



ELSEVIER

Journal of Financial Economics 64 (2002) 341–372

JOURNAL OF
Financial
ECONOMICS

www.elsevier.com/locate/econbase

Empirical pricing kernels[☆]

Joshua V. Rosenberg^{a,*}, Robert F. Engle^b

^a *Federal Reserve Bank of New York, New York, NY 10045, USA*

^b *Stern School of Business, New York University, New York, NY 10012, USA*

Received 15 December 1997; accepted 29 August 2001

Abstract

This paper investigates the empirical characteristics of investor risk aversion over equity return states by estimating a time-varying pricing kernel, which we call the empirical pricing kernel (EPK). We estimate the EPK on a monthly basis from 1991 to 1995, using S&P 500 index option data and a stochastic volatility model for the S&P 500 return process. We find that the EPK exhibits counter cyclical risk aversion over S&P 500 return states. We also find that hedging performance is significantly improved when we use hedge ratios based the EPK rather than a time-invariant pricing kernel. © 2002 Elsevier Science B.V. All rights reserved.

JEL classification: G12; G13; C50

Keywords: Pricing kernels; Risk aversion; Derivatives; Hedging

[☆]We are grateful for comments from David Bates, Mark Broadie, Peter Bossaerts, Peter Carr, Ravi Bansal, Steve Figlewski, Joel Hasbrouck, Gur Huberman, Jens Jackwerth, Bruce Lehmann, Jose Lopez, Jayendu Patel, Mark Rubinstein, William Schwert, Allan Timmermann, two anonymous referees at the Journal of Financial Economics, and seminar participants at New York University, Columbia University, Indiana University, Georgetown University, Northwestern University, Boston University, Case Western University, the University of Iowa, the Federal Reserve Board of Governors, and the Federal Reserve Bank of New York. The authors also appreciate the comments of participants at the Conference on Risk Neutral and Objective Probability Distributions, the 1998 Western Finance Association meetings, the 1998 Computational and Quantitative Finance conference, and the 1997 Time Series Analysis of High Frequency Data Conference. The authors wish to thank David Hait for research assistance.

*Corresponding author. Tel.: +1-212-720-6317; fax: +1-212-720-1773.

E-mail address: joshua.rosenberg@ny.frb.org (J.V. Rosenberg).

0304-405X/02/\$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.

PII: S 0304-405X(02)00128-9

1. Introduction

The asset pricing kernel summarizes investor preferences for payoffs over different states of the world. In the absence of arbitrage, all asset prices can be expressed as the expected value of the product of the pricing kernel and the asset payoff. Thus, the pricing kernel, when it is used with a probability model for the states, gives a complete description of asset prices, expected returns, and risk premia.

In this paper, we estimate the pricing kernel using current asset prices and a predicted asset payoff density. We define the empirical pricing kernel (EPK) as the preference function that provides the “best fit” to asset prices, given the forecast payoff density. By estimating the EPK at a sequence of points in time, we can observe and model the dynamic structure of the pricing kernel itself. From this analysis, we obtain improved option pricing relations, hedging parameters, and a better understanding of the pattern of risk premia.

We estimate the EPK each month from 1991 to 1995, using S&P 500 index option data and a stochastic volatility model for the S&P 500 return process. We find substantial evidence that the pricing kernel exhibits counter cyclical risk aversion over S&P 500 return states. Empirical risk aversion is positively correlated with indicators of recession (widening of credit spreads) and negatively correlated with indicators of expansion (steepening of term structure slope).

We develop an option hedging methodology to compare the accuracy of several pricing kernel specifications. Our tests measure relative performance in hedging out-of-the-money S&P 500 put options using at-the-money S&P 500 put options and the S&P 500 index portfolio. We find that hedge ratios formed using a time-varying pricing kernel reduce hedge portfolio volatility more than hedge ratios based on a time-invariant pricing kernel.

Although there is a large literature on pricing kernel estimation using aggregate consumption data, problems with imprecise measurement of aggregate consumption can weaken the empirical results of these papers. Hansen and Singleton (1982, 1983) postulate that the pricing kernel is a power function of aggregate U.S. consumption. They use maximum-likelihood estimation and the generalized method of moments to estimate the pricing kernel. Chapman (1997) uses functions of consumption and its lags as pricing kernel state variables, and he specifies the pricing kernel function as an orthogonal polynomial expansion. Hansen and Jagannathan (1991) derive bounds for the mean and standard deviation of the consumption-based pricing kernel in terms of the mean and standard deviation of the market portfolio excess returns.

Recently, Ait-Sahalia and Lo (2000) have used option data and historical returns data to non-parametrically estimate the pricing kernel projected onto equity return states. This technique avoids the use of aggregate consumption data or a parametric pricing kernel specification. Along similar lines, Jackwerth (2000) non-parametrically estimates the “risk aversion function” using option data and historical returns data.

Ait-Sahalia and Lo (2000) and Jackwerth (2000) estimate investor expectations about future return probabilities by smoothing a histogram of realized returns over

the past four years. Implicitly, these papers assume that investors form probability beliefs by equally weighting events over the prior four years and disregarding previous events. For example, using a four-year window, the October 1987 stock market crash influences probability beliefs until October 1991. In November 1991, the crash no longer has an effect on beliefs.

These assumptions are inconsistent with evidence from the stochastic volatility modeling literature—e.g., Bollerslev et al. (1992)—indicating that future state probabilities depend more on the recent events than long-ago events, but that long-ago events still have some predictive power. Misspecification of state probabilities induces error in the estimation of the pricing kernel, since the denominator of the state-price-per-unit probability is incorrectly measured.

In Ait-Sahalia and Lo (2000) and Jackwerth (2000), state prices and probabilities are averaged over time, so their estimates are perhaps best interpreted as a measure of the average pricing kernel over the sample period. Since the sample periods used are at least one year in length, neither paper detects time variation at less than an annual frequency. Average pricing kernels are also limited in their ability to price and hedge assets on an ongoing basis, since assets are correctly priced only when risk aversion and state probabilities are at their average level. As noted by Ait-Sahalia and Lo (2000), “In contrast, the kernel SPD estimator is *consistent across time* [emphasis added] but there may be some dates for which the SPD estimator fits the cross section of option prices poorly and other dates for which the SPD estimator performs very well.”

The remainder of the paper is organized as follows. Section 2 describes the theory and previous research related to the pricing kernel. Section 3 presents the empirical pricing kernel estimation technique, EPK specification, and hedge ratio specification. Section 4 describes the data used for estimation, and Section 5 presents the estimation results. Section 6 contains the hedging test results, and Section 7 concludes the paper.

2. Theory and previous research

Our initial discussion of asset pricing kernel theory and previous research introduces several pricing kernel specifications and discusses some potential estimation problems. We then consider the characteristics of pricing kernel projections.

2.1. The asset pricing kernel

The asset pricing kernel is also known as the stochastic discount factor, since it is a state-dependent function that discounts payoffs using time and risk preferences. Campbell et al. (1997) and Cochrane (2001) provide comprehensive treatments of the role of the pricing kernel in asset pricing. Other related papers include Ross (1978), Harrison and Kreps (1979), Hansen and Richard (1987), and Hansen and Jagannathan (1991).

In the absence of arbitrage, the current price of an asset equals the expected pricing-kernel-weighted payoff:

$$P_t = E_t[M_t X_{t+1}], \quad (1)$$

where P_t is the current asset price, M_t is the asset pricing kernel, and X_{t+1} is the asset payoff in one period.

In Lucas (1978) consumption-based asset pricing model, the pricing kernel is equal to the intertemporal marginal rate of substitution, so $M_t = U'(C_{t+1})/U'(C_t)$. Under the assumption of power utility, the pricing kernel is $M_t = e^{-\rho}(C_{t+1}/C_t)^{-\gamma}$, with a rate of time preference of ρ and a level of relative risk aversion of γ .

One of the basic characteristics of the pricing kernel is its slope, and standard risk-aversion measures are usually functions the pricing kernel slope. For example, the Arrow-Pratt (Arrow, 1964; Pratt, 1964) measure of absolute risk aversion is the negative of the ratio of the derivative of the pricing kernel to the pricing kernel. The Arrow-Pratt measure of relative risk aversion is absolute risk aversion multiplied by current consumption:

$$\gamma_t = -[C_{t+1}M'_t(C_{t+1})]/M_t(C_{t+1}). \quad (2)$$

Generally, the pricing kernel will depend not only on current and future consumption, but also on all variables that affect marginal utility. In the habit persistence models of Abel (1990), Constantinides (1990), or Campbell and Cochrane (1999), the pricing kernel depends on both past and current consumption. Eichenbaum et al. (1988) let the pricing kernel depend on leisure, while Startz (1989) uses durable goods purchases. Bansal and Viswanathan (1993) specify the pricing kernel as a function of the equity market return, the Treasury bill yield, and the term spread.

When the pricing kernel is a function of multiple state variables, the level of risk aversion can also fluctuate as these variables change. Campbell (1996) shows that a habit persistence utility function exhibits time-varying relative risk aversion, where relative risk aversion is decreasing in the amount that consumption exceeds the habit (the surplus consumption ratio). In Campbell (1996) model, we observe decreases in relative risk aversion during economic expansions when consumption is high relative to the habit. Furthermore, we observe increases in relative risk aversion during economic contractions when consumption falls closer to the habit. In contrast, the power utility function exhibits relative risk aversion that is time-invariant.

To investigate the characteristics of investor preferences, many researchers have used Eq. (1) as an identifying equation for the pricing kernel. For example, Hansen and Singleton (1982) identify the pricing kernel with an unconditional version of this equation:

$$0 = E[(P_{t+1}/P_t)M_t - 1]. \quad (3)$$

Hansen and Singleton (1982), using an approach followed in many subsequent papers, specify the aggregate consumption growth rate as a pricing kernel state variable. They measure consumption using data from the National Income and Products Accounts (NIPA).

However, measurement error in the NIPA consumption data can pose a significant problem. Ermini (1989), Wilcox (1992), and Slesnick (1998) discuss issues such as coding errors, definitional problems, imputation procedures, and sampling error. Ferson and Harvey (1992) consider problems introduced by the Commerce Department's seasonal adjustment technique. Breeden et al. (1989) address problems induced by use of time-aggregated rather than instantaneous consumption.

2.2. Projections of the pricing kernel

Because there is considerable debate among researchers over the state variables that enter into the pricing kernel, we examine a pricing kernel projection that can be estimated without specifying these variables. We are particularly interested in projecting the pricing kernel onto the payoffs of a traded asset (X_{t+1}). As discussed in Cochrane (2001), this projected pricing kernel has exactly the same pricing implications as the original pricing kernel for assets with payoffs that depend on X_{t+1} .

To allow complete generality, we write the original pricing kernel as $M_t = M_t(Z_t, Z_{t+1})$, where Z_t is a vector of pricing kernel state variables. We then rewrite Eq. (1) by factoring the joint density $f_t(X_{t+1}, Z_{t+1})$ into the product of the conditional density $f_t(Z_{t+1} | X_{t+1})$ and the marginal density $f_t(X_{t+1})$. We evaluate the expectation in two steps:

$$P_t = E_t[M_t^*(X_{t+1})X_{t+1}], \quad M_t^*(X_{t+1}) = E_t[M_t(Z_t, Z_{t+1}) | X_{t+1}]. \quad (4)$$

First, the pricing kernel is integrated using the conditional density, which gives the projected pricing kernel, $M_t^*(X_{t+1})$.¹ Second, the product of the projected pricing kernel and the payoff variable is integrated using the marginal density, which gives the asset price, P_t .

