ECONOMIC SIGNIFICANCE OF RISK PREMIUMS IN THE S&P 500 OPTION MARKET

R. BRIAN BALYEAT

Option pricing is complicated by the theoretical existence of risk premiums. This article utilizes a testable methodology to extract the pricing impact resulting from these risk premiums. First, option prices (based on the full dynamics of the underlying) are computed under the assumption that these risk premiums are not priced. The pricing methodology is independent of any particular option-pricing model or distributional assumptions on the return process for the underlying. The difference between the actual market prices and these "no-premium base case" prices reflects the effect of risk premiums. For at-the-money, 13-week S&P 500 options trading from 1989 until 1993, the effect of risk premiums is statistically significant and averages slightly over 20% (in units of Black–Scholes implied volatility). A simple delta-hedging strategy is used to demonstrate the economic significance of risk premiums, as the trading strategy provides enough profit to absorb a crash similar in magnitude to the 1987 crash once every 6.20 years. © 2002 Wiley Periodicals, Inc. Jrl Fut Mark 22:1147–1178, 2002

The author gratefully acknowledges the comments of Jeff Fleming, Cam Harvey, Mark Holder, David Hsieh, Pete Kyle, George Tauchen, Robert Webb, an anonymous referee, and the seminar participants at the University of Georgia, the University of Sydney, Texas A&M University, and the 2001 FMA meetings in Toronto. As usual, all remaining errors are the author's.

For correspondence, R. Brian Balyeat, Department of Finance, Lowry Mays College and Graduate School of Business, Texas A&M University, 4218 TAMU, College Station, Texas 77843; e-mail: bbalyeat@cgsb.tamu.edu

Received November 2001; Accepted May 2002

■ *R. Brian Balyeat is an Assistant Professor in the Department of Finance at Lowry Mays College and Graduate School of Business at Texas A&M University in College Station, Texas.*

The Journal of Futures Markets, Vol. 22, No. 12, 1147–1178 (2002)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fut.10051

INTRODUCTION

The pricing of options relative to the characteristics of the underlying or the characteristics of other derivative securities is complicated by the theoretical existence of risk premiums in option markets. Much of the historical literature on option pricing simply assumed that these risk premiums did not exist. More recently, this assumption has been criticized. However, testing for and modeling risk premiums either has been indirect or has required taking a stand on the theoretical option-pricing model. This article develops a methodology to test for and determine the economic significance of risk premiums in derivative markets. The testing methodology is independent of any particular option-pricing model or distributional assumptions on the return process for the underlying. In addition, the methodology provides easy metrics by which its calibration can be tested as well as the volatilities necessary to properly hedge a portfolio of options necessary to capture these risk premiums.

The compensation for risk premiums is defined as the difference between the market price of an option and the price that would be implied (on the basis of the full dynamics of the underlying) if risk premiums are not priced. The semi-non-parametric technique (SNP) of Gallant and Tauchen (1989) is used to model the full dynamics of the S&P 500 index via its conditional density. In this case, the lagged values of the returns on the S&P 500 index and the 3-month Treasury bill rate are used as inputs for the SNP technique. Using Girsanov's theorem, a simple change in measure is then used to forecast these dynamics of the S&P 500 and price options on the basis of the index in a risk-neutral setting. Because option prices are not used to determine the SNP conditional density for the underlying, the resulting option prices are risk-neutral with respect to any risk premiums normally found in option markets.

Thus, it does not matter if risk premiums are compensating the investor for volatility risk, jump-frequency risk, jump-size risk, or other risks in the option markets. The difference between the market price and the "no-premium base case" price simply represents compensation for undertaking those risks. In this study, the average effect of risk premiums (measured in units of Black–Scholes implied volatility) for 13-week, at-the-money S&P 500 index options over the no-premium base case is 20.9%.

After determining the statistical significance of risk premiums, the next step is to determine their economic importance. If the market is efficient, any decision rule on whether to buy or sell a properly hedged portfolio of options should return the risk-free rate in the long run (excluding transaction costs). However, the key to this procedure is calculating the correct volatility to use in the hedging ratio. Fleming (1999) and others generally used the volatility implied by the market price of the option. To the extent that the market price includes the effects of any risk premiums, the volatility implied from the market price is liable to overstate the volatility of the underlying. To extract the pricing effects of option risk premiums, the hedge should replicate the portion of the option value driven solely by the fundamentals of the underlying. Thus, the hedge should be calculated using the volatility implied by the forecast of the underlying dynamics and not from the volatility implied by the market price of the option. Hedging with the volatility implied by market prices results in hedge ratios whose magnitude is overstated; therefore, the average profit is understated. Because this article utilizes a methodology that computes option prices in the absence of option risk premiums, it is a relatively simple matter to extract the appropriate hedging volatility.

Therefore, the decision rule of buying the option if the no-premium base case price is greater than the market price and selling the option otherwise is implemented. The portfolio is then delta hedged (updated weekly) with S&P 500 futures and held until maturity. However, if the delta-hedged portfolio earns a positive return, it cannot be concluded that the risk premiums in the option prices are necessarily economically significant. The profit could simply represent the compensation for bearing risk in the option market. To help put this compensation into perspective, a comparison is made between the average profit that the decision rule generates and the losses incurred if crashes of differing magnitudes arise during the time period. This ratio provides a measure of the economic importance of risk premiums, as it illustrates the frequency that market makers can absorb the losses resulting from simulated crashes of varying magnitudes. The results show that the losses generated during the simulated crashes can be absorbed quickly (relative to the historical frequency of the crashes) by the profits generated during non-crash quarters. For example, the strategy only takes 6.20 years to absorb the losses generated during a simulated crash similar in magnitude to the 1987 crash.

The remainder of this article is organized as follows. The first section reviews the existing literature concerning option pricing and the corresponding risk premiums. The data series used in the analysis are outlined in the second section. The third section outlines the SNP methodology used to calculate the conditional density of the S&P 500 index and the corresponding option-evaluation technique, whereas the fourth section discusses the existence, magnitude, and several of the properties of the pricing effects of risk premiums in this market. Estimates of the economic significance of risk premiums are presented in the fifth section and the sixth section concludes.

LITERATURE REVIEW ON RISK PREMIUMS RELATIVE TO OPTION PRICING

Financial no-arbitrage theory dictates that if options are redundant securities and the market is frictionless, then the state-price density (SPD) implied by option prices must be identical to the SPD implied by the forecast of the underlying's dynamics over the maturity of the option. Thus, in this case either risk premiums do not exist or are not priced in equilibrium. Any deviation between the two SPDs indicates that there are either arbitrage opportunities in the market or that one of the markets is pricing additional sources of risk. These additional sources of risk can arise from frictions within either of the two markets, mispricing because of hedging pressures, liquidity risks, or one of the two markets (presumably the option market) putting additional weight on the higher moments of the underlying's distribution.

Risk premiums in option markets have three important implications for both theoretical and empirical finance. First, given that an exact option-pricing model has not been developed, a reasonable alternative is to price a set of options using the prices from other traded securities. Jackwerth and Rubinstein (2001) priced one set of options on the basis of the prices (and thus implied volatilities) from other sets of options. In the presence of significant risk premiums, this process is complicated by the fact that the dynamics of those premiums must also be considered. Green and Figlewski (1999) looked at a variant of this theme by examining delta-hedged option-writing strategies where the options were priced with a volatility markup relative to their best volatility forecast from historical data. They concluded that writing options with this type of volatility markup and hedging them using premarkup volatility forecasts significantly reduced the risks involved.

Second, option markets provide a natural mechanism to gain insight into the expected dynamics of their respective underlyings. Engle and Mustafa (1992) used a GARCH model to estimate the volatility process of the underlying on the basis of a series of traded options. Ferson, Heuson, and Su (2000) measured the predictability in stock market returns given the implied volatilities from option prices on the respective underlying. Bates (2000) developed a two-factor geometric jump-diffusion process to model futures prices that nests both stochastic volatility and stochastic volatility/jump-diffusion models. After calibrating the model to observed option prices, Bates (2000) concluded that a stochastic volatility model would require implausible inputs for the volatility of volatility and that the stochastic volatility/jump-diffusion model did not match the characteristics of the underlying. Poteshman (1998) developed a method to estimate a general stochastic variance model for the underlying from option prices. From the options data, Poteshman (1998) was able to reject many of the restrictions to the general stochastic variance model that are common in the asset pricing literature.

