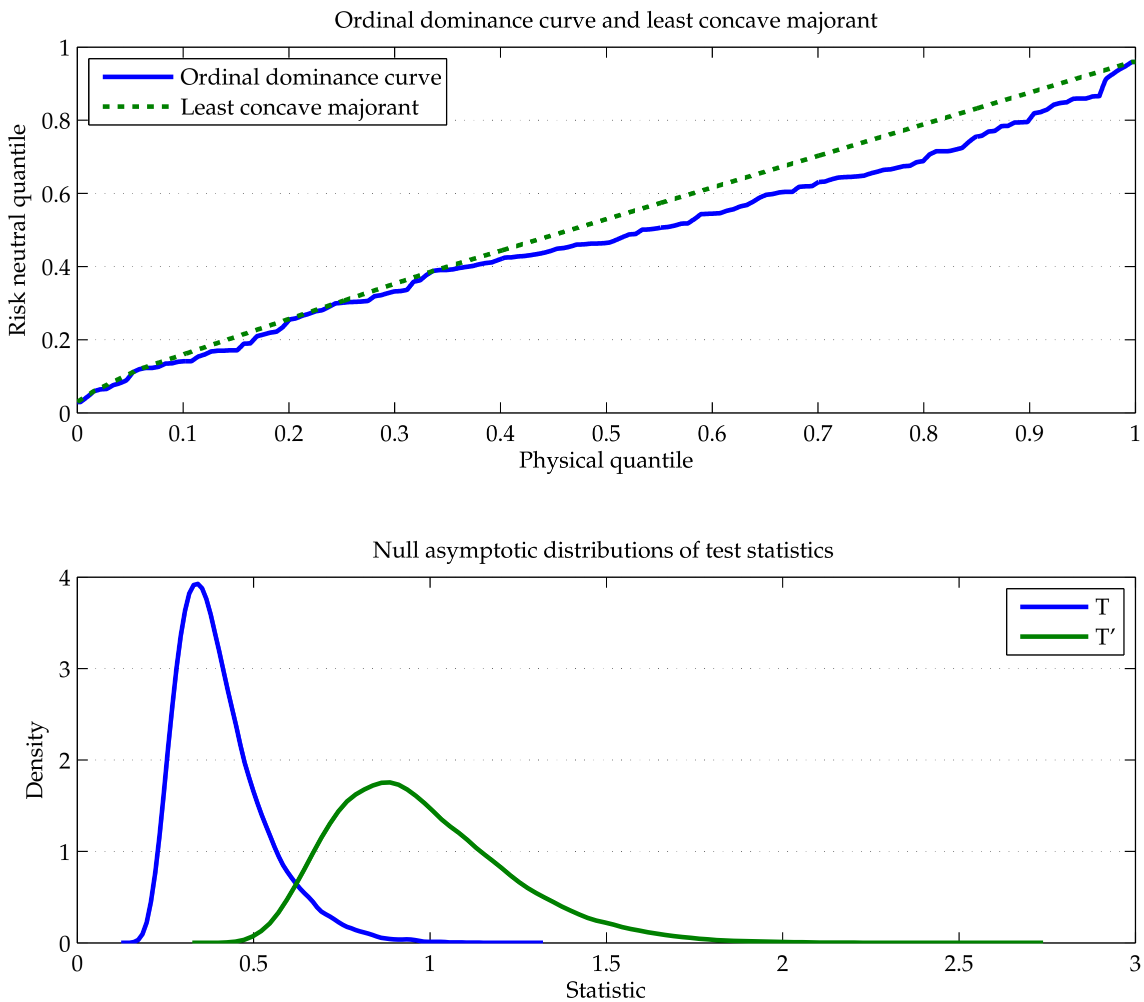

Figure 4.3: Monotonicity test results for December 16, 2009. $T = 0.576$ (p-value: 0.105) and $T' = 1.301$ (p-value: 0.108).