The original pricing kernel depends on the realization of the state vector (Z_{t+1}), while the projected pricing kernel depends on the realization of the asset payoff (X_{t+1}). Thus, for the valuation of an asset with payoffs that depend only on X_{t+1} , the pricing kernel is summarized as a function of the asset payoff. This univariate function can vary over time, reflecting time variation in the pricing kernel state variables.

Eq. (4) can also be used to identify the projected pricing kernel. For example, Ait-Sahalia and Lo (2000) and Jackwerth (2000) estimate pricing kernels projected onto equity return states using equity index option prices. These papers assume that investors have a finite horizon and that the equity index level is equal to the

¹By taking this conditional expectation, we do not assume that X_{t+1} is known at date t . Rather, we are making a statement about the value of next period's pricing kernel, for each possible realization of next period's payoff variable. Since there is not necessarily a deterministic relation between next period's pricing kernel and the payoff variable, we measure this relation by taking the expectation based on information known at date t . The projected pricing kernel depends on the state of the world next period through X_{t+1} in the same way that the original pricing kernel depends on the state of the world next period through Z_{t+1} .

aggregate wealth. Under these assumptions, a pricing kernel that is projected onto the equity index level is equal to the original pricing kernel. Earlier papers, such as Rubinstein (1976) and Brown and Gibbons (1985), derive conditions such that a pricing kernel that has the consumption growth rate as a state variable is equivalent to a pricing kernel that has the equity index return as a pricing kernel state variable.

In general, we can interpret the projected pricing kernel the same way we interpret the original pricing kernel, even though the two are not necessarily identical. When $M_t^*(X_{t+1})$ is constant, investors are indifferent to a unit payoff across payoff states. When $M_t^*(X_{t+1})$ is decreasing (increasing), investors show a decreasing (increasing) desire for a unit payoff across payoff states.

We define a measure of risk aversion (γ_t^*) for the projected pricing kernel that is related to the Arrow-Pratt measure of relative risk aversion. We set the projected pricing kernel risk aversion equal to the opposite of the normalized slope of the projected pricing kernel, such that

$$\gamma_t^* = -[X_{t+1}M_t^{*'}(X_{t+1})]/M_t^*(X_{t+1}). \quad (5)$$

Ait-Sahalia and Lo (2000) and Jackwerth (2000) use similar formulas in their definitions of relative and absolute risk aversion functions.

The level of projected risk aversion determines the relative preference for a unit payoff across payoff states. High levels of γ_t^* correspond to a steep, negatively sloped pricing kernel projection, i.e., a strong demand for hedging securities that pay off when the asset price is low.

3. Empirical pricing kernel estimation strategy

In this section, we describe our methodology for estimation of a time-varying pricing kernel projected onto asset return states. We propose an optimization technique that selects the pricing kernel that best fits traded asset prices, and we suggest two pricing kernel specifications. We then present the stochastic volatility model used to estimate payoff probabilities, and we derive option hedge ratios in a setting with time-varying probabilities and a time-varying pricing kernel.

3.1. Estimation technique

The accuracy of a candidate pricing kernel can be judged by how well it reproduces prices of traded assets. This criterion motivates our estimation procedure. We select the empirical pricing kernel as the function that provides the best fit to current derivative prices, given current expectations about future payoffs. Therefore, the EPK represents an estimate of the pricing kernel projection on a particular date, rather than an estimate of an average pricing kernel over a period of a year or more.

We begin by writing Eq. (4) for a derivative with a payoff that depends on the return to the underlying asset (r_{t+1}):

$$P_{i,t} = E_t[M_t^*(r_{t+1})g_i(r_{t+1})] = \int M_t^*(r_{t+1})g_i(r_{t+1})f_i(r_{t+1}) dr_{t+1}. \tag{6}$$

In Eq. (6), $P_{i,t}$ is the price of the i th asset with a payoff function of $g_i(r_{t+1})$, and $f_i(r_{t+1})$ is the probability density of one-period underlying asset returns. Eq. (6) also shows that the pricing kernel projection, $M_t^*(r_{t+1})$, is an implicit function of prices, payoffs, and probabilities.

Next, we rewrite Eq. (6) to find the formula for the fitted asset price ($\hat{P}_{i,t}$) using an estimated pricing kernel projection ($\hat{M}_t^*(r_{t+1})$) and estimated payoff density ($\hat{f}_i(r_{t+1})$) such that

$$\hat{P}_{i,t} = E_t[\hat{M}_t^*(r_{t+1})g_i(r_{t+1})] = \int \hat{M}_t^*(r_{t+1})g_i(r_{t+1})\hat{f}_i(r_{t+1}) dr_{t+1}. \tag{7}$$

We then estimate the pricing kernel projection as the function that makes fitted prices closest to observed prices, using the estimated payoff density. To simplify the estimation problem, we let the pricing kernel projection be a parametric function, $\hat{M}_t^*(r_{t+1}; \theta_{t+1})$, where θ_t is an $N \times 1$ parameter vector. We use the sum of squared errors as a distance measure.

We refer to the projected pricing kernel that solves the following optimization problem as the EPK:

$$\text{Min}_{\theta_t} \sum_{i=1}^L [P_{i,t} - \hat{P}_{i,t}(\theta_t)]^2, \tag{8}$$

where L represents the number of asset prices, and $\hat{P}_{i,t}(\theta_t)$ is the fitted price as a function of the pricing kernel parameter vector. To identify the pricing kernel parameter vector, we must observe at least as many derivative prices as there are parameters.

If we define the payoff density using a set of J realized (or simulated) returns, then we can estimate the fitted asset price using the following approximation to Eq. (7), where averaging replaces integration:

$$\hat{P}_{i,t}(\theta_t) \cong J^{-1} \sum_{j=1}^J [M^*(r_{t+1,j}; \theta_t)g_i(r_{t+1,j})]. \tag{9}$$

3.2. Pricing kernel specifications

We consider two specifications for the projected pricing kernel. In the first specification, the pricing kernel is a power function of the underlying asset’s gross return:

$$M^*(r_{t+1}; \theta_t) = \theta_{0,t}(r_{t+1})^{-\theta_{1,t}}. \tag{10}$$

In Eq. (10), the first parameter ($\theta_{0,t}$) is a scaling factor and the second parameter ($\theta_{1,t}$) determines the slope of the pricing kernel at date t . When $\theta_{1,t}$ is positive, the

pricing kernel is negatively sloped, which implies that the value of a unit payoff increases as the underlying asset return decreases. The level of projected risk aversion is $\gamma_t^* = \theta_{1,t}$. If $\theta_{1,t}$ changes over time, then this specification exhibits time-varying risk aversion.

Our second specification permits more flexibility in the shape of the pricing kernel, and also allows time variation in risk aversion.² We consider a pricing kernel with $N + 1$ parameters $(\theta_{0,t}, \dots, \theta_{N,t})$ and $N + 1$ polynomial terms $(T_0(r_{t+1}), \dots, T_N(r_{t+1}))$:

$$M^*(r_{t+1}; \theta_t) = \theta_{0,t}T_0(r_{t+1}) + \theta_{1,t}T_1(r_{t+1}) + \theta_{2,t}T_2(r_{t+1}) + \dots + \theta_{N,t}T_N(r_{t+1}). \quad (11)$$

If there are an infinite number of terms in this polynomial expansion, then the specification will accurately approximate any continuous function. However, in the context of our estimation problem, the number of observed asset prices places an upper bound on the order of the approximating polynomial. Thus, we use a class of polynomials that provides the most accurate approximations with the smallest possible number of terms.

Orthogonal polynomials are designed to provide more precise approximations using lower order expansions than alternative classes of polynomials. In an orthogonal polynomial expansion, each term is mutually orthogonal to all other terms. The number of required terms is minimized, since each term provides unique information that is not contained in previous terms.

Although there are several families of orthogonal polynomials (e.g., Legendre, Chebyshev, Laguerre, Hermite), the Chebyshev family provides an approximation that comes close to minimizing the maximum approximation error.³ As shown by Judd (1998), Chebyshev polynomials are nearly optimal polynomial approximations under the L^∞ norm.

The Chebyshev polynomial is defined over the domain $[-1, 1]$ with terms given by $T_n(x) = \cos(n \cos^{-1}(x))$. The first and second Chebyshev terms are $T_0 = 1$ and $T_1 = x$. The higher order terms are periodic functions. To obtain an accurate approximation over a closed domain $[a, b]$, we use the generalized Chebyshev polynomial. In this polynomial, $x = ((2r_{t+1} - a - b)/(b - a))$, where a and b are the endpoints of the approximation interval.

We adopt the generalized Chebyshev polynomial approximation for our second pricing specification. To ensure that the pricing kernel is strictly positive, we take the exponential of the polynomial expansion such that

$$M^*(r_{t+1}; \theta_t) = [\theta_{0,t}T_0(r_{t+1})] \exp[\theta_{1,t}T_1(r_{t+1}) + \theta_{2,t}T_2(r_{t+1}) + \dots + \theta_{N,t}T_N(r_{t+1})]. \quad (12)$$

²The power specification is only nested in an orthogonal polynomial specification with an infinite number of terms. With a finite number of orthogonal polynomial terms, the power specification might be a more accurate representation of the pricing kernel.

³Chapman (1997) estimates the asset pricing kernel as a function of aggregate consumption using a five-term Legendre polynomial expansion. Bansal et al. (1993) estimate an international asset pricing kernel as a function of powers of the Eurodollar interest rate and a world equity index return.

3.3. State probability density specification

In the subsequent empirical portions of the paper (Sections 4–6), we use the equity index return probability density for pricing kernel estimation. So, in this section of the paper, we develop a stochastic volatility model that incorporates the most important features of equity index return process. Previous studies, such as Ghysels et al. (1996), document that equity index return volatility is stochastic and mean-reverting, return volatility responds asymmetrically to positive and negative returns, and return innovations are non-normal.

Researchers often capture stochastic volatility in a discrete-time setting, using extensions of the autoregressive conditional heteroskedasticity (ARCH) model proposed by Engle (1982). Comprehensive surveys of ARCH and related models are given by Bollerslev et al. (1992) as well as Bollerslev et al. (1994). In a continuous-time setting, researchers commonly use stochastic volatility diffusions. Surveys of this literature include Ghysels et al. (1996) and Shephard (1996).

Our model of the equity index return process uses an asymmetric GARCH specification with an empirical innovation density. The GARCH specification of Bollerslev (1986) incorporates stochastic, mean-reverting volatility dynamics. The asymmetry term in our model is based on Glosten et al. (1993). Our empirical innovation density captures potential non-normalities in the true innovation density.