Third, and possibly the most important implication of risk premiums, is their effect in a theoretical option-pricing framework. If risk premiums do not exist or are not priced in equilibrium, then markets are complete, and any derivative security can be priced in a no-arbitrage framework. This leads to option-pricing models where the volatility of the underlying is a deterministic function of time and path of the underlying asset. Models of this variety include the original Black–Scholes (1973) model; the Cox and Ross (1976) constant elasticity-of-variance model; the Rubinstein (1994), Derman and Kani (1994), Duprie (1994), and Derman, Kani, and Chriss (1996) implied tree models; and the Aït-Sahalia and Lo (1998) kernel approach. However, if risk premiums do exist and they are priced, then the no-arbitrage framework no longer applies, necessitating the use of stochastic option-pricing models. Models of this nature include the stochastic interest rate models of Merton (1973) and Amin and Jarrow (1991); the jump-diffusion models of Merton (1976), Bates (1991), and Madan, Carr, and Chang (1998); the stochastic volatility models of Hull and White (1987), Scott (1987), Wiggens (1987), Melino and Turnbull (1990, 1995), Stein and Stein (1991), and Heston (1993); the stochastic volatility and stochastic interest rate models of Bailey and Stulz (1989), Amin and Ng (1993), and Bakshi and Chen (1997a,b); and the stochastic-volatility/jump-diffusion models of Bates (1996a), Scott (1997), Bakshi, Cao, and Chen (1997), and Bates (2000). However, each of these models requires the estimation of additional parameters to account for the nondeterministic nature of volatility.

The earliest evidence to support the existence of risk premiums in option markets is from analysis of the informational content of implied volatility from stock-index options relative to the corresponding realized future volatility. Canina and Figlewski (1993) found that volatility implied from S&P 100 options provided a very poor forecast of the future volatility of the index. However, Day and Lewis (1992); Fleming, Ostdiek, and Whaley (1995); and Fleming (1998) found a much stronger relationship between the implied volatility and the future realized volatility. The rough consensus from this literature is that the Black–Scholes implied volatility generally has predictive power, but it is an upwardly biased predictor of the volatility realized over the life of the option. This conclusion can be viewed as evidence that risk premiums exist in these markets. However, those articles are not a direct test for the existence of risk premiums, as they jointly tested a given option-pricing model and the existence of risk premiums. Buraschi and Jackwerth (2001) also examined the existence of risk premiums in the S&P 500 index option market indirectly by testing to see if options are truly redundant assets. If options are redundant assets, then risk premiums are not present because the returns on the underlying and the at-the-money options will span the prices in the economy. Their conclusion was that out-of-the-money and in-the-money options are not redundant securities when the at-the-money option and the underlying are used to span the payoff space. Therefore, risk premiums must exist in this market.

Another means by which the existence of risk premiums can be examined is by testing the martingale property. The martingale property simply states that the expected price of the asset under the risk-neutral SPD implied by the cross section of option prices, discounted by the risk-free rate, must equal the current price of the underlying. Two studies that directly tested the martingale restriction are Longstaff (1995) and Brenner and Eom (1997). In Longstaff (1995), a Hermite polynomial expansion was used to approximate the SPD for a collection of S&P 100 index options. Longstaff (1995) found that the index level implied by the pricing kernel was significantly larger than the corresponding index level and, thus, rejected the martingale restriction. Brenner and Eom (1997) used a Laguerre orthogonal polynomial series on S&P 500 index options and were unable to reject the martingale property. Aït-Sahalia, Wang, and Yared (2001) extended this line of research by comparing the SPD implied by S&P 500 index option prices to the SPD implied by the underlying. They tested to see if the information contained in the cross section of option prices at a given point in time was consistent with the information contained in the time series of the underlying up to that point in time. Their conclusion was that the two SPDs are significantly different because the SPD implied by the option prices exhibited excess skewness and kurtosis relative to the SPD implied by the index.

Two articles that examined the joint dynamics of the underlying and the associated option prices are Chernov and Ghysels (2000) and Pan

(2001). Chernov and Ghysels (2000) used the SNP methodology on S&P 500 index returns and the corresponding at-the-money Black–Scholes implied volatility as the moment-generating function to test the Heston (1993) stochastic volatility model with the efficient method of moments procedure of Gallant and Tauchen (1996b). Although they found that the data rejected Heston's (1993) model, they concluded that, for the purposes of option pricing, a univariate approach using only option data dominates the bivariate approach. This implies that, relative to the dynamics of the underlying, risk premiums play a more important role in option pricing. Pan (2001) used the aforementioned Bates (2000) model to examine the joint dynamics of the S&P 500 index and 1-month, at-the-money options. The procedure separately identified both a stochastic volatility and a relatively more important jump-risk premium component in the option-pricing framework.

However, none of these articles examined or modeled risk premiums directly without assuming a specific option-pricing model and parametric form for the underlying. As noted in Bates (2000) and Chernov and Ghysels (2000), frequently the data rejected the parametric specification for the underlying, the option-pricing model, or both. Before the effects of risk premiums in option prices can be recovered, the dynamics of the underlying must be carefully and adequately modeled. Thus, this article uses a methodology to test the economic significance of the effects of risk premiums in option markets that is independent of any particular option-pricing model or distributional assumptions on the return process for the underlying. In addition, the process allows for direct examination of the effects of risk premiums through time, across strikes, and by option maturity.

DATA

The primary data set consists of weekly observations on the S&P 500 index and the 3-month U.S. Treasury bill rate from January 1954 through August 1996 (2226 weekly observations). Weekly data are used to balance the need to estimate the dynamics of the underlying as accurately as possible with the desire for a parsimonious model that avoids the complications presented in the estimation day-of-week effects for the mean and volatility parameters of both series. The 3-month Treasury bill rate is included in the data set because previous empirical work [see Campbell (1987); Whitelaw (1994); Boudoukh, Richardson, & Whitelaw (1997); and Scruggs (1998)] shows that the equity series behavior depends on the level of the interest rate. To construct a comprehensive set of weekly observations for the two series, the day of the week with the fewest missing observations is initially used. However, although Wednesday has the fewest missing observations (86 total for both series), almost 30% (25 of the total) of the missing observations occurred in the last 30 weeks of the 1968 equity market. As noted in French and Roll (1986), the 1968 equity markets were experiencing technical difficulties, and the backlog necessitated closing an extra day each week for order processing and updating equipment. Thus, because of the clustering of the missing Wednesday observations, the day with the next largest number of observations is used. Tuesdays have the next fewest missing observations (106 total for both series). If the Tuesday observation is missing for either series, the Wednesday data for both observations are used. Wednesday data exist for each of the missing Tuesdays. As noted in the third section, the model is originally calibrated to Tuesday data (fewest missing observations) and is then recalibrated to Friday data to correspond to S&P 500 index option maturities.

After collecting the weekly data set, two transformations are performed. For the S&P 500 index, geometric percentage returns are calculated; for the 3-month Treasury bill rate, the natural logarithm is used to prevent the simulation of negative interest rates. These transformations left 2225 observations for both series, 25 of which are dropped from the beginning of the data set to serve as start-up values.

The option data use Friday last-trade prices (as opposed to quotes) from March 17, 1989 to December 31, 1993 for all traded strikes on the S&P 500 index with maturities of at least one week. S&P 500 index options are chosen because they are European; thus, the analysis is not complicated by an early exercise premium. In addition, to avoid a nonsynchronous trading bias, the level of the S&P 500 index and the corresponding futures price at the time of each trade are also recorded. Because the S&P 500 index provides a dividend stream, a dividend series is needed for the option evaluations. The dividend series is derived by value-weighting the annualized, most recent dividend for each company in the index, summing them, and dividing by the current index level for that day.

METHODOLOGY

SNP Methodology

To derive the joint conditional dynamics of the stock and bond series, the SNP methodology of Gallant and Tauchen (1989) is used. SNP utilizes a Hermite polynomial expansion to approximate the density function of the underlying time series, y_t. Letting x denote a vector of lagged values of the time series up to a final lag L [i.e., $x = (y'_{t-L}, \ldots, y'_{t-2}, y'_{t-1})'$] and $z = R^{-1}(y - \mu)$ be the standard location-scale adjustment, the SNP density function with parameter vector θ is of the form

$$f(y \mid x, \theta) \propto [P(z, x)]^2 \, \mathbf{n}_M(\mathbf{y} \mid \boldsymbol{\mu}_x, \boldsymbol{\Sigma}) \tag{1}$$

with a constant of proportionality of $1/\int [P(s, x)]^2 \phi(s)\, ds$. In addition, $P(z, x)$ is a multivariate Hermite polynomial of degree K_z in z, where each of the coefficients can be a polynomial of degree K_x in x and can depend on L_p lags in x. The multivariate Gaussian density of dimension M, with mean vector μ_x, and variance–covariance vector Σ_x, is denoted $\mathbf{n}_M(\mathbf{y} \mid \boldsymbol{\mu}_x, \Sigma_x)$, and the standard multivariate Gaussian distribution with a zero mean vector and an identity variance–covariance matrix is represented by $\phi(\mathrm{z})$.