Figure 4.3 presents our test results for December 16, 2009. The top panel plots the estimated odc and its lcm. The estimated odc appears to be roughly convex to the right of the center of the physical distribution, consistent with a U-shaped pricing kernel. This shape for the odc was fairly common in the months where we rejected the null hypothesis. The test statistics T and T' are equal to 0.576 and 1.301 respectively. To calculate critical values and p-values, we simulated 50,000 observations from the estimated null asymptotic distribution of each test statistic, as described in Section 3.3. The simulated densities are

Year	Number of months	Discounted mean			Annualized volatility		
		Avg	Min	Max	Avg	Min	Max
1997	12	1.002	0.998	1.005	0.208	0.170	0.292
1998	11	1.003	1.000	1.005	0.223	0.161	0.381
1999	11	1.003	0.999	1.005	0.220	0.176	0.279
2000	9	1.004	0.998	1.007	0.211	0.153	0.301
2001	2	1.002	1.002	1.002	0.293	0.261	0.326
2002	4	1.000	0.999	1.002	0.239	0.188	0.286
2003	11	1.000	0.999	1.000	0.196	0.150	0.294
2004	12	0.999	0.997	1.003	0.145	0.115	0.187
2005	12	1.001	0.998	1.003	0.124	0.098	0.140
2006	12	1.002	1.000	1.003	0.123	0.106	0.167
2007	12	1.002	1.001	1.004	0.164	0.104	0.259
2008	8	1.000	0.996	1.005	0.231	0.181	0.292
2009	12	0.998	0.996	1.002	0.293	0.201	0.447

Table 4.3: Moments of estimated risk neutral distributions.

plotted in the lower panel of Figure 4.3. The p-values associated with the statistics T and T' are 0.105 and 0.108 respectively. We conclude that there is weak evidence against the null hypothesis of a nonincreasing pricing kernel at this date.

4.3 Full sample results

For each month in the sample period, we select the date with the largest number of option prices with a single time-to-maturity, and test the monotonicity of the pricing kernel at that date, with that time-to-maturity. Recall that we only include options with time-to-maturity between 18 and 22 days in our sample. We exclude 24 months in which there are no days with prices for at least 40 options with a common time-to-maturity, and the last four months of 2008 due to the extremely high realized volatility levels associated with the financial crisis. This leaves us with 128 dates at which we apply our tests.

Table 4.3 presents summary statistics for the moments of the estimated risk neutral densities by year. The mean of the estimated risk neutral distribution, discounted by the risk-free rate, acts as a simple specification check on our cubic spline model. It should

	Number of	T statistic				T' statistic			
Year	months	10%	5%	1%	0.1%	10%	5%	1%	0.1%
1997	12	25	17	8	0	17	8	8	8
1998	11	36	36	18	0	45	27	0	0
1999	11	45	18	9	0	36	27	9	0
2000	9	22	11	0	0	33	0	0	0
2001	2	50	50	50	0	50	50	50	0
2002	4	0	0	0	0	0	0	0	0
2003	11	18	18	9	0	18	18	9	9
2004	12	0	0	0	0	25	8	0	0
2005	12	25	17	0	0	67	67	58	25
2006	12	42	17	0	0	67	67	50	8
2007	12	33	8	0	0	67	58	42	25
2008	8	25	12	12	0	25	25	12	0
2009	12	25	17	8	8	25	25	8	8
All Years	128	27	16	6	1	38	30	19	8

Table 4.4: Rejection rates (%) for monotonicity tests.

equal one if a direct investment in the S&P 500 index is correctly priced. We can see that the discounted mean is in all cases very close to one. Table 4.3 also summarizes the annualized volatilities associated with our estimated risk neutral distributions. As we would expect, the overall pattern of the volatilities roughly matches that of the option implied volatilities, plotted in Figure 4.1.