The asymmetric GARCH model is specified as follows:

$$\ln(S_t/S_{t-1}) - rf = \mu + \varepsilon_t, \quad \varepsilon_t \sim f(0, \sigma_{t|t-1}^2) \tag{13}$$

and

$$\sigma_{t|t-1}^2 = \omega_1 + \omega_2 I + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1|t-2}^2 + \delta \text{Max}[0, -\varepsilon_{t-1}]^2. \tag{14}$$

In Eq. (13), the log-return net of the riskless rate of interest, $\ln(S_t/S_{t-1}) - rf$, has a constant mean (μ). Although a constant expected return is not usually compatible with time-varying risk aversion, the effect over a short period (e.g., up to one month) has a negligible effect on probability estimates. Therefore, Eq. (13) works as an approximation. We draw return innovations (ε_t) from an empirical density function (f) with stochastic variance ($\sigma_{t|t-1}^2$).

Eq. (14) defines conditional return variance ($\sigma_{t|t-1}^2$) as a function of two constants (ω_1 and ω_2), the lagged squared innovation (ε_{t-1}^2), and a non-linear function of the lagged return ($\text{Max}[0, -\varepsilon_{t-1}]^2$). The second constant (ω_2) permits a shift in long-run volatility using an indicator variable ($I = 0$ or 1) to mark the different time periods.

We estimate the model parameters using maximum likelihood with a normal innovation density. Bollerslev and Wooldridge (1992) show conditions that allow this technique to provide consistent parameter estimates even when the true innovation density is non-normal.

We model the empirical innovation density (f) by factoring the innovation density into time-varying and time-invariant components. To separate these components, we define a *standardized innovation* as the ratio of a return innovation (ε_t) and its conditional standard deviation ($\sigma_{t|t-1}$). The standardized innovation density—i.e.,

the set of standardized innovations—is the time-invariant component of empirical innovation density. The conditional standard deviation ($\sigma_{t|t-1}$) is the time-varying component of the empirical innovation density. On a particular date, we construct the empirical innovation density by multiplying each standardized innovation by the conditional standard deviation.

To estimate the standardized innovation density, we take the ratio of each return innovation and its conditional standard deviation using the estimated stochastic volatility model. This collection of estimated standardized innovations forms a density function that incorporates excess skewness, kurtosis, and other extreme return behavior that is not captured in a normal density.

After we estimate the stochastic volatility model, we use Monte Carlo simulation to determine the future return density over any desired time horizon. For example, we can create the one-period return density by simulating many one-period return realizations. We obtain a simulated one-period log-return ($\mu + rf + \varepsilon_{t+1}$) and a simulated one-period simple return [$\exp(\mu + rf + \varepsilon_{t+1}) - 1$] by randomly selecting an innovation (ε_{t+1}) from the empirical innovation density.

We can create a multi-period return density by simulating many multi-period return paths. We obtain a 20-period return by drawing the first return innovation (ε_{t+1}), updating the conditional variance ($\sigma_{t+2|t+1}$), drawing the second return innovation (ε_{t+2}), updating the conditional variance ($\sigma_{t+3|t+2}$), and continuing through the twentieth innovation. The one-period simulated log-return is equal to $\sum_{i=1, \dots, 20} (\mu + rf + \varepsilon_{t+i})$, and the one-period simulated simple return is equal to $\exp[\sum_{i=1, \dots, 20} (\mu + rf + \varepsilon_{t+i})]$.

3.4. Hedge ratio specification

We develop a hedge ratio estimation technique that does not depend on a particular specification of the state probability density or the pricing kernel. Our hedge ratios neutralize an option portfolio to the first- and second-order effects of changes in the underlying price. In a continuous-time diffusion setting, a first-order hedge will eliminate all randomness in the hedge portfolio and provide a minimum-variance hedge. In a discrete-time setting with stochastic volatility, first- and second-order hedges will reduce, but not eliminate hedge portfolio variability.

We derive these hedge ratios using a Taylor series expansion of the option pricing formula. The put option price change after one day ($Put_{t+1} - Put_t$) is approximately equal to the following function of the underlying price change ($S_{t+1} - S_t$):

$$Put_{t+1} - Put_t \cong \frac{\partial Put_{t+1}}{\partial S_{t+1}}(S_{t+1} - S_t) + \frac{1}{2} \frac{\partial^2 Put_{t+1}}{\partial S_{t+1}^2}(S_{t+1} - S_t)^2. \quad (15)$$

The first and second partial derivatives ($\partial Put_{t+1} / \partial S_{t+1}$, $\partial^2 Put_{t+1} / \partial S_{t+1}^2$) in Eq. (15) measure the sensitivity of the put price to first- and second-order changes in the underlying price. These price sensitivities are commonly called the option delta and the option gamma.

To form an option portfolio that is hedged against a first-order underlying price change, another security must be purchased that moves in the opposite manner from

the option. This hedging security has sensitivity to a first-order underlying price change equal to $-\partial Put_{t+1}/\partial S_{t+1}$. To further hedge against second-order effects of the underlying price change, a second security with second-order price sensitivity equal to $-\partial^2 Put_{t+1}/\partial S_{t+1}^2$ is required.

We refer to the number of units of the hedging security that are purchased to hedge a single option as the *hedge ratio*. To estimate hedge ratios in a setting with an arbitrary pricing kernel, we generalize the Engle and Rosenberg (1995) methodology for estimation of price sensitivities. Therefore, we consider three possible one-day underlying price changes. The stock price could rise by one standard deviation to $S_t + \varepsilon$, remain constant at S_t , or fall by one standard deviation to $S_t - \varepsilon$. Each underlying price change results in a different date $t + 1$ put option price: $Put_{t+1|S_t+\varepsilon}$, $Put_{t+1|S_t}$, or $Put_{t+1|S_t-\varepsilon}$.

We calculate approximations to the first and second partial derivatives of the option pricing formula using centered finite difference approximations:

$$\frac{\partial Put_{t+1}}{\partial S_{t+1}} \cong \frac{Put_{t+1|S_t+\varepsilon} - Put_{t+1|S_t-\varepsilon}}{2\varepsilon}$$

$$\frac{\partial^2 Put_{t+1}}{\partial S_{t+1}^2} \cong \frac{Put_{t+1|S_t+\varepsilon} - 2Put_{t+1|S_t} + Put_{t+1|S_t-\varepsilon}}{\varepsilon^2} \tag{16}$$

To evaluate these approximate derivatives, we find the value of the put option next period at different underlying price levels. We measure the current put price using the pricing kernel projected onto the T -period underlying asset return ($r_{t,t+T}$) and the put’s payoff function:

$$Put_t = E_t[M^*(r_{t,t+T}; \theta_t) \text{Max}[0, K - S_{t+T}]] \tag{17}$$

Next period, the underlying price is equal to S_{t+1} . The new underlying price affects next period’s put price through the payoff probability density and the pricing kernel. We write the conditional expectation using the updated payoff density as $E_{t+1|S_{t+1}}[\cdot]$.⁴ The parameters of the pricing kernel can also change when the underlying price changes, so we write next period’s parameter vector as $\theta_{t+1|S_{t+1}}$ such that

$$Put_{t+1|S_{t+1}} = E_{t+1|S_{t+1}}[M^*(r_{t+1,t+T}; \theta_{t+1|S_{t+1}}) \text{Max}[0, K - S_{t+T}]] \tag{18}$$

We then form an approximate pricing equation by replacing next period’s parameter vector with its conditional expectation in the current period such that

$$Put_{t+1|S_{t+1}} \cong E_{t+1|S_{t+1}}[M^*(r_{t+1,t+T}; E_t[\theta_{t+1}]) \text{Max}[0, K - S_{t+T}]] \tag{19}$$

We represent the pricing kernel parameter vector as a linear function of a constant vector (α), lagged parameter vectors ($\theta_{t-1}, \dots, \theta_{t-K}$), and an error vector (e_{t+1}):

$$\theta_{t+1} = \alpha + \beta_0\theta_t + \dots + \beta_K\theta_{t-K} + \dots + e_{t+1} \tag{20}$$

⁴If the underlying price follows a geometric Brownian motion, then the underlying price at date $t + 1$ (S_{t+1}) provides additional information about the expected price at date $t + T$ (S_{t+T}). If the underlying price follows an asymmetric GARCH process, then the underlying price at date $t + 1$ provides additional information about the expected price as well as higher moments of the price distribution.

Thus, we can evaluate the conditional expectation $E_t[\theta_{t+1}]$ by

$$E_t[\theta_{t+1}] = \alpha + \beta_0\theta_t + \dots + \beta_K\theta_{t-K}. \quad (21)$$

We use Monte Carlo simulation to estimate tomorrow's option price, conditional on tomorrow's underlying asset price, as

$$Put_{t+1|S_{t+1}} \cong J^{-1} \sum_{j=1}^J [M^*(r_{t+1,t+T,j}; [\alpha + \beta_0\theta_t + \dots + \beta_K\theta_{t-K}]) \text{Max}[0, K - S_{t+T,j}]]. \quad (22)$$

4. Data

We develop our options dataset from a subset of Berkeley Options Database covering the period 1991 through 1995. One of the advantages of this dataset is that option quotes are time-stamped and recorded along with the simultaneously measured underlying price making it easier to construct a database of time-synchronized daily option “closing” prices.

To create our database of option closing prices, we first collect “end-of-day” option prices for all contracts. We do so by averaging the last recorded bid–ask quote of the day between 2:00 and 3:00 PM Central Time. The cross-section of midquotes from the last hour of trading is not entirely synchronized, since the S&P 500 index level can change over the last hour of trading.

To correct for this effect, we calculate a Black and Scholes (1973) implied volatility each day for each option contract. We then find the closing price for each contract by evaluating the Black-Scholes formula. We use the same inputs and the implied volatility, except that the closing S&P 500 index level replaces the synchronized S&P 500 level. Finally, we average each call (put) price with the synthetic put (call) price that we determine using put–call parity adjusted for dividends. This average price is used in the estimation procedure.

This technique does not require the Black-Scholes model to be correct. It simply uses the Black-Scholes formula as an extrapolation device to calculate an option price adjustment when the S&P 500 level changes from the time of the last option quote to the close of option trading.

Our implied volatility calculation uses the end-of-day option midquote, the contemporaneous S&P 500 index level, the riskless interest rate, time until expiration (in trading days), and dividend yield. We measure the riskless interest rate using Datastream's bid and ask discount rates for U.S. Treasury Bills with maturities of one, three, and six months. The riskless rate for a particular option is calculated by linear interpolation of the interest rates of Treasury Bills that straddle the option expiration date. We calculate the dividend yield over the life of each option contract by taking the present value of future S&P 500 dividends and dividing by the current index level.

To eliminate data errors and ensure that closing option prices are representative of market conditions at the end of the trading day, we use several screening criteria. We

base some of these on Bakshi et al. (1997). In our sample, we include options with moneyness $-0.10 \leq (K/S_t - 1) \leq 0.10$, mid quotes greater than \$3/8 and less than \$50, annualized implied volatilities greater than 5% or less than 90%, and prices that satisfy the no-arbitrage lower bound ($P_t \geq \text{Max}[0, \exp[-rf(T - t)]K - S_t + D_{t,T}]$ or $C_t \geq \text{Max}[0, S_t - \exp[-rf(T - t)]K - D_{t,T}]$). We delete from the sample cross-sections of calls (puts) that violate the no-arbitrage condition that option premia are decreasing (increasing) in the exercise price, and options that violate the maximum vertical spread premium condition ($C_t(K_1) - C_t(K_2) \leq K_2 - K_1$; $P_t(K_2) - P_t(K_1) \leq K_2 - K_1$). Finally, we include in the sample only dates on which at least eight options (both calls and puts) satisfy the preceding criteria.