To decrease the number of terms in the Hermite polynomial, additional parameterization of the mean and variance of the Gaussian term is provided. The specification for the mean is a simple autoregressive function of the form $\mu_x = b_0 + Bx_{t-1}$. The number of lags in the autoregressive term is denoted L_μ. The variance specification is similar to the ARCH specification proposed by Nelson (1991) and Davidian and Carroll (1987), where $\boldsymbol{\Sigma}_x = \mathbf{R}_x\mathbf{R}'_x$, $vech(\mathbf{R}_x) = \boldsymbol{\rho}_0 + \mathbf{r}\,|\mathbf{x}_{t-1} - \boldsymbol{\eta}_{t-2}|$, and $|\mathbf{x}_{t-1} - \boldsymbol{\eta}_{t-2}|$ are vectors of error terms $y_{t-L_r} - \mu_{x_{t-L_r-1}}$ to $y_{t-1} - \mu_{x_{t-2}}$ of length L_r.

In addition to the tuning parameters already discussed (L_μ, L_r, L_p, K_z, and K_x), the SNP methodology provides two additional parameters, I_z and I_x. Both of these terms are designed to suppress interactions between the data series in the Hermite polynomial. Fixing I_z and I_x both equal to zero does not suppress any of the terms in the polynomial. Setting I_z to a positive integer suppresses all interaction terms in the polynomial of degree (in z) higher than $K_z - I_z$. The case is the same for I_x.

The parameters of the SNP specification are set in the usual maximum likelihood manner by minimizing

$$s_n(\theta) = -(1/n) \sum_{t=1}^{n} \log[f(y_t \mid x_t, \theta)] \tag{2}$$

The size of the specification (i.e., the individual values of the tuning parameters) is set by minimizing the Schwarz–Bayes information criterion (Schwarz, 1978)

$$\mathrm{BIC} = s_n(\hat{\theta}) + (1/2)(p_\theta/n)\log(n) \tag{3}$$

where p_θ represents the number of parameters, and n is the number of observations in each series.

As noted in Fenton and Gallant (1996), the BIC generally tends to be too conservative and, therefore, selects under-parameterized models. Thus, each model is subjected to a battery of specification tests designed to measure the predictability of the residuals. The specification tests, first proposed by Gallant, Rossi, and Tauchen (1992), are designed to determine if the scaled residuals have any power to predict the mean or variance of the data series. The mean test regresses the scaled residuals on 20 lags of the linear, squared, and cubic terms of the lagged data series. Likewise, the variance test regresses the square of the scaled residuals on 20 lags of the linear, squared, and cubic terms of the lagged data series. Any predictability of the residuals signals an underspecified model. Logical steps to correct predictability in the mean include adding lags in the autoregressive specification (i.e., increasing L_μ) or increasing the dependence of the Hermite polynomial on lagged values of the data by increasing K_x or L_p. For the variance, adding lagged values in the ARCH specification or increasing the dependence of the Hermite polynomial on lagged values of the data are logical steps to remedy predictability in the squared scaled residuals.

Estimation of the Joint Conditional Density

The estimation of the bivariate specification follows the procedure outlined in Gallant and Tauchen (1996a). A detailed analysis of the specification search is available from the author upon request. The final model is determined by a combination of the BIC criteria and the specification tests detailed previously. The preferred model is characterized by one nondiagonal lag in the autoregressive specification, 13 lags in the ARCH stock specification, 12 lags in the ARCH bond specification, and a Hermite polynomial of degree 4 whose coefficients at the linear level are linear functions of the lagged values of the data. Also, the ARCH specification is diagonal except for the contemporaneous constant, and interactions in the Hermite polynomial above the linear level are suppressed ($I_z = 3$). This model passes the mean and variance specification tests for both the bond and stock series with p values all easily above 0.01. Simpler models either fail one or more of the specification tests (p values of less than 0.01) or are dominated (either by a simple LRT for nested models or by the BIC for non-nested models) by the preferred model.

The large number of lags in the ARCH specification for both the stock return and the bond series is characteristic of the persistence in the volatility for both of these series. The fourth degree Hermite polynomial allows for excess skewness and kurtosis in these series relative to the normal. In addition, the coefficients on the linear terms in the Hermite polynomial are functions of the lagged data. This allows the shape of the distribution to depend directly on the path of the series in fairly complex ways.

When estimating option values by simulating the underlying, it is imperative that the conditional density of the underlying be estimated as accurately and flexibly as possible. There are three very appealing characteristics of the SNP methodology that make its use ideal in this setting. First, the SNP methodology nests many of the traditional financial econometric models. For example, setting $L_\mu = L_r = L_p = K_z = 0$ produces an i.i.d. Gaussian characterization. By then moving $L_\mu > 0$, a Gaussian VAR is produced, and then by setting $L_r > 0$, a Gaussian AR-ARCH is parameterized.

Second, according to Gallant and Nychka (1987), the SNP methodology is dense over the class of probability density functions (pdf). Hence, asymptotically, the SNP density, $f(y \mid x, \theta)$, can be made arbitrarily close to the true density, p. In other words, $\lim_{k\to\infty} \min_\theta \|p(\cdot|\cdot) - f_K(\cdot|\cdot, \theta)\| = 0$. These results also hold for densities based solely on the Hermite polynomial expansion. Therefore, the autoregressive and ARCH features of the SNP process cannot hurt the fit and can only aid in the model-selection process. The autoregressive and ARCH features of the SNP estimation are similar to the prewhitening process of Andrews and Monahan (1992).

Third, by setting $K_x > 0$ ($K_x = 1$ in this case), the shape of the error density becomes a function of the lagged data. Consequently, not only location and scale but also the basic shape of the pdf can depend on lagged values of the data. Thus, the SNP pdf is able to capture stochastic volatility or jump probabilities that are inherent in the dynamics of the underlying index. This heterogeneity and flexibility of the SNP methodology is an advantage not found in many other estimation techniques and is essential in any option-pricing setting.

A 25,000-week simulation from the bivariate specification is presented in Figure 1. The top two panels represent the equity series, and the bottom two stand for the bond series. The left-hand panels present the entire 480+ years in the simulation, whereas the right-hand panels display only the first 2200 observations from the simulation. The stock-return simulation has the "electrocardiogram" pattern that is expected

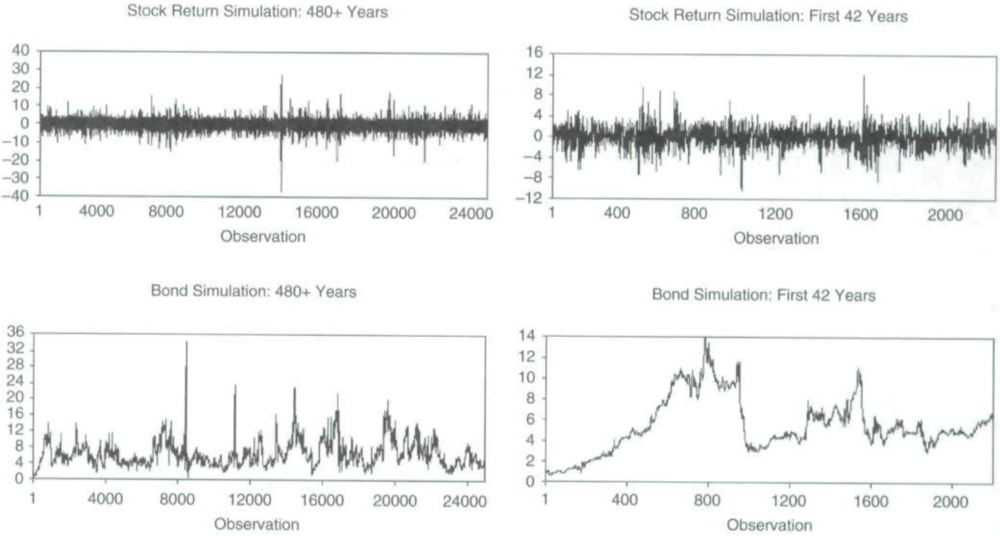

FIGURE 1
Bivariate pdf simulations. The top two panels represent the equity series, and the bottom two stand for the bond series. The left-hand panels present a 480+ year simulation, and the right-hand panels display only the first 2200 observations from the simulation.

for series of this nature. In addition, there are clearly times when the market experiences sharp declines similar in magnitude to the crash of 1987, and the 2200 point graph displays evidence of volatility clustering.