Table 4.4 and Figure 4.4 present the results of our monotonicity tests. In Table 4.4 we report the rejection frequency obtained with our two test statistics in each year, at 10%, 5%, 1% and 0.1% significance levels. We find that our test statistics, particularly the T' statistics, exceed their asymptotic critical values more frequently than we would expect under the null hypothesis of monotonicity. The null hypothesis is rejected 16% and 30% of the time at the 5% significance level with the T and T' statistics respectively. The T' statistic generates rejections at the 5% significance level in half of the months during the years 2005-2009. In Figure 4.4 we plot the p-values using the two test statistics over time. As we would expect, the two series track each other fairly closely, having a correlation of 0.90. Some p-values are very close to zero, indicating that we may confidently reject the null hypothesis of monotonicity at certain dates.

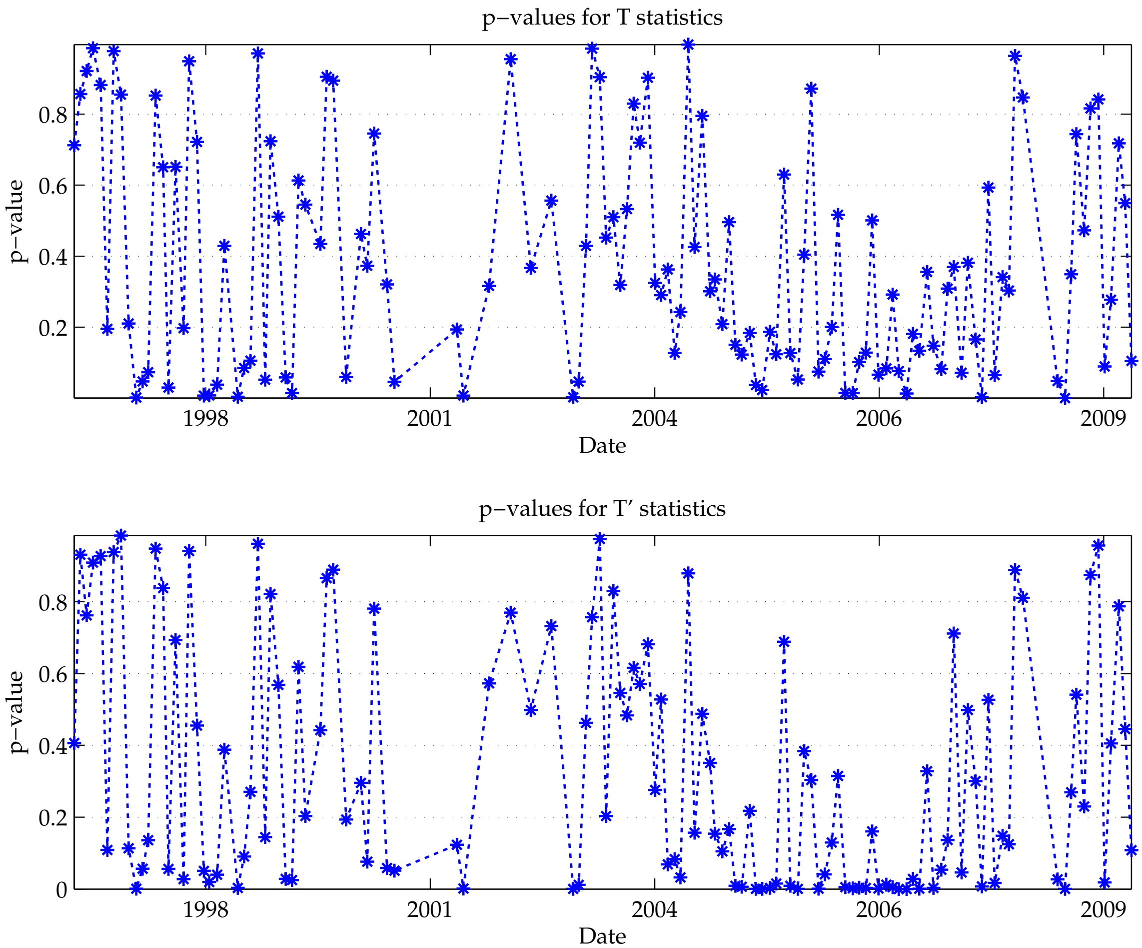

Figure 4.4: p-values for monotonicity tests.