We use the following procedure to construct a dataset of options with one month (20 trading days) until expiration. We first eliminate all options with greater than 24 or fewer than 16 trading days until expiration. For each trading date, we choose the option series with time until expiration closest to 20 days. We are left with a single cross-section of call and put options each month (around the twentieth of the month) with a time-until-expiration of approximately one month. This sampling methodology is similar to that of Christensen and Prabhala (1998).

In Table 1, we report the properties of the one-month option contracts that we use for pricing kernel estimation. In the sample, there are 53 months (of 60 total) for which we have a cross-section of options that satisfies our screening criteria. In 39 months, there is an option series with exactly 20 days until expiration. We use a series with 21 days until expiration eight times, and the remaining six dates have option series with 18, 19, and 22 days until expiration. There is no satisfactory data in seven of the 60 months of the sample.

On a given estimation date, there are between 8 and 13 options available. There are roughly equal numbers of options with moneyness $(K/S_t - 1)$ from 3% to 0%, 0% to -3%, and -3% to -6%. There are somewhat fewer options with moneyness

Table 1
Summary of option data used for pricing kernel estimation

The source of our data is the Berkeley Options Database (1991–1995). Each month, we extract from the database a cross-section of options (both puts and calls) with approximately one month until expiration. We include options with moneyness between 10% and -10%. Additional data screening criteria are used to ensure that our option closing prices are representative of market conditions at the end of the trading day. There are 53 estimation dates in the sample, and the number of observations per estimation date ranges from eight to 13.

Option moneyness ($K/S_t - 1$) (%)	Number of observations	Average implied volatility (%)	Average call price	Average put price
-6 to 10	66	20.85	\$34.83	\$1.19
-3 to 6	133	16.58	\$21.81	\$1.92
0 to 3	136	13.67	\$10.76	\$4.13
0 to 3	137	11.36	\$3.25	\$9.77
3 to 6	37	11.61	\$0.90	\$16.66

between -6% and -10% . The smallest number of options has moneynesses ranging from 3% to 6% , and there are no options available with moneyness greater than 6% . In our sample, option contracts with higher moneyness generally have lower implied volatilities, a pattern known as a “volatility skew”.

Due to put–call parity, we gain no additional information if we include a call and a put with the same exercise price (or moneyness) in our estimation procedure. Therefore, we estimate pricing kernels using only out-of-the-money put options (moneyness $\leq 0\%$) and out-of-the-money call options (moneyness $> 0\%$).

Fig. 1 graphs five representative cross-sections of one-month option closing prices for June of 1991 through June of 1995 against percent moneyness $(K/S_t - 1)100$. The curves exhibit an inverted-V shape, since put premia increase in exercise price and call premia decrease in exercise price. The variation in the slope and height reflects differences in investor probability beliefs and risk aversion over time.

Table 2 reports summary statistics for the daily S&P 500 index returns series (1970 through 1995) used for estimation of the state probability model. Over this period, the average annualized S&P 500 index return (capital appreciation only) is 7.55% , and the annualized S&P 500 return standard deviation is 14.79% . S&P 500 returns exhibit negative skewness and positive kurtosis, and there is evidence of return serial correlation.

5. Estimation of the empirical pricing kernel projected onto S&P 500 return states

In this section, we estimate a monthly pricing kernel using a cross-section of S&P 500 index option prices and the S&P 500 return density function. We then analyze the relationship between empirical risk aversion and business conditions.

5.1. Estimation of S&P 500 return state probability densities

To find the most accurate model of the S&P 500 return density, we estimate and test three nested GARCH models: ARCH(1), GARCH(1, 1), and asymmetric GARCH(1, 1).⁵ We define the asymmetric GARCH (1,1) model using Eqs. (13) and (14). We define the other two models using the same equations, but set $\delta = 0$ for the GARCH(1, 1) model and both $\beta = 0$ and $\delta = 0$ for the ARCH(1) model. We use a likelihood ratio test to measure the statistical significance of the increase in likelihood for each model generalization. Table 3 reports the model estimates.

In Table 3, we find that the GARCH model offers a statistically significant improvement over the ARCH model with a likelihood ratio test p -value less than 0.0001 . Our tests also show that the asymmetric GARCH model provides a better fit than the GARCH model. We see that the robust t -statistic for the volatility asymmetry parameter (δ) is 2.41 , which confirms the presence of an asymmetric

⁵We also estimate the GARCH components model of Engle and Lee (1999) but do not find that it is a statistically significant improvement over the GARCH(1, 1) model. To conserve space, we do not report these results.

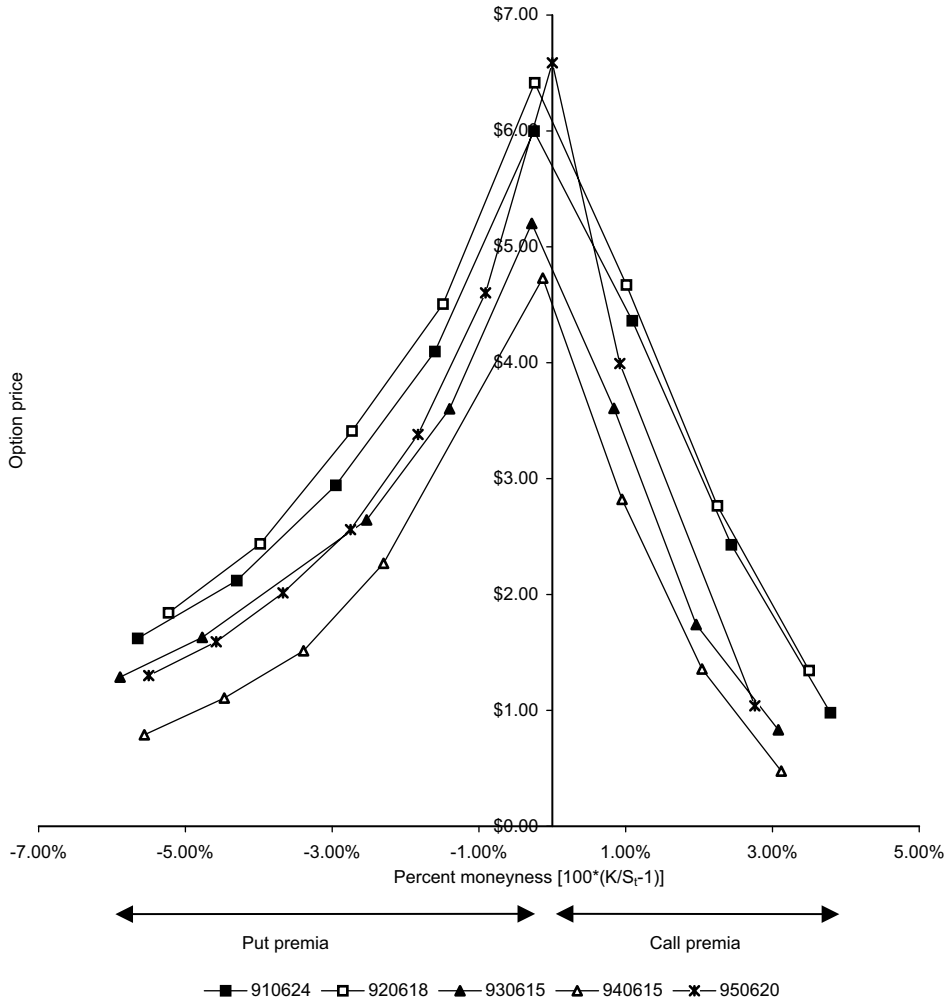


Fig. 1. One-month S&P 500 index option prices. This figure graphs the cross-section of one-month S&P 500 option prices from June 1991 to 1995 against percent moneyness (option exercise price/closing S&P 500 index level - 1) 100. For positive (negative) percent moneyness, we report call (put) premia.

volatility effect. We set the indicator variable equal to one during the EPK estimation period (1991 through 1995) and equal to zero for the rest of the sample. This variable is not significant in the GARCH or asymmetric GARCH models.

We create the standardized innovations $(\varepsilon_t/\sigma_{t|t-1})$ for the asymmetric GARCH model by taking the ratio of each return innovation and its conditional standard deviation. If we have correctly specified the stochastic volatility model, then the standardized innovations will be free of time dependence. We perform specification tests on the standardized innovations to measure autocorrelation in the standardized

Table 2

S&P 500 daily log-return summary statistics, 1970–1995

The normality test p -value is the p -value of the Jarque and Bera (1980) normality test statistic, which measures the closeness of the empirical S&P 500 log-return density to a normal density. The serial correlation p -value is the p -value of the Ljung-Box Q -statistic (1978), which measures serial correlation in the innovations using ten lagged values. The ARCH test p -value is the p -value of the Engle (1982) ARCH LM statistic, which measures the presence of stochastic volatility as represented by persistence in return magnitudes. We use ten return lags for this test.

Number of observations	6571
Annualized mean	7.55%
Annualized std. dev.	14.79%
Skewness	-2.31
Excess kurtosis	60.20
Normality test p -value	<0.0001
Serial correlation test p -value	<0.0001
ARCH test p -value	<0.0001

Table 3

Estimated state probability models

We estimate an ARCH(1), GARCH(1,1), and asymmetric GARCH(1,1) model by maximizing the likelihood function of daily S&P 500 log-returns from 1970 to 1995. We calculate robust t -statistics using the Bollerslev and Wooldridge (1992) method. We define the asymmetric GARCH model as

$$\ln(S_t/S_{t-1}) - rf = \mu - \varepsilon_t, \quad \varepsilon_t \sim f(0, \sigma_{t|t-1}^2), \quad \sigma_{t|t-1}^2 = \omega_1 + \omega_2 I + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1|t-2}^2 + \delta \text{Max}[0, -\varepsilon_{t-1}]^2.$$

The ARCH p -value is the p -value of the Engle (1982) ARCH LM statistic using ten lags. The LR (likelihood ratio) test p -value measures the statistical significance of the improvement in expanding the model specification. The reported p -value is the p -value for twice the difference between the log-likelihood of the unrestricted and restricted models.

	ARCH(1)		GARCH(1,1)		Asymmetric GARCH(1,1)	
	Coefficient	Robust t -statistic	Coefficient	Robust t -statistic	Coefficient	Robust t -statistic
μ	0.0004	3.09	0.0004	4.10	0.0003	2.41
θ	0.1483	6.71	0.1363	10.62	0.1361	10.60
ω_1	7.06E - 05	17.85	1.22E - 06	3.45	1.26E - 06	3.80
ω_2	-3.41E - 05	-8.52	-4.64E - 07	-1.77	-4.24E - 07	-1.69
α	0.1964	2.86	0.0663	2.41	0.0262	3.12
β			0.9212	35.88	0.9264	53.83
δ					0.0672	2.41
Log-likelihood	21,941.39		22,332.09		22,368.10	
ARCH p -value	<0.0001		0.3426		0.7004	
LR test p -value	N/A		<0.0001		<0.0001	

innovations and in the squared standardized innovations. In Table 4, Panel A, we report results of these specification tests.