The bond simulation derived from the bivariate specification represents a very plausible forecast. What is arguably the most striking feature of these graphs is the ability of the model to produce bond simulations that appear to exhibit the regime-switching characteristics of Hamilton (1988) and Gray (1996). In the graph of the first 2200 observations, it appears as if the process has entered into a new regime on or about the 1000th observation. However, the SNP process is intrinsically of one regime.

Option-Evaluation Methodology

The 5-year option data set includes 22 dates in which 13-week options are traded. Two of these dates, May 22, 1992 and November 22, 1993, are dropped from the analysis because options on these dates only traded over a few strikes (10 and 5, respectively) and appeared to have low liquidity, as many of the last trades of the day occurred before noon. The remaining 20 dates correspond to the March, June, September, and December contracts and have expirations corresponding to the S&P futures contracts. The 13-week options are chosen for the analysis for three reasons. First, the 13-week options have reasonable liquidity because most strikes trade throughout the day. Second, by choosing 13-week options, the number of times during the analysis when it is necessary to switch maturity dates is kept to a reasonable level. Finally, as previously noted, these contracts have maturities that correspond to the S&P futures contracts, which will be used later for delta hedging.

To avoid an out-of-sample bias, a new SNP optimization is rerun for each of the 20 quarters previously mentioned, conditioning only upon the weekly data up to and including that date. The SNP specification is assumed to be the same size as the preferred Tuesday data model examined previously, only now re-estimated upon the corresponding Friday weekly data. The switch to Friday data is necessary to match the maturities of the S&P 500 options. As shown in Table I, all 20 of the SNP specifications passed the mean and variance specification tests for both the stock and bond series, even at the 5% level.

To compute option prices based on the SNP specification for a given day, a simulation from the bivariate conditional density is run W weeks ahead (W is the number of weeks to maturity for the option). For options with less than 13 weeks to maturity, the corresponding SNP specification with 13 weeks to maturity is used, but the conditioning vector is updated to include observations up to and including the date in question. A vector of 10,000 independent realizations of the W weeks-ahead equity return is collected. The mean of the vector is adjusted by subtracting the average return from each component and then adding back the cost of carry (the risk-free rate minus the dividend rate). This adjustment to the return vector does not affect the volatility (or any of the other higher moments) of the simulated returns and ensures that the options calculated from the simulations do not violate the put-call parity. The adjusted vector of equity returns corresponds roughly to the equity values of a 13-week binomial tree with time-varying volatility. Risk-neutral asset pricing is then applied to value the options over each of the 10,000 realizations of the process that are then averaged and discounted at the risk-free rate to calculate the option value. This methodology is motivated by Girsanov's theorem, as expanded by Harrison and Kreps (1979). When changing from the actual density to the risk-neutral SPD implied by the dynamics of the S&P 500, the (instantaneous) diffusion functions are identical under both measures, and only the drift needs to be adjusted.

The estimation process is repeated for each index level for which an is option traded. This is necessary because on any given day the options across strikes do not trade simultaneously and are thus priced in the market on the basis of different levels of the underlying index. Therefore,

TABLE I

SNP Estimates of Bivariate Weekly S&P 500 Returns and Logged 3-Month T-Bill Rates: Re-estimated for Each Quarter

									Specification Test: 20 Lags							
									S&P 500 Returns				*Logged 3-Month T-Bill Rate*			
	Tuning Parameters								*Mean*		*Variance*		*Mean*		*Variance*	
Date	L_μ	L_r	L_p	K_z	I_z	K_x	I_x	l_θ	*F Test*	*P Value*	*F Test*	*P Value*	*F Test*	*P Value*	*F Test*	*P Value*
890317	1	13/12	1	4	3	1	0	48	1.100	0.281	1.084	0.310	1.215	0.128	0.642	0.985
890616	1	13/12	1	4	3	1	0	48	1.092	0.295	1.096	0.289	1.215	0.128	0.645	0.984
890915	1	13/12	1	4	3	1	0	48	1.093	0.293	1.093	0.293	1.215	0.127	0.648	0.983
891215	1	13/12	1	4	3	1	0	48	1.161	0.189	0.947	0.591	1.192	0.152	0.596	0.994
900316	1	13/12	1	4	3	1	0	48	1.178	0.168	0.961	0.563	1.203	0.139	0.615	0.991
900622	1	13/12	1	4	3	1	0	48	1.172	0.175	0.952	0.582	1.204	0.139	0.619	0.990
900921	1	13/12	1	4	3	1	0	48	1.189	0.155	0.967	0.548	1.206	0.137	0.620	0.990
901214	1	13/12	1	4	3	1	0	48	1.175	0.171	0.969	0.544	1.216	0.127	0.627	0.989
910322	1	13/12	1	4	3	1	0	48	1.167	0.181	0.995	0.487	1.228	0.115	0.628	0.988
910621	1	13/12	1	4	3	1	0	48	1.170	0.177	0.998	0.480	1.221	0.121	0.628	0.989
910920	1	13/12	1	4	3	1	0	48	1.165	0.184	1.012	0.452	1.235	0.109	0.629	0.988
911220	1	13/12	1	4	3	1	0	48	1.181	0.164	0.981	0.517	1.249	0.097	0.590	0.995
920320	1	13/12	1	4	3	1	0	48	1.153	0.199	0.963	0.557	1.241	0.104	0.602	0.993
920619	1	13/12	1	4	3	1	0	48	1.139	0.220	0.972	0.538	1.256	0.092	0.608	0.992
920918	1	13/12	1	4	3	1	0	48	1.117	0.252	0.987	0.505	1.293	0.067	0.583	0.996
921218	1	13/12	1	4	3	1	0	48	1.107	0.269	1.002	0.472	1.294	0.066	0.594	0.994
930319	1	13/12	1	4	3	1	0	48	1.114	0.258	1.013	0.450	1.286	0.071	0.592	0.995
930618	1	13/12	1	4	3	1	0	48	1.114	0.257	1.013	0.450	1.296	0.065	0.599	0.994
930917	1	13/12	1	4	3	1	0	48	1.103	0.276	1.019	0.436	1.287	0.070	0.602	0.993
931217	1	13/12	1	4	3	1	0	48	1.100	0.282	1.024	0.426	1.291	0.068	0.606	0.993

Note. The results for the SNP bivariate specification estimated upon the corresponding Friday weekly data are presented for the 20 quarters under study. The specifications are characterized by 1 nondiagonal lag in the autoregressive specification, 13 lags in the ARCH stock specification, 12 lags in the ARCH bond specification, and a hermite polynomial of degree 4 whose coefficients at the linear level are linear functions of the lagged values of the data. Also, the ARCH specification is diagonal except for the contemporaneous constant, and interactions in the hermite polynomial above the linear level are suppressed ($I_z = 3$). Even at the 5% level, the SNP specifications pass the mean and variance specification tests for both the stock and bond series for each of the 20 quarters.

all of the options are evaluated on the basis of an adjusted simulation that conditions upon the history, which includes the return generated by the index level of the underlying at the time of the trade as the most recent lag of the equity series. To maintain comparability between option evaluations on the same day (on the basis of a different index level), each simulation uses the same vector of random numbers. However, using a unique vector of random numbers for the various simulations in a given day usually only affects the option evaluations by less than a penny.

Mathematically, for a put and call with strike X, time to maturity T, initial index level S_0, risk-free rate r_f, dividend rate d_T, and N simulated returns each denoted as R_i

$$c = \frac{\sum_{i=1}^{N} \max[0, (S_{T,i} - X)]e^{-r_f T}}{N}$$

and

$$p = \frac{\sum_{i=1}^{N} \max[0, (X - S_{T,i})]e^{-r_f T}}{N} \tag{4}$$

where

$$S_{T,i} = S_0^*(R_i - \mu_T + r_f - d_T) \quad \text{and} \quad \mu_T = \frac{1}{N}\sum_{i=1}^{N} R_i$$

This is simply a discrete approximation of

$$f = e^{-r_f T}\int \psi_f\, p(S_T)\, dS_T \tag{5}$$

where ψ_f is the terminal value function for the option in question (i.e., $\max[S_T - X, 0]$ for call options and $\max[X - S_T, 0]$ for puts), and $p(S_T)$ is the risk-neutral SPD implied by the dynamics of the S&P 500 index.