4.4 Robustness checks

In Table 4.5 we report the overall rejection rates obtained using our two test statistics when a number of changes are made to our use of data and estimation of the risk neutral distribution. Here we consider four variations on the risk neutral estimator.

1. The risk neutral distribution is assumed known. In this case we use the same risk neutral estimator, but compare our statistics to critical values obtained under the assumption $\xi = 0$, as discussed at the end of Section 3.3.

Panel A: -15% ≤ option moneyness ≤ 15%

Specification	Number of months	T statistic 10%	5%	1%	0.1%	T' statistic 10%	5%	1%	0.1%
Benchmark	128	27	16	6	1	38	30	19	8
Risk neutral known	128	28	18	6	2	38	33	19	8
Bid-ask weights	128	36	22	8	1	43	34	22	9
Equal weights	128	29	21	6	1	40	30	20	9
9 knots	128	26	13	5	1	38	28	15	9

Panel B: No moneyness restrictions

Specification	Number of months	T statistic 10%	5%	1%	0.1%	T' statistic 10%	5%	1%	0.1%
Benchmark	142	27	17	6	1	35	28	18	6
Risk neutral known	142	28	18	6	1	37	29	18	8
Bid-ask weights	142	30	23	7	1	39	32	17	6
Equal weights	142	30	20	6	1	39	28	16	9
9 knots	142	25	14	5	1	33	25	13	6

Table 4.5: Overall rejection rates (%) under alternative specifications of the risk neutral distribution estimator, with and without moneyness restrictions.

2. Bid-ask spreads are used in place of trading volumes to weight the squared pricing errors when estimating the parameters of the cubic spline.

3. Equal weights are applied to the squared pricing errors when estimating the parameters of the cubic spline.

4. A total of 9 spline knots are used, rather than 11.

The empirical results reported in Sections 4.2 and 4.3 excluded options with moneyness outside the range ±15% from the sample. We consider the four variations on the risk neutral estimator with this moneyness restriction in place, and also with options of all moneyness included in the sample.

Table 4.5 reveals that the effects of making these changes are modest. Treating the risk neutral distribution as known, which amounts to using smaller critical values for our tests, leads to a very small increase in the rejection rate. This reflects the fact, noted also by Golubev et al. (2008), that sampling uncertainty about the risk neutral distribution is minor compared to sampling uncertainty about the physical distribution. Using equal weights or bisk-ask spreads in place of trading volumes to estimate the parameters of the cubic spline appears to increase the rejection rate by a modest amount. Reducing the number of spline knots from 11 to 9, or removing the moneyness restrictions on options, leads to a small overall reduction in rejection rates.

5 Final remarks

We have proposed a new test of pricing kernel monotonicity, and applied it to a sample of thirteen years of options market data for the S&P 500 index. Statistically significant violations of monotonicity are detected at a substantial proportion of the dates considered. Our results provide empirical support for the growing literature on the so-called "pricing kernel puzzle", indicating that well documented nonmonotonicities in estimated pricing kernels cannot always be attributed to sampling uncertainty.

The robustness checks reported in Table 4.5, and further checks that are unreported, reveal that our results are relatively insensitive to the way in which the risk neutral distribution is specified. Uncertainty about the physical distribution plays a dominant role in determining the outcome of our monotonicity tests. In particular, a sensible approach to modeling time variation in volatility is required to obtain meaningful results. We saw in Figure 4.1 that our realized volatility forecasts track the option-implied volatilities quite closely. When the two volatilities part company, we are much more likely to see rejections of pricing kernel monotonicity, with the ratio of densities potentially adopting a U- or inverted U-shape as one density becomes much more concentrated or disperse than the other. Any empirical study of the shape of the pricing kernel must be closely informed by the much larger literature on volatility modeling. Our adoption of realized volatility methods based on high frequency data represents our best attempt to keep astride of current developments in this area.

A Mathematical appendix

A.1 Technical conditions for risk neutral distribution estimation

Assumption A.1. The following statements are true.