We use the Ljung and Box (1978) *Q*-statistic to test for autocorrelation in the standardized innovations. The asymmetric GARCH model passes this test with a *p*-value of 0.7378. Next, we use Engle (1982) ARCH LM test to test for autocorrelation in the squared standardized innovations. The asymmetric GARCH model passes this test with a *p*-value of 0.7004. Since the asymmetric GARCH model provides the best fit to the return data and passes the specification tests, we choose this model for state probability estimation.

We then estimate standardized innovation density using the collection of the standardized innovations from asymmetric GARCH model. Table 4, Panel A shows that this density exhibits negative skewness (−0.36) and positive excess kurtosis (4.26) compared to a normal density.

In Table 4, Panel B, we compare extreme return probabilities using the standardized innovation density and a standard normal density. Under the standard normal assumption, innovations of magnitude greater than five or ten standard deviations almost never occur (less than 1:1,000,000). In practice, the probability of these extreme events is non-negligible. For example, empirical return innovations less than −5 standard deviations are observed six times in 10,000, while empirical return innovations less than −10 standard deviations are observed three times in 10,000.

Table 4

Properties of the standardized innovations using the estimated asymmetric GARCH model

We calculate the standardized innovations by dividing each ordinary innovation (ϵ_t) by its conditional standard deviation ($\sigma_{t|t-1}$). Both variables are taken from the estimated asymmetric GARCH model. In Panel A, we use Engle (1982) ARCH test statistic to measure unexplained stochastic volatility, and we use the Ljung and Box (1978) test statistic to measure unexplained serial correlation. We compare probabilities of extreme innovations for the standardized innovation density and a standard normal density in Panel B.

Panel A. Summary statistics

Number of observations	Mean	Std. dev.	Skewness	Excess kurtosis	Normality test <i>p</i> -value	ARCH test <i>p</i> -value	Serial correlation test <i>p</i> -value
6,571	0.0004	1.0000	−0.36	4.26	<.0001	0.7004	0.7378

Panel B. Probabilities of extreme events

Probability	Empirical density	Standard normal
Innovation < − 10 std. dev.	0.00030	<0.000001
Innovation < − 5 std. dev.	0.00061	<0.000001
Innovation < − 3 std. dev.	0.00427	0.00135
Innovation > 3 std. dev.	0.00290	0.00135
Innovation > 5 std. dev.	0.00015	<0.000001
Innovation > 10 std. dev.	<0.000001	<0.000001

Fig. 2 uses 200,000 Monte Carlo simulation replications to graph the estimated one-month state probability densities each June. The time variation in state probabilities is apparent. As we can see, there are higher probabilities for large negative return states in June 1991 and June 1992 than in June of subsequent years. The annualized return standard deviation estimates (in chronological order for June 1991 through June 1995) are: 15.55%, 13.03%, 11.74%, 9.98%, and 10.88%. The return density skewness estimates are -0.36 , -0.43 , -0.46 , -0.46 , and -0.47 . The return density kurtosis estimates are 5.06, 5.35, 5.66, 5.54, and 5.73.

Fig. 3 graphs the conditional volatility forecasts using the asymmetric GARCH model. Over this period, the estimated annualized S&P 500 volatility ranges from 6.75% to 112.86% with a standard deviation of 6.36%. The highest volatility forecasts over this period are around the time of the October 1987 market crash.

5.2. Estimation of the S&P 500 empirical pricing kernel

For estimation of the S&P 500 empirical pricing kernel, we use a power specification and a four-parameter orthogonal polynomial specification. We set the return domain for orthogonal polynomial equal to the range of option moneyness (-10 – 10%). Outside of this domain, we set the pricing kernel equal to its estimated value at -10% or 10% .

Once per month, we identify the pricing kernel that best fits the cross-section of one-month S&P 500 option premia using each specification. In the optimization

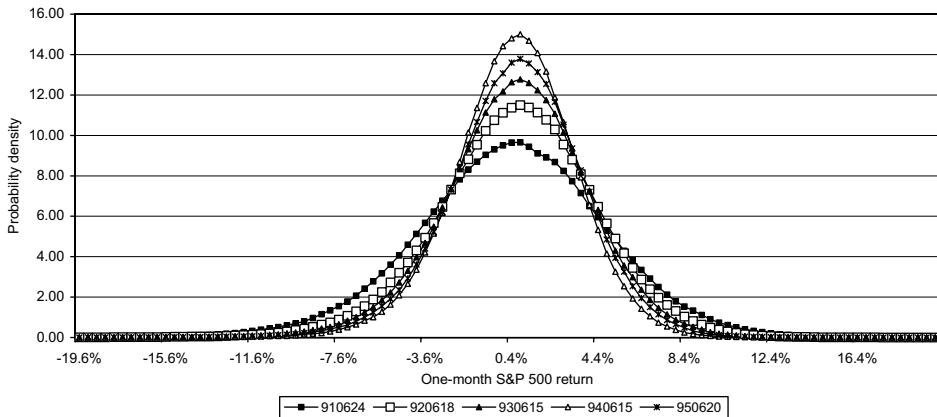


Fig. 2. Estimated S&P 500 return density functions in mid-June 1991–1995. We estimate the density using Monte Carlo simulation (200,000 replications) of the estimated asymmetric GARCH model with an empirical innovation density. We obtain density functions using a Gaussian kernel density estimator with bandwidth equal to $0.9N^{-1/5}\sigma$, where σ is the return standard deviation and N is the number of replications.

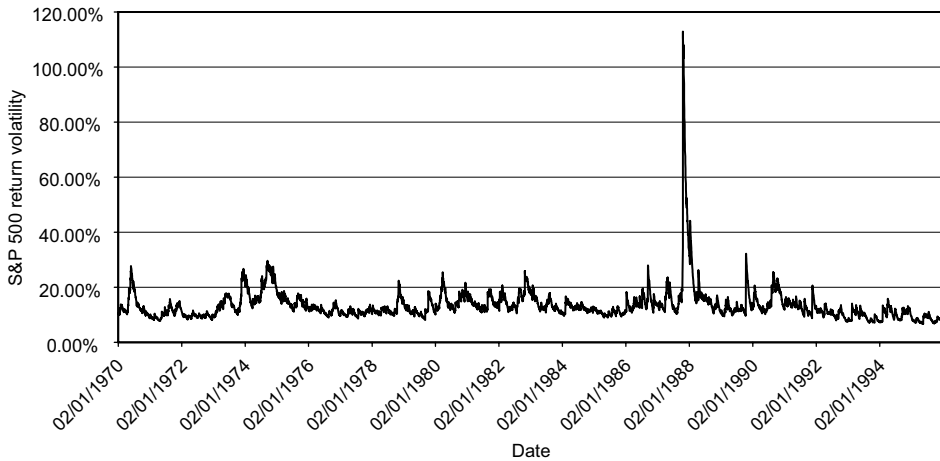


Fig. 3. Annualized S&P 500 conditional volatility, using the asymmetric GARCH model over the period 1970–1995. Dates (x-axis) follow the month/day/year format.

procedure, we use the estimated asymmetric GARCH model to create a simulated one-month probability density with 200,000 replications. To ensure that the estimated pricing kernel accurately prices a riskless one-month bond, we also set the scaling factor $(\theta_{0,t})$ to satisfy the pricing equation $B_t = E_t[M^*(r_{t+1}; \theta_t)]$.

Table 5 reports the estimation results. The orthogonal polynomial specification fits S&P 500 option prices more closely than the power specification. The average pricing error standard deviation for the orthogonal polynomial specification is \$0.09 with a minimum of \$0.03 and a maximum of \$0.24. The average forecast error standard deviation for the power specification is \$0.63 with a minimum of \$0.28 and a maximum of \$1.34.

In Fig. 4, we graph power pricing kernel estimates each June of the sample period. We find that the level of risk aversion varies across time, as illustrated by the changing slope of the estimated pricing kernels. The negatively sloped pricing kernel estimates show that investors experience declining marginal utility over S&P 500 return states.

Fig. 5 graphs orthogonal polynomial pricing kernel estimates. Compared to the power pricing kernels, the orthogonal polynomial pricing kernels assign greater value to large negative S&P 500 return states and lesser value to large positive S&P 500 return states. The state-price-per-unit probability for large negative return states is especially volatile, and could reflect time-varying demand for insurance against a significant market decline. There is also some evidence of a region of increasing marginal utility for small positive S&P 500 return states.

We then estimate an average power pricing kernel and an average orthogonal polynomial kernel by evaluating each specification at the average parameter estimates. Table 5 reports characteristics of the parameter estimates (including their

Table 5

Empirical pricing kernel estimation results

We estimate one-month pricing kernels projected onto S&P 500 return states over the period 1991–1995 using two specifications. We report results for the power specification, $M_t(r_{t+1}) = \theta_{0,t}(r_{t+1})^{-\theta_{1,t}}$, in Panel A. We report results for the orthogonal polynomial specification, $M_t(r_{t+1}) = \theta_{0,t}T_0(r_{t+1}) \exp[\theta_{1,t}T_1(r_{t+1}) + \theta_{2,t}T_2(r_{t+1}) + \theta_{3,t}T_3(r_{t+1})]$, in Panel B.

$N = 53$	Mean	Standard deviation	Min	Max
<i>Panel A. Power specification parameters</i>				
$\theta_{0,t}$	1.0051	0.0063	0.9866	1.0185
$\theta_{1,t}$	7.36	2.58	2.36	12.55
Std. dev. of pricing errors	\$0.63	\$0.26	\$0.28	\$1.34
<i>Panel B. Orthogonal polynomial specification parameters</i>				
$\theta_{0,t}$	0.19	0.10	0.04	0.40
$\theta_{1,t}$	-2.25	1.06	-4.38	-0.25
$\theta_{2,t}$	-0.88	0.68	-2.52	0.19
$\theta_{3,t}$	-1.08	0.42	-1.94	-0.19
Std. dev. of pricing errors	\$0.09	\$0.05	\$0.03	\$0.24

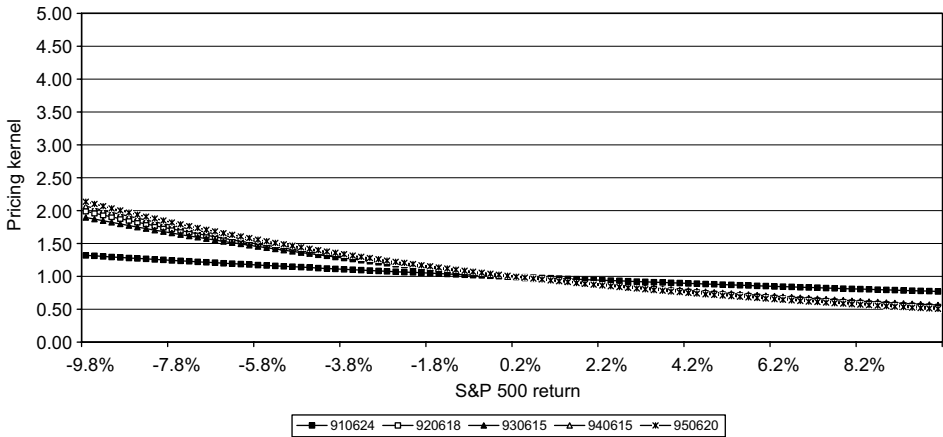


Fig. 4. Empirical pricing kernel using power specification in mid-June 1991–1995. We define this pricing kernel as $M_t(r_{t+1}) = \theta_{0,t}(r_{t+1})^{-\theta_{1,t}}$. The time-varying slope of the pricing kernel estimates reflects fluctuations in empirical risk aversion.

means), and Fig. 6 graphs the average pricing kernels. The average orthogonal polynomial EPK has some similarities to the estimate of Ait-Sahalia and Lo (2000). Both pricing kernels are steeply upward sloping for large negative returns and downward sloping for large positive returns, and both pricing kernels have a region of increasing marginal utility.