This process has several advantages over the canonical approach in Stutzer (1996). Like Stutzer (1996), this approach uses a long time series to compute the density for the underlying. However, unlike Stutzer (1996), the shape of the SNP density is conditional and is able to simulate previously unobserved realizations.

Characteristics of the Volatility Series and the Effect of Risk Premiums

The first step in the analysis is to determine the magnitude of risk premiums and examine several of their characteristics. However, before risk premiums and their effects can be examined, it is necessary to calculate an implied volatility for each market option in the data set and for the corresponding options derived from the SNP simulations. The

Black–Scholes option-pricing formula is inverted, as originally proposed by Latané and Rendleman (1976), to obtain an implied volatility for each market option on the basis of the level of the index at the time of the trade. The same is done for each of the corresponding simulated option prices. The Black–Scholes model is chosen because it is a convenient metric for scaling and allowing comparisons of option prices across strikes, option maturities, and time, not because it is deemed to be the appropriate option-pricing model.

The option prices created from the dynamics of the underlying are not competing with the Black–Scholes model (or any other option-pricing model) in a "horse race;" rather these option prices are specifically designed to be "no-premium base case" prices. By their very construction, these option prices assume that risk premiums are either zero or not priced. As mentioned in the previous section, the SNP methodology computes a conditional density at time t on the basis of the lagged values of the variables in question up to and including time $t - 1$. Because the option prices that result from the SNP simulations of the index returns do not involve actual market-option prices, they are risk-neutral with respect to any risk premiums implied by the option market.

If there are risk premiums in the S&P 500 index option market, then the no-premium base case prices should be significantly less than the corresponding market prices for these options. This ordering is then preserved when the prices are transformed into their corresponding Black–Scholes implied volatilities. The difference between the two implied volatility series represents the effect of risk premiums. As noted in Bates (1996b), ". . . risk-neutral and actual expected average variances will diverge in the presence of a substantial risk premium." Figures 2 and 3 graph the 1991 implied Black–Scholes volatilities that correspond to the market and no-premium base case prices. Because each point is evaluated on the basis of the index level at the time of trade, the implied volatility smirks are not necessarily smooth.

From these graphs, two very important factors become evident. First, the implied volatilities match the "smile," or more accurately, the "smirk" pattern noted in the literature for the postcrash S&P 500 index options [see Rubinstein (1994)]. In other words, option-implied volatilities increase for strikes less than the current index level (out-of-the-money puts and in-the-money calls), flatten for near-the-money strikes, and begin to rise slowly for strikes above the at-the-money level (out-of-the-money calls and in-the-money puts). This effect has been attributed to institutional factors such as the market charging a premium for buying out-of-the-money puts as portfolio insurance and a thin market for

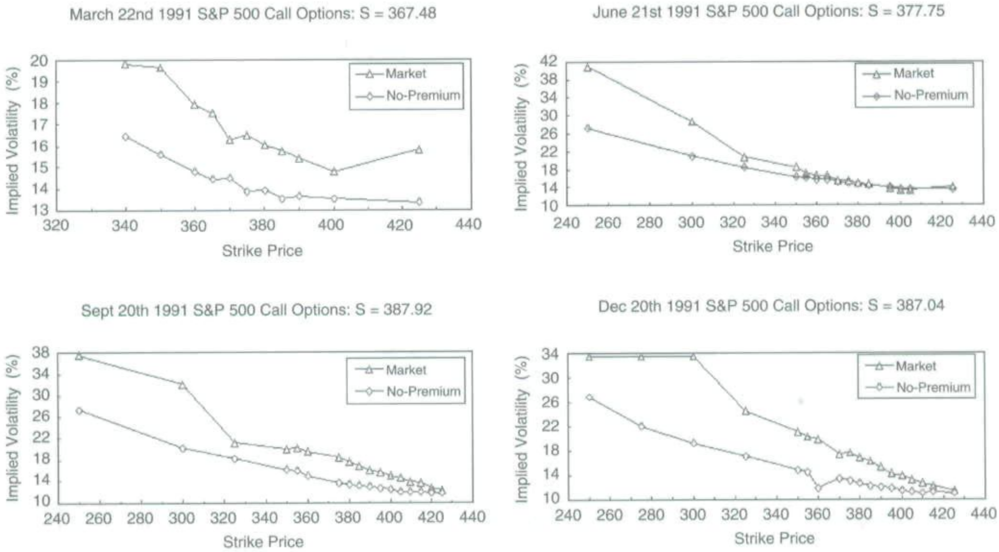

FIGURE 2

1991 call-implied volatility. The four panels in this figure represent the four quarters in 1991. Each panel graphs the Black–Scholes implied volatility derived from each strike in 1991 for which call options were traded in the market against the implied volatility for the corresponding no-premium base case options on the basis of the SNP simulations of the index. Because each point is evaluated based on the index level at the time of trade, the implied volatility lines are not necessarily smooth. The graphs for calls in other years look similar and are not presented.

out-of-the-money calls. However, part of this "smirking" effect appears to be endogenous and most likely results from an asymmetric volatility effect similar to the "leverage" effect of Black (1976) and Christie (1982). Thus, the existence of the implied volatility smirk is not solely due to risk premiums in option prices. In the simulations, if the realization begins to move away from the at-the-money level, the process begins to condition upon larger price movements, thus increasing the expected variance for the next week as a result of volatility clustering. However, this process is asymmetric, as negative price movements increase the expected volatility by a greater amount than do corresponding positive price movements. Hence, for the realizations of the W-week process that produce value for the out-of-the-money puts, a rise in the volatility of the series would be expected. The same is true for in-the-money calls. However, because the market does not seem to be as interested in selling deep out-of-the-money calls as it is in buying deep out-of-the-money puts, this pattern does not always fully materialize because high-enough strikes are not always traded.

The second striking feature of the implied volatility graphs is that the implied volatilities from the no-premium base case prices are consistently

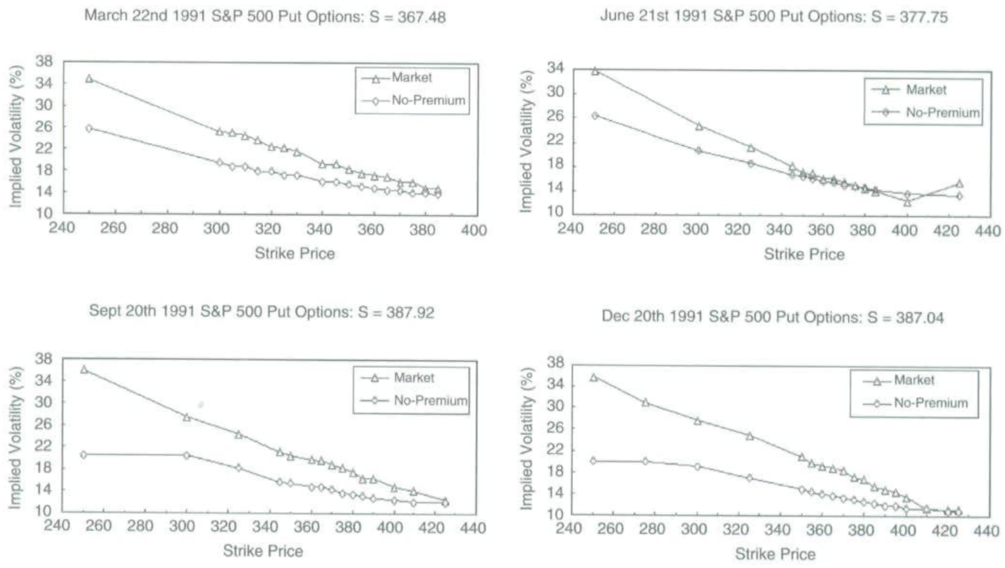

FIGURE 3
1991 put-implied volatility. The four panels in this figure represent the four quarters in 1991. Each panel graphs the Black–Scholes implied volatility derived from each strike in 1991 for which put options were traded in the market against the implied volatility for the corresponding no-premium base case options on the basis of the SNP simulations of the index. Because each point is evaluated based on the index level at the time of trade, the implied volatility lines are not necessarily smooth. The graphs for puts in other years look similar and are not presented.

less than the market-implied volatilities. Jackwerth and Rubinstein (1996) made a similar observation by comparing at-the-money implied volatilities to the historical volatility of the S&P 500 index. In this article, the results not only hold up across strikes but also against a forecast of the underlying's volatility. Yet, this is what would be expected if the option market is applying risk premiums. This simply means that options are not redundant assets. Thus, options are priced with additional source(s) of risk relative to the corresponding underlying spot market. These additional risk sources could result from frictions in the markets, hedging pressures, jump risks, or additional weight placement on the higher moments of the underlying.