(a) f is a cubic spline of the form (3.1), with $\theta^* \in \Theta$.

(b) $\{(d_i, s_i, v_i) : i \in \mathbb{N}\}$ is a collection of nonrandom elements of $\{0,1\} \times (0,\infty) \times (0,\infty)$, while $\{p_i : i \in \mathbb{N}\}$ is a collection of real valued random variables.

(c) For each $i \in \mathbb{N}$, $p_i = h(d_i, s_i; \theta) + e_i$, with $\{e_i : i \in \mathbb{N}\}$ a mutually independent collection of real valued random variables, each with $E(e_i) = 0$.

(d) $\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^m v_i z_i z_i' = \Xi$ for some positive definite matrix Ξ.

(e) $\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^m v_i^2 z_i z_i' E(e_i^2) = \Sigma$ for some positive semidefinite matrix Σ.

(f) $\sup_{i\in\mathbb{N}} E|v_i z_{i,j} e_i|^{2+\delta} < \infty$ for each $j = 1, \ldots, 4k$ and some $\delta > 0$.

(g) f is strictly positive on (x_0, x_k), with strictly positive right-derivative at x_0 and strictly negative left-derivative at x_k.

Part (a) of Assumption A.1 ensures that f is a smooth cubic spline. Parts (b)-(f) are standard regularity conditions for weighted linear regression with fixed regressors. Part (g) ensures that the inequality restrictions in (3.2) have an asymptotically negligible impact upon the distribution of $\sqrt{n}(\theta_n - \theta^*)$.

A.2 Mathematical proofs

Proof of Theorem 2.1. To prove part (a), we begin by showing that

$$\sqrt{n}(\phi_n - \phi) \rightsquigarrow \mathscr{G} \quad \text{in } \ell^\infty[0,1] \tag{A.1}$$

when $F = G$. Let $\mathcal{I}$ denote the identity function on $[0,1]$. Observe that

$$\sqrt{n}\,(\phi_n - \phi) = \sqrt{n}\,(F_n - F) \circ G_n^{-1} + \sqrt{n}\left(F \circ G_n^{-1} - F \circ G^{-1}\right). \tag{A.2}$$

Focusing first on the second term in (A.2), we use the assumption $F = G$ to obtain

$$\sqrt{n}\left(F \circ G_n^{-1} - F \circ G^{-1}\right) = \sqrt{n}\left((G_n \circ G^{-1})^{-1} - \mathcal{I}\right). \tag{A.3}$$

Under Assumption 2.2, we have $\sqrt{n}(G_n \circ G^{-1} - \mathcal{I}) \rightsquigarrow \zeta$ in $\ell^\infty[0,1]$. Using the Hadamard differentiability of the inverse operator at the identity function (Lemma 3.9.23(ii) in van der Vaart and Wellner, 1996) – which holds since ζ is continuous under Assumption 2.2 – we may apply the functional delta method (Theorem 3.9.4 in van der Vaart and Wellner, 1996) to obtain

$$\sqrt{n}((G_n \circ G^{-1})^{-1} - \mathcal{I}) \rightsquigarrow -\zeta \quad \text{in } \ell^\infty[0,1]. \tag{A.4}$$

This takes care of the second term on the right-hand side of (A.2). The first term on the right-hand side of (A.2) may be written as

$$\sqrt{n}\,(F_n - F) \circ G_n^{-1} = \left(\sqrt{n}(F_n \circ F^{-1} - \mathcal{I})\right) \circ \left(F \circ G_n^{-1}\right).$$

Under Assumption 2.2, we have $\sqrt{n}(F_n \circ F^{-1} - \mathcal{I}) \rightsquigarrow \xi$ in $\ell^\infty[0,1]$. Also, from (A.3) and (A.4) we have $F \circ G_n^{-1} \rightsquigarrow F \circ G^{-1}$ in $\ell^\infty[0,1]$. Therefore, applying the continuous mapping theorem to the operation of composition – which is justified by the uniform continuity of ξ under Assumption 2.2 – we find that

$$\left(\sqrt{n}(F_n \circ F^{-1} - \mathcal{I})\right) \circ \left(F \circ G_n^{-1}\right) \rightsquigarrow \xi \quad \text{in } \ell^\infty[0,1].$$

This takes care of the second term on the right-hand side of (A.2), and proves (A.1).