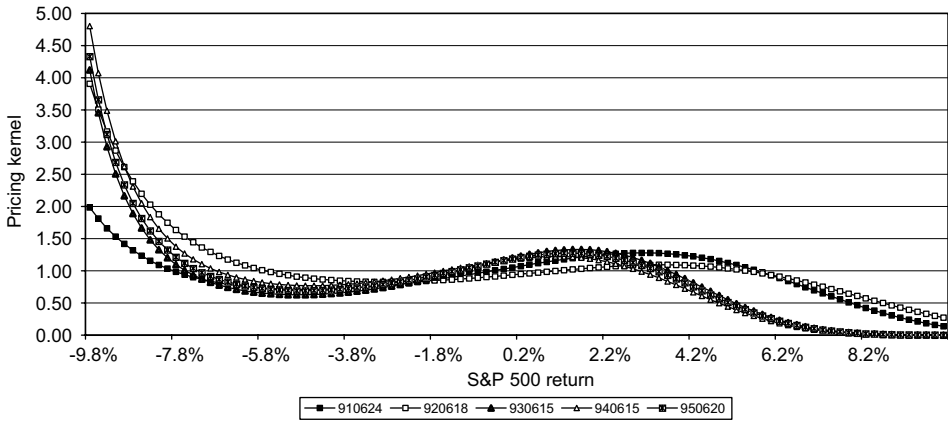


Fig. 5. Empirical pricing kernel using orthogonal polynomial specification in mid-June 1991–1995. We use a four-term expansion: $M_t(r_{t+1}) = \theta_{0,t}T_0(r_{t+1})\exp[\theta_{1,t}T_1(r_{t+1}) + \theta_{2,t}T_2(r_{t+1}) + \theta_{3,t}T_3(r_{t+1})]$. The pricing kernel estimates using this specification assign greater weight to large negative S&P 500 return states and lesser weight to large positive S&P 500 return states than do the estimates using the power specification.

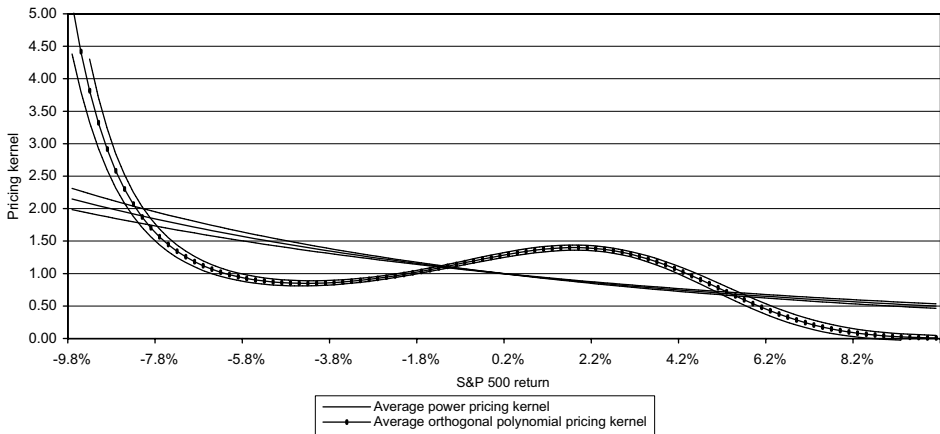


Fig. 6. Time-invariant (average) pricing kernels in mid-June 1991–1995 using the power and orthogonal polynomial specifications. We obtain each average pricing kernel by evaluating the specification at the average parameter estimates for the period from 1991 to 1995. The figure also shows ± 2 standard error bounds for the pricing kernels.

Jackwerth’s absolute risk-aversion function (Jackwerth, 2000) is closely related to the pricing kernel and can be expressed as the negative of the ratio of the first derivative of the pricing kernel and the pricing kernel ($-M'_t(r_{t+1})/M_t(r_{t+1})$). Jackwerth notes two key empirical findings for the absolute risk aversion functions: “... post-crash risk aversion functions are negative around the center

[return states close to zero]...[and] risk aversion functions rise for wealth levels greater than about 0.99 [return states greater than -1%].”

We do not find either of these characteristics for our estimates of the power pricing kernel. In this specification, the absolute risk aversion function equal to the exponent of the power function multiplied by the inverse of the gross return $[\theta_{1,t}^*(r_{t+1})^{-1}]$. Thus, the power pricing kernel exhibits declining (but positive) absolute risk aversion as long as the exponent ($\theta_{1,t}$) is positive. This is the result that we find over the period 1991 through 1995.

Our estimates of the orthogonal polynomial pricing kernel exhibit some of the risk-aversion characteristics noted by Jackwerth (2000). We estimate the average absolute risk-aversion function using the average orthogonal polynomial pricing kernel graphed in Fig. 6. We find that there is a region of negative absolute risk aversion over the range from -4% to 2% and that absolute risk aversion increases for returns greater than -4% . The shape of our estimated average absolute risk aversion function is similar to Jackwerth’s estimate over a similar time period.

5.3. Linking empirical risk aversion to business conditions

We use the risk aversion of the estimated power pricing kernel ($\theta_{1,t}$) as our measure of empirical risk aversion. We analyze the time series of $\theta_{1,t}$ estimates to gain insight into risk aversion dynamics and links between risk aversion and the business cycle.

Table 6 provides summary statistics for empirical risk aversion. Over the sample period, empirical risk aversion averages 7.36. However, the level fluctuates substantially, ranging from 2.26 to 12.55. Empirical risk aversion is positively autocorrelated ($\rho = 0.45$) and mean-reverting. Fig. 7 graphs the time series of empirical risk aversion estimates.

Other studies, such as Fama and French (1989), show that risk premia are correlated with the business cycle. Specifically, risk premia are lowest at business cycle peaks and highest at business cycle troughs. We provide evidence of time-varying risk aversion through the business cycle, which supports the Fama and French (1989) results.

Table 6

Summary statistics for empirical risk aversion

We estimate empirical risk aversion over the period 1991–1995, using the exponent of the estimated power pricing kernel. There are seven months for which there is insufficient data for estimation of the EPK, leaving 53 observations of empirical risk aversion.

Number of observations	53
Mean	7.36
Std. dev.	2.58
Min	2.36
Max	12.55
First-order autocorrelation	0.45

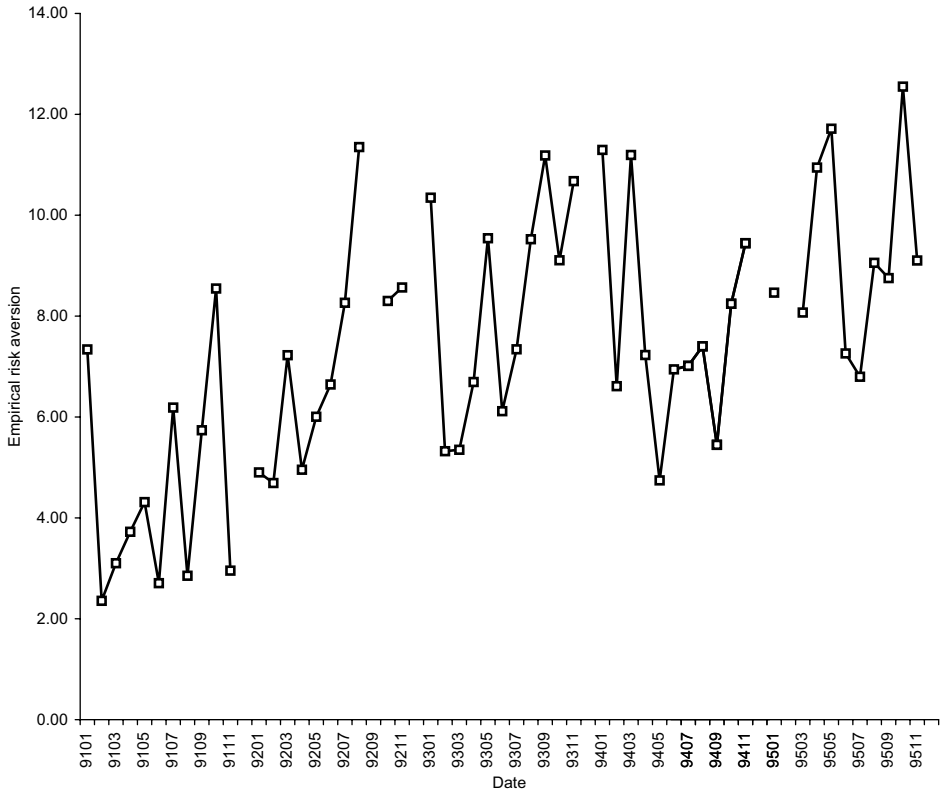


Fig. 7. Empirical risk aversion each month from 1991 through 1995. Empirical risk aversion is the exponent of the estimated power pricing kernel ($\theta_{1,t}$). The seven missing monthly observations are indicated by absence of a data point as well as by a gap in the connecting curve.

To measure the relation between empirical risk aversion and the business cycle, we construct several variables that reflect current and expected business conditions. We calculate one-month percentage changes in business condition indicators by using the indicator measured on the current and previous pricing kernel estimation date.

Fama and French (1989) and Lahiri and Wang (1996) use credit spreads (the difference between the yield on risky and riskless bonds) as an indicator of business conditions. As the economy weakens (strengthens), credit spreads widen (narrow) to compensate investors for an increased probability of default.

We calculate credit spreads using a risky bond yield equal to Moody’s long-term Baa corporate bond yield index. This index measures the average yield-to-maturity of approximately one hundred seasoned corporate bonds with maturities as close as possible to 30 years and at least 20 years. We set the riskless bond yield equal to the Federal Reserve’s 30-year constant maturity Treasury yield. We collect both data items from the Federal Reserve’s H.15 release.

Estrella and Hardouvelis (1991) show that the slope of the yield curve (term spread) is pro cyclical. Steepening of the slope indicates expansion, while flattening of the slope indicates contraction. We use the H.15 release information to measure the yield curve slope as the 30-year constant maturity Treasury yield minus the three-month constant maturity Treasury yield.

Estrella and Hardouvelis (1991) also suggest that the level of short-term interest rates might reveal the state of the business cycle. They summarize the view that high short-term rates are associated with a tight monetary policy, low current investment opportunities, and low output. They present empirical evidence that demonstrates this relation. We use the percentage change in the three-month constant maturity Treasury yield as proxy for this indicator.