For both the puts and calls, the difference between the two implied volatility measures generally decreases as the strikes approach the at-the-money level. However, given risk premiums, this pattern of convergence is expected. Options with strikes at or slightly above the money are the most sensitive to changes in the volatility of the underlying and, hence, would not demand as great a premium. For example, in the Black–Scholes option-pricing model, the vega of an option (the first derivative of the option price relative to the volatility of the underlying) decreases monotonically as it moves away from the at-the-money level, but it decreases more quickly for strikes below the at-the-money level than for strikes above that level. In Heston's (1993) stochastic volatility model, risk premiums are proportional to the current variance of the underlying. Thus, given that realizations of the underlying away from the at-the-money level are more likely to have come from a distribution with a larger variance, the Heston (1993) model would predict that away-from-the-money options would have higher risk premiums than at-the-money options. In addition, Heston's (1993) model would predict that risk premiums are not a function of the time to maturity of the option. What Heston's (1993) model does not explain is why the implied volatilities for the market and the no-premium base case continue to converge for strikes slightly above the at-the-money level and only begin to diverge again for strikes well above the at-the-money level. This is most likely the result of a jump-risk premium for large negative realizations. A risk premium for large negative realizations would imply that the difference in the two implied volatility graphs should decrease as the strike price increases.

To assess the magnitude of this undervaluation for the entire sample and to get a better understanding for the dynamics of the implied volatilities, weekly series are derived for the market and no-premium base case-implied volatilities. These series are then compared to the volatility that is realized for the underlying over the remaining life of the option in question. The market-implied volatility series is computed by averaging the annualized implied volatilities of the at-the-money put and call that expire at the end of the quarter in question. Hence, the first observation is based on an option with 13 weeks to maturity, the second is with 12 weeks, and so forth. This pattern is continued until only 1 week before maturity remains; then, the next contract is selected (usually with 13 weeks until maturity). To maintain consistency, the implied volatilities for the no-premium base case are calculated in a similar manner. The realized series is derived by annualizing the square root of the average daily squared returns that occur during the life of the contract. As a result of Jensen's Inequality, the Black–Scholes implied volatility will understate the actual volatility implied by the SPD. However, as noted in Feinstein (1989), the Black–Scholes formula is nearly linear in volatility for at-the-money options. Thus, the bias due to the strict concavity of the Black–Scholes function relative to volatility is of little concern in this application. The resulting series are graphed in Figure 4. For clarity in the graph, the lines are discontinuous at the points where volatilities on the basis of results from one quarter end and a new quarter begins. Table II contains descriptive statistics for the three implied volatility series, excluding the transitions between quarters.

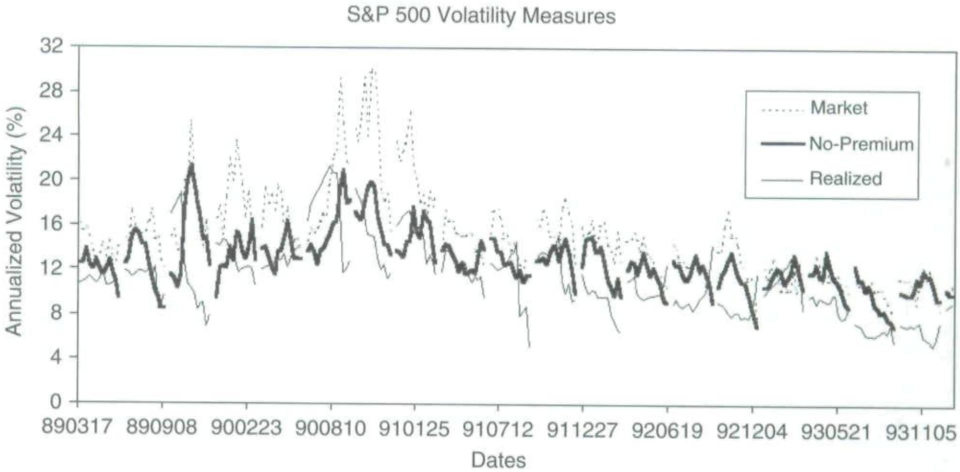

FIGURE 4

Volatility measures. The market-implied volatility series is computed by averaging the annualized implied volatilities of the at-the-money put and call that expire at the end of the quarter in question. Hence, the first observation is based on an option with 13 weeks to maturity, the second is with 12 weeks, and so forth. This pattern is continued until only 1 week before maturity remains; then, the next contract is selected (usually with 13 weeks until maturity). To maintain consistency, the implied volatilities for the no-premium base case are calculated in a similar manner. The realized series is derived by annualizing the square root of the average daily squared returns that occur during the life of the contract. For clarity in the graph, the lines are discontinuous at the points where volatilities on the basis of results from one quarter end and a new quarter begins.

Figure 4 and Table II demonstrate that the market-implied volatility generally exceeds the implied volatility from the no-premium base case, which is in turn greater than the realized volatility. This relationship is as expected. The no-premium base case-implied volatilities are conditioning upon the entire range of possible price movements, which naturally includes the possibility of large price movements as dictated by the dynamics of the SNP conditional pdf. Because these rare, large price movements are not realized during most of the sample period, it is logical to expect that the volatility series from the estimated no-premium base case will generally exceed the realized volatility. The market-implied volatilities are even greater than the implied volatilities from the no-premium base case because, not only are they allowing for the possibility of a large movement as dictated by the historical dynamics of the S&P 500 index, they are also applying risk premiums.

Likewise, the variance of the market-implied volatility series is almost three times that of the no-premium base case series and is over one and a half times larger than the variance from the realized estimations. On the surface, this would appear to violate a variance bound, as

TABLE II
Descriptive Statistics on Implied Volatility Series

	Market	*No Premium*	*Realized*
	Levels		
Mean	0.1505	0.1277	0.1131
Standard error	0.0026	0.0016	0.0021
Median	0.1459	0.1257	0.1106
Standard deviation	0.0413	0.0253	0.0329
Sample variance	0.0017	0.0006	0.0011
Kurtosis	1.6838	1.0907	0.4941
Skewness	1.1776	0.6938	0.7119
Range	0.2232	0.1446	0.1622
Minimum	0.0759	0.0694	0.0511
Maximum	0.2991	0.2140	0.2133
Sum	37.7771	32.0445	28.3804
Count	251	251	251
	Correlations		
Market	1		
No premium	0.8031	1	
Realized	0.6095	0.4835	1

Note. Summary statistics are presented for the three implied volatility series (excluding the transitions between quarters) for the period of Mar. 17, 1989 until Dec. 31, 1993. The market series consists of the average of the at-the-money put- and call-implied volatility from the market prices for the nearby S&P 500 index options maturing on the June, September, December, and March cycle. The no-premium base case series is calculated similarly using option prices (derived from the SNP simulations of the index) with the same time to maturity and based on the same index level as the corresponding market prices. The realized series is calculated by annualizing the square root of the average daily squared returns that occur in the underlying during the life of the contract. In addition, the correlations between each of the series are detailed.

the market variance is expected to be less than the realized variance. However, if the market is applying risk premiums, this bound is no longer necessarily violated. Also, there is probably a larger measurement error when calculating implied volatilities as compared with realized volatilities. The excess kurtosis of the market implied volatility series is also considerably larger than that of the no-premium base case series, which in turn is larger than the realized. This again would be expected, as the density estimation of the index returns is conditioning upon the recent history and includes the possibility of large price movements as dictated by the dynamics of the underlying. Furthermore, the market is doing the same thing and is applying risk premiums.

In addition to showing descriptive statistics for these three implied volatility series, Table II presents the pairwise correlations among the three series. The volatility series from the market and no-premium base case have a correlation of over 80%. This suggests that the market could be using more information than just the history of the series. The potential importance of this additional information for option-pricing applications is left for further research. The realized volatility series has a correlation of 60.95% with the market and 48.35% with the no-premium base case. Hence, the market appears to do a better job predicting the volatility. However, this is to be expected because the market is conditioning upon a larger information set than the SNP estimations of the index returns.