We complete the proof of (a) with another application of the continuous mapping theorem. One may show without difficulty that the lcm operator $\mathcal{M} : \ell^\infty[0,1] \to \ell^\infty[0,1]$ and related operator $\mathcal{D} : \ell^\infty[0,1] \to \ell^\infty[0,1]$ are positive homogenous of degree one. One may also show that $\mathcal{M}(\gamma_1 + \gamma_2) = \mathcal{M}\gamma_1 + \gamma_2$ and $\mathcal{D}(\gamma_1 + \gamma_2) = \mathcal{D}\gamma_1$ whenever $\gamma_1, \gamma_2 \in \ell^\infty[0,1]$ and γ_2 is affine; see p. 168 in Carolan and Tebbs (2005). Since ϕ is affine when $F = G$, we may therefore write

$$T_n = \int_0^1 \mathcal{D}\left(\sqrt{n}(\phi_n - \phi)\right)(u)\mathrm{d}u \quad \text{and} \quad T_n' = \sup_{u\in[0,1]} \mathcal{D}\left(\sqrt{n}(\phi_n - \phi)\right)(u).$$

In view of (A.1) and the continuity of the integral and supremum operations, part (a) of our theorem now follows from the continuous mapping theorem if we can show that $\mathcal{D}$ is continuous. In fact, continuity follows immediately from Marshall's Lemma (Robertson et al., 1988, ch. 7), which states that $\|\mathcal{M}\gamma_1 - \mathcal{M}\gamma_2\| \leq \|\gamma_1 - \gamma_2\|$ for any $\gamma_1, \gamma_2 \in \ell^\infty[0,1]$.

Next we briefly sketch a proof of part (b). If ϕ is not concave, then we have $\mathcal{D}\phi(u) > 0$ for all u in some open interval A with endpoints in $(0,1)$. Under Assumption 2.2, an application of the continuous mapping theorem yields $\phi_n \to_p \phi$ uniformly on A. It follows that $\int_A \mathcal{D}\phi_n(u)\mathrm{d}u \to_p \int_A \mathcal{D}\phi(u)\mathrm{d}u > 0$ and $\sup_A \mathcal{D}\phi_n \to_p \sup_A \mathcal{D}\phi > 0$. Our desired result follows easily. □

Proof of Theorem 3.1. First we shall prove part (a). The main inconvenience here is the presence of the inequality restrictions (3.2), which we temporarily dispose of. Let $\tilde{\Theta}$ be the collection of $\theta \in \mathbb{R}^{4k}$ for which $R\theta = 0$, and let

$$\tilde{\theta}_n = \operatorname*{argmin}_{\theta\in\tilde{\Theta}} \sum_{i=1}^m v_i \left(p_i - h(d_i, s_i; \theta)\right)^2 = \operatorname*{argmin}_{\theta\in\tilde{\Theta}} \sum_{i=1}^m v_i \left(p_i - z_i'\theta\right)^2.$$

$\tilde{\theta}_n$ is simply a weighted least squares estimator subject to linear equality restrictions. With some elementary algebra we may show that

$$\sqrt{n}\left(\tilde{\theta}_n - \theta^*\right) = \sqrt{\frac{n}{m}} M_n \Xi_n^{-1} \frac{1}{\sqrt{m}} \sum_{i=1}^m v_i z_i e_i,$$

where Ξ_n is invertible for n sufficiently large under Assumption A.1(d). This condition also ensures that $\Xi_n \to \Xi$ and $M_n \to M$ as $n \to \infty$. An application of the Liapounov central limit theorem gives $m^{-1/2}\sum_{i=1}^m v_i z_i e_i \to_d N(0,\Sigma)$ under Assumption A.1(b,c,e,f). Since $n/m \to \lambda$, we conclude that $\sqrt{n}(\tilde{\theta}_n - \theta^*) \to_d \Psi^{1/2}N$.