In our analysis, we use the one-month percentage change in the S&P 500 level as reported in the CRSP database as a potential business cycle indicator. We also include the aggregate U.S. consumption growth rate, since it is used in many studies of the stochastic discount factor such as Hansen and Singleton (1982). We measure consumption growth using per-capita non-durable goods and services (monthly, real, seasonally adjusted) from the Federal Reserve's FRED database. Our monthly U.S. resident population estimates are from the Census Bureau.

To measure autocorrelation in risk aversion, we include the one-month lag of empirical risk aversion. In addition, we use the difference between at-the-money implied and objective volatility as a proxy for risk aversion. This volatility spread measures the mark-up of an at-the-money option price above the price that a risk-neutral investor would accept.

Table 7 reports a multiple regression of empirical risk aversion on all of the above variables. We present univariate correlations and their p -values in the last two columns. We measure independent variables in percent (except lagged empirical risk aversion). Therefore, a one basis point (0.01) increase in the independent variable raises empirical risk aversion by one one-hundredth of the regression coefficient.

The Table 7 results show that empirical risk aversion varies counter cyclically with business conditions. One business cycle indicator variable is significant in the multiple regression (credit spread), and two business cycle variables have significant correlation with empirical risk aversion (credit spread and term structure slope).⁶ The signs of the regression parameter estimates and correlation coefficient estimates are all consistent with counter cyclical risk aversion.

For example, in the multiple regression, the credit spread is statistically significant and has a positive estimate of 9.95. Hence, a one-basis point widening of the credit spread increases empirical risk aversion by 0.0995. The correlation between empirical risk aversion and credit spreads is also positive and statistically significant ($\rho = 0.50$, p - value = 0.0004) in the univariate analysis. Since the credit spread is a

⁶Two other variables, lagged risk aversion and volatility spread, are significant in the regression and correlation analysis. The positive coefficient on lagged risk aversion shows that risk aversion is positively auto correlated. The positive coefficient on the volatility spread shows that an increase in empirical risk aversion is associated with a widening of the spread between implied and objective volatility.

Table 7

Analysis of empirical risk aversion

We report estimates from a multiple regression of empirical risk aversion on business cycle indicators. We also show univariate correlations of empirical risk aversion with business cycle indicators. The seven months in which one-month lagged empirical risk aversion is missing are dropped from the regression, leaving 46 observations. We measure independent variables on the same date on which we calculate empirical risk aversion. We also show OLS *t*-statistics and *p*-values.

Variable	Regression coefficient	<i>t</i> -Statistic	<i>p</i> -Value	Univariate correlation	Correlation <i>p</i> -value
Intercept	3.99	3.88	0.0004		
One-month lagged empirical risk aversion	0.48	3.54	0.0011	0.45	0.0018
S&P 500% return	0.04	0.29	0.7732	0.03	0.8555
Three-month Treasury yield change	0.68	0.37	0.7130	0.24	0.1058
Credit spread change	9.95	2.12	0.0405	0.50	0.0004
Term structure slope change	0.23	0.12	0.9035	-0.36	0.0129
Implied vol. – objective vol. change	0.52	3.36	0.0018	0.39	0.0069
Consumption growth change	-0.04	-0.72	0.4789	-0.10	0.5224
Adjusted R^2	46.10%				

counter cyclical indicator, these results show that empirical risk aversion is counter-cyclical.

The slope of the term structure is not statistically significant in the multiple regression, but it has a statistically significant negative correlation with risk aversion ($\rho = -0.36$, p -value = 0.0129). This result is further evidence of counter cyclical risk aversion, since risk aversion is negatively correlated with a pro cyclical indicator.

Our empirical findings provide some support for habit persistence models of investor utility. Habit models such as Campbell (1996) and Campbell and Cochrane (1999) predict that Arrow-Pratt relative risk aversion is counter cyclical. In an economic expansion, the surplus consumption ratio (the proportion that current consumption exceeds the habit) is high and risk aversion is low. In a recession, the reverse is true.

If our measure of empirical risk aversion is positively correlated with Arrow-Pratt relative risk aversion, then our finding of counter cyclical empirical risk aversion shows that Arrow-Pratt relative risk aversion might be counter cyclical. In this setting, our finding would provide empirical support for a key implication of habit models.

6. Hedging tests

In this section, we use hedging performance to measure the importance of time variation in the pricing kernel and to compare the accuracy of the power and orthogonal polynomial specifications.

To implement our tests, we create hedge portfolios for a \$100 position in out-of-the-money (OTM) S&P 500 index options. We estimate hedge ratios using time-invariant and time-varying pricing kernels with power and orthogonal polynomial specifications. We use at-the-money (ATM) put options and/or the S&P 500 index portfolio as hedging instruments.

We construct a hedging sample using the same screening criteria as we did for estimation of the pricing kernels. However, the hedging sample comprises options with approximately one month (from 16 to 24 days) until expiration, instead of exclusively one-month options.

On each sample date, we select a put with moneyness $(K/S_t - 1)$ closest to zero but no more than 1% in absolute value. This is the at-the-money put. As the out-of-the-money put, we select the option with moneyness closest to -3% , but no greater than -3% . When we cannot find suitable options with closing prices on the sample date and the next trading date, we exclude the sample date from the analysis. Using these criteria, we have 243 observations available for the hedging tests. Table 8 summarizes the sample characteristics.

To form one-day ahead option prices that are the basis for the hedge ratio estimates, we require estimates of the expected one-day ahead pricing kernel. For the time-invariant pricing kernels, the expected pricing kernel parameter vector is equal to the average parameter vector using the 53 observations from EPK estimation. We show the average parameter estimates in Table 9.

Table 8

Summary of option data used for hedging tests

Our hedging sample comprises at-the-money (ATM) and out-of-the-money (OTM) options with approximately one month (from 16 to 24 days) until expiration from the Berkeley Options Database (1991–1995). On each sample date, the ATM put is the closest-to-the-money put (within 1% of the money). The OTM put is the put with moneyness closest to -3% , but no greater than -3% . When we cannot find suitable options with reported prices on the sample date and next trading date, we exclude the sample date from the analysis. We measure the option time-until-expiration in trading days.

	ATM put option	OTM put option
Number of observations	243	243
Average price	\$6.34	\$2.39
Std. dev. price	\$1.04	\$0.77
Average price change	-\$0.11	-\$0.10
Std. dev. price change	\$1.38	\$0.56
Average time to maturity (days)	19.88	19.88
Std. dev. time to maturity (days)	2.40	2.40
Average moneyness $(K/S_t - 1)$	0.00%	-3.61%
Std. dev. moneyness $(K/S_t - 1)$	0.32%	0.48%
Average implied volatility	12.90%	16.36%
Std. dev. of implied volatility	2.79%	2.76%
Minimum implied volatility	9.04%	11.69%
Maximum implied volatility	30.96%	34.39%

Table 9

Time-invariant models of the pricing kernel parameter vector

We estimate the time-invariant power model by averaging the scaling factor ($\theta_{0,t}$) and the risk-aversion coefficient ($\theta_{1,t}$) using the 53 monthly estimates from the EPK estimation. We estimate the time-invariant orthogonal polynomial model by averaging the scaling factor ($\theta_{0,t}$) and generalized Chebyshev coefficients ($\theta_{1,t}, \theta_{2,t}, \theta_{3,t}$) using the 53 monthly estimates from the EPK estimation.

$N = 53$	Estimate
<i>Panel A. Time-invariant power specification</i>	
$\theta_{0,t}$	1.0051
$\theta_{1,t}$	7.3629
<i>Panel B. Time-invariant orthogonal polynomial specification</i>	
$\theta_{0,t}$	0.1945
$\theta_{1,t}$	-2.2457
$\theta_{2,t}$	-0.8791
$\theta_{3,t}$	-1.0752

For the empirical pricing kernel, we estimate a model in which tomorrow's parameter vector is a linear function of today's parameter vector (Eq. (20) with one lag). We use the hedging sample for estimation, but we drop sequential observations more than one day apart, leaving 133 observations. Table 10 shows that this model predicts reasonably well, with adjusted R^2 that range from 12% to 52%.

Panel A, Table 10, reports the $\theta_{1,t+1}$ model estimates for the power pricing kernel specification. In this model, $\theta_{1,t+1}$ is forecast by the sum of 2.26 and $0.69\theta_{1,t}$. Both parameters (α, β_1) are statistically significant, and the regression adjusted R^2 is 50%. This model generates an average one-day ahead forecast of $\theta_{1,t+1}$ equal to 7.49 with a standard deviation of 1.87.

We form hedge one-day ratios for S&P 500 put options by measuring the impact of a one-day change in the S&P 500 index on the put price. We create three possible realizations for the next day's S&P 500 level and calculate three corresponding put prices. These put prices depend on the expected pricing kernel and the payoff density function (estimated by 200,000 Monte Carlo replications using the asymmetric GARCH model).

For each pricing kernel specification, we use the corresponding hedge ratios to calculate the time series of hedging errors and the standard deviation of hedge portfolio prices. We also compare the relative hedging performance using a statistic similar to Diebold and Mariano (1995). Our relative performance measurement is equal to the t -statistic for the difference between squared hedging errors for each model. We calculate standard errors using the heteroskedasticity and autocorrelation consistent covariance matrix of Newey and West (1987).

In Table 11, we report the hedging test results. We find strong evidence that the pricing kernel is time-varying. When we use a time-varying pricing kernel, we improve hedging performance from 1% to 3% over a time-invariant pricing kernel.

Table 10

Dynamic models of the pricing kernel parameter vector

We estimate dynamic models of the pricing kernel parameter vector using a linear regression of pricing kernel parameter estimates on one-day lagged parameter estimates. We construct a sample of estimated pricing kernels (power and orthogonal polynomial) on each of the 243 hedging sample dates. We drop from the sample the hedging dates that are not one day apart. We use the remaining 133 observations in the forecasting model estimation regressions. We obtain parameter forecasts each of the 243 hedging dates using the estimated models.

Panel A. EPK power specification

$\theta_{0,t+1} = \alpha + \beta_0\theta_{0,t} + \beta_1\theta_{1,t} + e_{0,t+1}$ $\theta_{1,t+1} = \alpha + \beta_1\theta_{1,t} + e_{1,t+1}$				Adj- R^2 (%)	Average one-day parameter forecasts	Std. dev. one-day parameter forecasts
	α	β_0	β_1			
$\theta_{0,t+1}$	0.8163	0.1885	0.0000	12.25	1.0053	0.0030
p -val.	<0.0001	<0.0001	0.8632			
$\theta_{1,t+1}$	2.2608		0.6947	49.92	7.4872	1.8717
p -val.	<0.0001		<0.0001			

Panel B. EPK orthogonal polynomial specification

$\theta_{0,t+1} = \alpha + \beta_0\theta_{0,t} + \beta_1\theta_{1,t} + e_{0,t+1}$ $\theta_{1,t+1} = \alpha + \beta_1\theta_{1,t} + \beta_2\theta_{2,t} + \beta_3\theta_{3,t} + e_{1,t+1}$ $\theta_{2,t+1} = \alpha + \beta_1\theta_{1,t} + \beta_2\theta_{2,t} + \beta_3\theta_{3,t} + e_{2,t+1}$ $\theta_{3,t+1} = \alpha + \beta_1\theta_{1,t} + \beta_2\theta_{2,t} + \beta_3\theta_{3,t} + e_{3,t+1}$						Adj- R^2 (%)	Average one-day parameter forecast	Std. dev. one-day parameter forecasts
	α	β_0	β_1	β_2	β_3			
$\theta_{0,t+1}$	0.1756	0.4321	0.0290	-0.0075	-0.0065	39.45	0.2210	0.0698
p -val.	0.0739	0.0996	0.1568	0.8530	0.8961			
$\theta_{1,t+1}$	0.1845		0.8545	-0.4116	0.2034	51.72	-2.0464	0.7858
p -val.	0.0429		<0.0001	0.0556	0.6154			
$\theta_{2,t+1}$	-0.1454		0.1384	0.3496	0.0240	33.60	-0.7196	0.4056
p -val.	0.2665		0.2823	0.0217	0.9330			
$\theta_{2,t+1}$	-0.2535		0.1455	-0.1492	0.5352	45.67	-0.9768	0.3011
p -val.	0.0024		0.0729	0.1166	0.0033			

The hedging improvement is statistically significant in three of six cases and marginally significant in one case.