Because the observations graphed in Figure 4 potentially exhibit serial correlation, an analysis grouping the observations by time to maturity is presented in Table III. Grouping the observations by maturity also controls for any time to maturity effects in the data. For a given maturity, each of the data points is from a different maturity date (for the market series) or SNP specification (for the no-premium base case). Thus, the data points for a given maturity level should not be correlated. For each of the 13 weekly periods, the average realized volatility is less than the no-premium base case-implied volatility, which in turn is less than the market-implied volatility. At the 13-week horizon, the difference between the market-implied volatility and the no-premium base case is a little over 20%. This difference remains relatively stable over the remaining 12 maturities. Table III also presents the t tests for the null hypothesis that the average difference between the respective series is zero. At an alpha of 0.05 and a critical value of 1.645 for a one-tailed test, the market-implied volatility is always significantly above the realized volatility from the same period. The same is true for the market-implied volatility as compared with the no-premium base case.

ECONOMIC SIGNIFICANCE OF RISK PREMIUMS

Calculation Methodology

To determine the economic significance of risk premiums, first, the simple trading rule of buying the option if the no-premium base case price is greater than the market price, and selling the option otherwise is implemented. As noted by Whaley (1986), to ensure convergence, the options are held to maturity and delta-hedged weekly. The positions are delta-hedged so that the resulting profits are due to a difference in the option values and not to fortuitous market moves.

To provide a more accurate hedge and to avoid calculating a dividend position weekly, the position is hedged with the S&P 500 index futures. For hedging purposes, it is assumed that the traditional Black–Scholes delta hedge holds locally for each option. The simple Black–Scholes delta hedge is chosen because, as noted in Dumas, Fleming, and Whaley (1996), it

TABLE III

Differences in Average Volatility Measures by Weeks to Maturity.

	Market		No-Premium		Realized			Market versus No-Premium Average Difference	Market versus Real Average Difference	No-Premium versus Real Average Difference
	Mean	*Standard Deviation*	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*	*Count*	*t Test*	*t Test*	*t Test*
13 weeks	0.1535	0.0357	0.1262	0.0176	0.1212	0.0345	20	4.68	6.23	0.64
12 weeks	0.1524	0.0354	0.1275	0.0176	0.1217	0.0350	20	4.65	5.52	0.91
11 weeks	0.1568	0.0371	0.1273	0.0171	0.1214	0.0367	20	5.07	5.03	0.76
10 weeks	0.1624	0.0445	0.1287	0.0195	0.1220	0.0377	19	4.75	4.67	0.89
9 weeks	0.1578	0.0443	0.1329	0.0232	0.1177	0.0342	19	3.55	5.74	2.27
8 weeks	0.1591	0.0479	0.1390	0.0268	0.1167	0.0339	19	2.90	4.57	3.04
7 weeks	0.1631	0.0532	0.1370	0.0295	0.1164	0.0351	19	3.88	4.40	2.61
6 weeks	0.1557	0.0427	0.1322	0.0252	0.1155	0.0348	19	3.86	4.65	2.23
5 weeks	0.1514	0.0365	0.1302	0.0255	0.1113	0.0334	19	4.18	5.78	2.75
4 weeks	0.1453	0.0429	0.1290	0.0273	0.1033	0.0274	19	2.60	6.44	5.04
3 weeks	0.1408	0.0410	0.1234	0.0306	0.0990	0.0232	19	3.49	5.56	4.76
2 weeks	0.1350	0.0327	0.1181	0.0289	0.1005	0.0236	19	4.27	4.48	2.83
1 week	0.1219	0.0334	0.1072	0.0278	0.0982	0.0278	18	3.13	2.67	1.15

Note. The first six columns of the table contain averages and standard deviations of the three volatility series by weeks to maturity. The market series consists of the average of the at-the-money put- and call-implied volatility from the market prices for the nearby S&P 500 index options maturing on the June, September, December, and March cycle. The no-premium base case series Is calculated similarly using option prices (derived from the SNP simulations of the Index) with the same time to maturity and based on the same index level as the corresponding market prices. The realized series is calculated by annualizing the square root of the average daily squared returns that occur in the underlying during the life of the contract. The count column is the number of distinct maturity dates in the sample. The last three columns provide the value of the t tests under the null hypotheses that the average difference in the Indicated two series is different from zero.

performes better as a simple hedge on the S&P 500 than all of the other competing deterministic volatility models out of sample. In addition, the positions are based on one unit of the underlying as opposed to a contract size of 500, and mark-to-market effects for the futures are ignored. Aït-Sahalia, Wang, and Yared (2001) adopted a similar option decision rule on an unhedged portfolio.

It is important to examine which volatility is used to compute the hedge ratio. The two choices are the implied volatility from market prices or the implied volatility from the no-premium base case prices. Using the implied volatility from market prices leads to two problems. First, options of a given strike do not always trade from week to week, especially if there is a large move in the market. Thus, market prices are not always available for these options across the quarter; hence, the implied volatilities required to compute and update the hedging ratios would have to be estimated. Second, and more importantly, as explained previously, the market prices do not provide the correct implied volatilities even when the options do trade. The purpose of this trading rule exercise is to capture the pricing differential between the market price and the no-premium base case price. It is the no-premium base case option that is being replicated in the futures market; thus, its implied volatility must be used. Without the no-premium base case prices, it is impossible to accurately hedge the market options.

Results from the Hedging Strategy

This hedging strategy is expected to make profit over the 20 quarters because the realized volatility is almost always less than the volatility implied by the dynamics of the index, which is usually less than the volatility implied by market prices. Hence, the strategy frequently involves writing delta-hedged options that are not exercised. This strategy is similar to Fleming (1999) who used both delta-neutral hedged portfolios and butterfly spreads to test for mispricing in the S&P 100 index option market. For this sample, the decision rule only buys 27 of the 237 calls and 23 of the 297 puts. The options purchased under this strategy are almost exclusively either deep in-the-money calls or deep in-the-money puts. The purchase of these deep in-the-money options does not necessarily signal a sign change in risk premiums, as the pricing differences between the market prices and the no-premium base case prices are frequently within an eighth and more likely represent a bid-ask bounce phenomenon. In addition, these options are almost guaranteed to finish in-the-money and, therefore, must be priced as the discounted value of their intrinsic value. Thus, they are not really candidates for the examination of pricing effects of risk premiums. In either case, they have deltas of near one and are always exercised; thus, they have a minimal impact on the following analysis.

The results from the buy–sell decision rule are shown in the third column of Table IV. The trading strategy makes money in 16 of the 20 quarters in question (and comes very close to breaking even in an additional quarter). In each of the 16 profitable quarters, the volatility realized over the quarter is less than the volatility implied by market prices, which in turn is greater than the volatility implied by the estimated

TABLE IV
Hedged Portfolio Profits

	Date	Base Case Total Profit	10% Crash Total Profit	1987 Crash (27.33%) Total Profit	50% Crash Total Profit
1	89/03/17	33.59	−42.76	−841.15	−2306.80
2	89/06/16	16.32	−106.63	−951.72	−2405.47
3	89/09/15	−120.26	−322.85	−1304.34	−2949.46
4	89/12/15	24.18	−58.56	−408.65	−952.02
5	90/03/16	−0.01	−81.12	−872.89	−2262.25
6	90/06/22	−1.82	−39.67	−205.65	−480.89
7	90/09/21	104.94	−90.17	−833.83	−1836.52
8	90/12/14	34.92	−35.55	−981.99	−2946.46
9	91/03/22	83.73	−179.97	−1451.64	−3445.87
10	91/06/21	26.05	−168.05	−1137.18	−2688.93
11	91/09/20	76.72	−166.87	−1145.31	−2713.74
12	91/12/20	20.36	−259.28	−1749.30	−4051.99
13	92/03/20	99.57	−159.38	−1296.40	−2937.77
14	92/06/19	63.16	−225.41	−1475.71	−3337.58
15	92/09/18	57.86	−181.54	−1102.79	−2433.06
16	92/12/18	31.22	−291.06	−1903.64	−4270.88
17	93/03/19	58.23	−153.81	−1217.54	−2703.40
18	93/06/18	−13.78	63.49	−112.06	−493.84
19	93/09/17	68.11	−252.43	−1434.28	−3243.48
20	93/12/17	42.98	−210.63	−1362.31	−3029.81
	Average	35.30	−148.11	−1089.42	−2574.51

Note. The third column of the table illustrates the profit per quarter on the basis of the simple trading rule: buy the option if the market price is below the no-premium base case price, and sell the option otherwise. As a result of the presence of risk premiums in the option market, over 90% of the options are written. To ensure that the resulting profit is truly due to a difference in the option values, the position is delta hedged (updated weekly) using the volatility implied by the no-premium base case prices, and the options are held until maturity. The remaining columns illustrate the profits (losses) from the same strategy conditional upon a crash of the indicated magnitude occurring during the last week of the given quarter. The results for any given quarter are only conditional on a crash occurring in the last week of that quarter and not upon crashes having occurred in previous quarters. Given that the original option strategy involved writing options, this strategy naturally loses money during quarters containing a simulated crash, with larger simulated crashes experiencing greater losses.

dynamics of the index. Thus, the delta-neutral option-writing strategy is profitable during these quarters.