If the probability of the inequality restrictions (3.2) binding vanishes in the limit, so that $P(\tilde{\theta}_n = \theta_n) \to 1$, then we will have $\sqrt{n}(\theta_n - \theta^*) = \sqrt{n}(\tilde{\theta}_n - \theta^*) + o_p(1)$, and the proof of

(a) will be complete. Under Assumption A.1(a,g) we may choose $\epsilon > 0$ and $\delta > 0$ such that $f'(\cdot;\theta^*) \geq \delta$ on $(x_0, x_0+\epsilon)$, $f(\cdot;\theta^*) \geq \delta$ on $[x_0+\epsilon, x_k-\epsilon]$, and $f'(\cdot;\theta^*) \leq -\delta$ on $(x_k-\epsilon, x_k)$. If $f(\cdot;\tilde{\theta}_n) \to_p f(\cdot;\theta^*)$ and $f'(\cdot;\tilde{\theta}_n) \to_p f'(\cdot;\theta^*)$ uniformly on (x_0, x_k), then the three quantities

$$P\left(\inf_{x\in(x_0,x_0+\epsilon)} f'(x;\tilde{\theta}_n) > 0\right),\ P\left(\inf_{x\in[x_0+\epsilon,x_k-\epsilon]} f(x;\tilde{\theta}_n) > 0\right),\ P\left(\inf_{x\in(x_k-\epsilon,x_k)} f'(x;\tilde{\theta}_n) < 0\right)$$

will all converge to one as $n \to \infty$. But then we must also have $f(\cdot;\tilde{\theta}_n) \geq 0$ on (x_0, x_k) with probability approaching one, so that $P(\tilde{\theta}_n = \theta_n) \to 1$. To prove uniform convergence in probability of $f(\cdot;\tilde{\theta}_n)$ to $f(\cdot;\theta^*)$ and of $f'(\cdot;\tilde{\theta}_n)$ to $f'(\cdot;\theta^*)$, we note that $f(x;\theta)$ and $f'(x;\theta)$ are linear in θ for each $x \in \mathbb{R}$ under Assumption A.1(a). Thus we may write $f(x;\tilde{\theta}_n) = f(x;\theta^*) + \left(\frac{\partial}{\partial\theta} f(x;\theta)|_{\theta=\theta^*}\right)' (\tilde{\theta}_n - \theta^*)$ and $f'(x;\tilde{\theta}_n) = f'(x;\theta^*) + \left(\frac{\partial}{\partial\theta} f'(x;\theta)|_{\theta=\theta^*}\right)' (\tilde{\theta}_n - \theta^*)$. Uniform convergence in probability now follows from the uniform boundedness of $\frac{\partial}{\partial\theta} f(x;\theta)|_{\theta=\theta^*}$ and $\frac{\partial}{\partial\theta} f'(x;\theta)|_{\theta=\theta^*}$ and the fact that $\tilde{\theta}_n \to_p \theta^*$. This proves (a).

To prove (b) we note that, under Assumption A.1(a), $F(x;\theta)$ is linear in θ for each $x \in \mathbb{R}$. Thus we may write

$$\sqrt{n}(F_n(x) - F(x)) = \sqrt{n}(F(x;\theta_n) - F(x;\theta^*)) = \left(\frac{\partial}{\partial\theta} F(x;\theta)\bigg|_{\theta=\theta^*}\right)' \sqrt{n}(\theta_n - \theta^*)$$

for each $x \in \mathbb{R}$ and $n \in \mathbb{N}$. Our desired result now follows from part (a) and the fact that $\frac{\partial}{\partial\theta} F(x;\theta)|_{\theta=\theta^*}$ is uniformly bounded in x. □

Proof of Theorem 3.2. This result is merely an application of Donsker's theorem to the sequence of random variables $\{\exp(\frac{\sigma_\tau}{\sigma_t}(q_t - r_t) + r_\tau) : t \in \mathbb{N}\}$, which are iid with cdf G under Assumption 3.1. □

References

Aït-Sahalia, Y. and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance* **53** 499–547.