When the ATM put is used as a hedging instrument, the hedge portfolio formed using the EPK power specification has a standard deviation of \$11.10 per day. This is the least volatile hedge portfolio. The hedge portfolio formed using time-invariant power pricing kernel has a standard deviation of \$11.21. The t -statistic for the hedging performance difference is 2.82.

Compared to hedging using the ATM put, hedging using the S&P 500 portfolio somewhat diminishes performance for all the specifications. Once again, the EPK power specification creates the most effective hedge ratios. Its hedge portfolio standard deviation is \$12.11, while the hedge based on the time-invariant power specification has a standard deviation of \$12.41.

Table 11

Hedging test results

We conduct hedging tests to compare the (time-varying) EPK power and orthogonal polynomial pricing kernel specifications with the time-invariant (average) power and orthogonal polynomial specifications. The tests cover the period 1991–1995 (243 sample dates). We form hedges of a \$100 position in one-month out-of-the-money (OTM) S&P 500 index put options using one-month at-the-money (ATM) put options, the S&P 500 index portfolio, or both. We construct hedge ratios to neutralize the hedge portfolio sensitivity to the first-order (and in one case, second order) effects of underlying price changes.

Portfolios—pricing kernel specification	Hedge portfolio standard deviation	Reduction in standard deviation (EPK versus time-invariant specification) (%)	Robust <i>t</i> -statistic (EPK versus time-invariant specification)
<i>No hedge</i>			
\$100 OTM written put position	\$22.56		
<i>Hedge using underlying</i>			
Time-invariant power	\$12.41		
EPK power	\$12.11	2.39	1.16
Time-invariant orthog. poly.	\$13.45		
EPK orthogonal polynomial	\$13.13	2.36	1.95
<i>Hedge using ATM put</i>			
Time-invariant power	\$11.21		
EPK power	\$11.10	0.95	2.82
Time-invariant orthog. poly.	\$11.99		
EPK orthogonal polynomial	\$11.64	2.90	1.74
<i>Hedge using underlying and ATM put</i>			
Time-invariant power	\$11.36		
EPK power	\$11.29	0.63	2.94
Time-invariant orthog. poly.	\$12.08		
EPK orthogonal polynomial	\$11.67	3.39	2.12

We also see that hedging with both the S&P 500 portfolio and the ATM put is inferior to hedging with the ATM put alone. If we know the true hedge ratios, it is always better to add more hedging instruments. However, including additional hedging instruments can decrease hedging performance when we estimate hedge ratios with error. Of the four tested specifications, we find that EPK power specification has lowest hedge portfolio standard deviation (\$11.29).

Our tests show that the hedging performance of the power specification is consistently superior to the performance of the orthogonal polynomial specification. Using any of the hedging instruments, the power specification reduces the hedge portfolio standard deviation more than the orthogonal polynomial specification. The same is true for the time-invariant hedges. These performance differences are statistically significant at the 5% level in five of six pair-wise comparisons (not reported in Table 11).

7. Conclusions

Our paper uses the no-arbitrage relationship between asset prices, payoff probabilities, and the pricing kernel to estimate an empirical pricing kernel. The empirical pricing kernel is the preference function that most closely reproduces observed asset prices, based on a forecast payoff density. Using a sequence of forecast asset payoff densities and cross-sections of asset prices, we can construct a time series of empirical pricing kernels and empirical risk aversion estimates.

We use S&P 500 index option prices and estimated S&P 500 return densities to estimate the empirical pricing kernel and empirical risk aversion each month from 1991 to 1995. Our analysis shows that empirical risk aversion is counter cyclical. Empirical risk aversion is positively correlated with indicators of recession such as widening of credit spreads and negatively correlated with indicators of expansion such as steepening of the term structure slope. This finding supports the results of Fama and French (1989).

We analyze two parametric specifications for the empirical pricing kernel. We observe that the orthogonal polynomial pricing kernel specification fits option prices better than does the power specification. However, the power pricing kernel specification is superior in terms of hedging performance. We conclude that the evidence is mixed with respect to the correct functional form for the empirical pricing kernel.

References

- Abel, A., 1990. Asset prices under habit formation and catching up with the Joneses. *American Economic Review* 80, 38–42.
- Ait-Sahalia, Y., Lo, A.W., 2000. Nonparametric risk management and implied risk aversion. *Journal of Econometrics* 94, 9–51.
- Arrow, K.J., 1964. The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies* 31, 91–96.
- Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. *Journal of Finance* 52, 2003–2049.
- Bansal, R., Viswanathan, S., 1993. No arbitrage and arbitrage pricing—a new approach. *Journal of Finance* 48, 1231–1262.
- Bansal, R., Hsieh, D.A., Viswanathan, S., 1993. A new approach to international arbitrage pricing. *Journal of Finance* 48, 1719–1747.
- Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.
- Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Economics Letters* 12, 143–172.
- Bollerslev, T., Chou, R.Y., Kroner, K.F., 1992. ARCH modeling in finance—a review of the theory and empirical evidence. *Journal of Econometrics* 52, 5–59.
- Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. ARCH models. In: Engle, R.F., McFadden, D.L. (Eds.), *Handbook of Econometrics*. North-Holland, Amsterdam, pp. 2959–3038.
- Breedon, D.T., Gibbons, M.R., Litzenberger, R.H., 1989. Empirical tests of the consumption-oriented CAPM. *Journal of Finance* 44, 231–262.

- Brown, D.P., Gibbons, M.R., 1985. A simple econometric approach for utility-based asset pricing models. *Journal of Finance* 40, 359–381.
- Campbell, J., 1996. Consumption and the stock market: interpreting international experience. *Swedish Economic Policy Review* 3, 251–299.
- Campbell, J., Cochrane, J., 1999. By force of habit: a consumption based explanation of aggregate stock market behavior. *Journal of Political Economy* 107, 205–251.
- Campbell, J., Lo, A., MacKinlay, C., 1997. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.
- Chapman, D., 1997. Approximating the asset pricing kernel. *Journal of Finance* 52, 1383–1410.
- Christensen, B.J., Prabhala, N.R., 1998. The relation between implied and realized volatility. *Journal of Financial Economics* 50, 125–150.
- Cochrane, J.H., 2001. *Asset Pricing*. Princeton University Press, Princeton, NJ.
- Constantinides, G.M., 1990. Habit formation: a resolution of the equity premium puzzle. *Journal of Political Economy* 98, 519–543.
- Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253–263.
- Eichenbaum, M.S., Hansen, L.P., Singleton, K.J., 1988. A time series analysis of representative agent models of consumption and leisure choice under uncertainty. *Quarterly Journal of Economics* 103, 51–78.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987–1007.
- Engle, R.F., Lee, G., 1999. A permanent and transitory component model of stock return volatility. In: Engle, R.F., White, H. (Eds.), *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger*. Oxford University Press, Oxford, England, pp. 475–497.
- Engle, R.F., Rosenberg, J.V., 1995. GARCH gamma. *Journal of Derivatives* 2, 47–59.
- Ermini, L., 1989. Some new evidence on the timing of consumption decisions and their generating process. *Review of Economics and Statistics* 71, 643–650.
- Estrella, A., Hardouvelis, G.A., 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46, 555–576.
- Fama, E.F., French, K.R., 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25, 23–49.
- Ferson, W.E., Harvey, C.R., 1992. Seasonality and consumption-based asset pricing. *Journal of Finance* 47, 511–552.
- Ghysels, E., Harvey, A.C., Renault, E., 1996. Stochastic volatility. In: Maddala, G.S., Rao, C.R. (Eds.), *Handbook of Statistics*. North-Holland, Amsterdam, pp. 119–191.
- Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–1801.
- Hansen, L.P., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. *Journal of Political Economy* 99, 225–262.
- Hansen, L.P., Richard, S., 1987. The role of conditioning information in deducing testable restrictions implied by asset pricing models. *Econometrica* 55, 587–613.
- Hansen, L.P., Singleton, K.J., 1982. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica* 50, 1269–1286.
- Hansen, L.P., Singleton, K.J., 1983. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy* 91, 249–265.
- Harrison, J.M., Kreps, D.M., 1979. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* 20, 381–408.
- Jackwerth, J., 2000. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* 13, 433–451.
- Jarque, C.M., Bera, A.K., 1980. Efficient Tests for normality, homoscedasticity, and serial independence of regression residuals. *Economics Letters* 6, 255–259.
- Judd, K., 1998. *Numerical Methods in Economics*. MIT Press, Cambridge.

- Lahiri, K., Wang, J.G., 1996. Interest rate spreads as predictors of business cycles. In: Maddala, G.S., Rao, C.R. (Eds.), *Handbook of Statistics*. Elsevier Science, New York, pp. 297–315.
- Ljung, G.M., Box, G.E.P., 1978. On a measure of lack of fit in time series models. *Biometrika* 65, 297–303.
- Lucas, R., 1978. Asset prices in an exchange economy. *Econometrica* 46, 1429–1445.
- Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- Pratt, J.W., 1964. Risk aversion in the small and in the large. *Econometrica* 32, 122–136.
- Ross, S.A., 1978. A simple approach to the valuation of risky streams. *Journal of Business* 51, 453–475.
- Rubinstein, M., 1976. The strong case for the generalized logarithmic utility model as the premier model of financial markets. *Journal of Finance* 31, 551–571.
- Shephard, N., 1996. Statistical aspects of ARCH and stochastic volatility. In: Cox, D.R., Barndorff-Neilson, O.E., Hinkley, D.V. (Eds.), *Time Series Models in Econometrics, Finance, and other Fields*. Chapman & Hall, London, England, pp. 1–67.
- Slesnick, D.T., 1998. Are our data relevant to the theory? The case of aggregate consumption. *Journal of the American Statistical Association* 16, 52–61.
- Startz, R., 1989. The stochastic behavior of durable and non-durable consumption. *Review of Economics and Statistics* 71, 356–363.
- Wilcox, D., 1992. The construction of U.S. consumption data: some facts and their implications for empirical work. *American Economic Review* 82, 922–941.