Aït-Sahalia, Y. and Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics* **94** 9–51.

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.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008). Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* **76** 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2009). Realized kernels in practice: trades and quotes. *Econometrics Journal* **12** C1–C33.

Barone-Adesi, G., Engle, R. F. and Mancini, L. (2008). A GARCH option pricing model with filtered historical simulation. *Review of Financial Studies* **21** 1223–1258.

Barone-Adesi, G. and Dall'O, H. (2010). Is the price kernel monotone? Swiss Finance Institute Research Paper Series No. 10–03.

Beare, B. K. (2011). Measure preserving derivatives and the pricing kernel puzzle. UCSD Economics Working Paper No. 2010-01.

Carolan, C. A. and Tebbs, J. M. (2005). Nonparametric tests for and against likelihood ratio ordering in the two sample problem. *Biometrika* **92** 159–171.

Chabi-Yo, F., Garcia, R. and Renault, E. (2008). State dependence can explain the risk aversion puzzle. *Review of Financial Studies* **21** 973–1011.

Constantinides, G. M., Jackwerth, J. C. and Perrakis, S. (2009). Mispricing of S&P 500 index options. *Review of Financial Studies* **22** 1247–1277.

Detlefsen, K., Härdle, W. and Moro, R. (2007). Empirical pricing kernels and investor preferences. Sonderforschungsbereich 649 Discussion Paper 2007-017.

Dybvig, P. H. (1988). Distributional analysis of portfolio choice. *Journal of Business* **63** 369–393.

GOLUBEV, Y., HÄRDLE, W. AND TIMONFEEV, R. (2008). Testing monotonicity of pricing kernels. Sonderforschungsbereich 649 Discussion Paper 2008-001.

HANSEN, P. R., HUANG, Z. AND SHEK, H. H. (2011). Realized GARCH: A joint model for returns and realized measures of volatility. To appear in the *Journal of Applied Econometrics*.

HÄRDLE, W., KRÄTSCHMER, V. AND MORO, R. (2009). A microeconomic explanation of the EPK paradox. Sonderforschungsbereich 649 Discussion Paper 2009-010.

HÄRDLE, W., OKHRIN, Y. AND WANG, W. (2010). Uniform confidence bands for pricing kernels. Sonderforschungsbereich 649 Discussion Paper 2010-003.

HENS, T. AND REICHLIN, C. (2010). Three solutions to the pricing kernel puzzle. Swiss Finance Institute Research Paper Series No. 10–14.

HSIEH, F. AND TURNBULL, B. W. (1996). Nonparametric and semiparametric estimation of the receiver operating characteristic curve. *Annals of Statistics* **24** 25–40.

JACKWERTH, J. C. (2000). Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* **13** 433–451.

JACKWERTH, J. C. AND RUBINSTEIN, M. (1996). Recovering probability distributions from option prices. *Journal of Finance* **51** 1611–1631.

MONTEIRO, A. M., TÜTÜNCÜ, R. H. AND VICENTE, L. N. (2008). Recovering risk-neutral probability density functions from options prices using cubic splines and ensuring nonnegativity. *European Journal of Operational Research* **187** 525–542.

PATTON, A. J. AND TIMMERMANN, A. (2010). Monotonicity in asset returns: new tests with applications to the term structure, the CAPM, and portfolio sorts. *Journal of Financial Economics* **98** 605–625.

ROBERTSON, T., WRIGHT, F. AND DYKSTRA, R. (1988). *Order Restricted Statistical Inference.* Wiley, New York.

ROSENBERG, J. V. AND ENGLE, R. F. (2002). Empirical pricing kernels. *Journal of Financial Economics* **64** 341–372.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, New York.

Ziegler, A. (2007). Why does implied risk aversion smile? *Review of Financial Studies* **20** 859–904.