For the third quarter of 1989 and the second quarter of 1990, although the volatility implied by the estimated dynamics of the index is less than the volatility implied by the market, the volatility realized in the market is higher than the market's prediction. Hence, the trading strategy produces losses. For the second quarter of 1993, the situation is reversed. Although the realized volatility is low, the estimated no-premium base case prices imply a slightly greater volatility than the market prices. Hence, the trading strategy also produces losses in this quarter. The first quarter of 1990 is unusual because the relative implied volatilities predict that the trading strategy should make money, but the result is a slight loss.

As the third column in Table IV shows, the delta-hedged strategy is profitable over the sample period. This does not necessarily indicate a market inefficiency or arbitrage opportunity; rather, the average profit simply represents the return for bearing risk in this market. However, before it can be concluded that this profit is economically significant, the profit needs to be compared to the losses that would be incurred if crashes of various magnitudes occurred during the quarters in question. To do this, crashes of various magnitudes are simulated during the last week of each quarter. The effect of each crash is independent of the other crashes. In other words, the results for the loss incurred if a crash occurs in the third quarter of 1992 conditions upon the actual history up to and until the 12th week of the quarter and not upon a crash already having occurred in the second quarter of 1992. The rest of Table IV presents the losses incurred for simulated crashes of 10, 27.33, and 50%, respectively, for the 20 quarters in the sample. The 10% drop is chosen to represent a market correction; the 27.33% decrease represents the 1987 crash; and the 50% drop represents a catastrophic event. As expected, as the crash severity increases, the strategy loses ever-increasing sums.

To further interpret the economic significance of risk premiums, the average loss that occurs as a result of a specific crash is divided by the average profit received during the noncrash quarters (i.e., the base case). The resulting quotient is simply the mean number of quarters that the market would require between crashes of the given magnitude to break even. This methodology is not intended to predict the exact dates when crashes will occur, nor is it designed to test to see if specific crashes are predicted by the market as in Bates (1991). Because the third quarter of 1989 experienced a market correction of over 7%, it cannot be classified as a noncrash quarter. Therefore, the results that follow exclude the results from the third quarter of 1989.

TABLE V
Crash Survivability

	Base Case	87 Crash (27.33%)		50% Crash		10% Crash	
	Average Profit	Average Loss	Years to Break Even	Average Loss	Years to Break Even	Average Loss	Years to Break Even
89–93	43.49	−1078.11	6.20	−2554.78	14.69	−138.92	0.80
89	24.70	−733.84	7.43	−1888.10	19.11	−69.32	0.70
90	34.51	−723.59	5.24	−1881.53	13.63	−61.63	0.45
91	51.71	−1370.86	6.63	−3225.13	15.59	−193.54	0.94
92	62.95	−1444.64	5.74	−3244.82	12.89	−214.35	0.85
93	38.88	−1031.55	6.63	−2367.63	15.22	−138.34	0.89
89–90	30.30	−727.98	6.01	−1884.34	15.55	−64.92	0.54
91–93	51.18	−1282.35	6.26	−2945.86	14.39	−182.08	0.89

Note. This table compares the no-premium base case profits to the losses incurred during simulated crashes of various magnitudes, as computed in Table IV. Assuming that the option market is efficient and is therefore a zero sum game, the negative of the ratio of losses incurred during simulated crashes to profits gained in the no-premium base case is the average number of quarters between crashes of various magnitudes before the hedging strategy breaks even. Thus, if the general dynamics of risk premiums continue with this pattern and a crash on the order of the 1987 crash does not occur every 6.20 years, then the risk premiums are economically significant.

Table V summarizes the average quarterly profit from the hedging strategy and the average quarterly loss that would occur if the respective crash occurred in the last week of the quarter in question. Assuming that the option market is efficient and is therefore a zero sum game, the negative of the ratio of losses incurred during simulated crashes to profits gained in the base case is the average number of quarters between crashes of various magnitudes before the hedging strategy breaks even. Thus, if the general dynamics of risk premiums continue with this pattern and it is reasonably assumed that a crash on the order of the 1987 crash does not occur every 6.20 years, then the pricing effects of risk premiums are economically significant. The degree of the economic significance depends on the frequency at which large negative realizations are forecasted in the physical measure for the underlying. The strategy also yields enough profit to absorb a 10% crash once every 0.80 years or a 50% crash (which has never been realized) once every 14.69 years. With the possible exception of 1990, the results are amazingly consistent across time. The 1990 options are priced with higher risk premiums that probably resulted from either increased fears because of the October 1989 market correction previously discussed or to Iraq's invasion of Kuwait.

CONCLUSIONS

This article has examined existence, magnitude, properties, and economic importance of risk premiums in the S&P 500 index option market. The compensation for these risk premiums has been analyzed by examining the difference between the volatility implied by the market price for the option and the volatility implied by the price of the option that would result if these premiums are not priced. The no-premium base case option prices are calculated from risk-neutralized simulations of the underlying. Because these simulations do not include any option prices in the calculation methodology, they are by definition risk-neutral with respect to any option-risk premiums. An interesting result is that the no-premium base case option prices alone are able to replicate the pattern of the implied volatility smile. Thus, the implied volatility smile is not solely due to option-risk premiums. The results show that the S&P 500 index option market is trading at a volatility level that is above both the realized and no-premium base case volatility levels. The average pricing effect (in units of Black–Scholes implied volatility) of risk premiums for 13-week, at-the-money options over the no-premium base case is 20.9%. This result is reasonably stable across maturities.

Delta-neutral portfolios are formed by selling (buying) the option when the market price is above (below) the corresponding no-premium base case price. With the exception of a few deep in-the-money options, this strategy usually involves selling the options in question. These portfolios are then hedged using the Black–Scholes delta at the volatility implied by the no-premium base case level. This strategy consistently produces profits above the break-even level. The economic significance of these profits is measured by simulating crashes of varying magnitudes during the quarters in question. The results confirm that the losses generated during the simulated crashes could be absorbed quickly (relative to the historical frequency of the crashes) by the profits generated during noncrash quarters. For example, profits generated under the strategy can absorb the losses from a 10% crash, a 27.33% crash on the order of the '87 crash, and a 50% crash once every 0.80, 6.20, and 14.69 years, respectively.

The next step in the research is to begin to explicitly consider the existence and characteristics of risk premiums when using options data to compute or imply dynamics or prices both for the underlying and for other options and when trying to develop theoretical option-pricing models.

BIBLIOGRAPHY

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

Aït-Sahalia, Y., Wang, Y., & Yared, F. (2001). Do option markets correctly price the probabilities of movements of the underlying asset? Journal of Econometrics, 102, 67–110.

Amin, K. I., & Jarrow, R. A. (1991). Pricing foreign currency options under stochastic interest rates. Journal of International Money and Finance, 10, 310–329.

Amin, K. I., & Ng, V. (1993). Options evaluation with systematic volatility. Journal of Finance, 48, 881–910.

Andrews, D., & Monahan, J. C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. Econometrica, 60, 953–966.

Bailey, W., & Stulz, R. (1989). The pricing of stock index options in a general equilibrium model. Journal of Financial and Quantitative Analysis, 24, 1–12.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. Journal of Finance, 52, 2003–2049.

Bakshi, G., & Chen, Z. (1997a). An alternative valuation model for contingent claims. Journal of Financial Economics, 44, 123–165.

Bakshi, G., & Chen, Z. (1997b). Equilibrium evaluation of foreign exchange claims. Journal of Finance, 52, 799–826.

Bates, D. S. (1991). The crash of '87: Was it expected? The evidence from option markets. Journal of Finance, 46, 1009–1044.

Bates, D. S. (1996a). Jumps and stochastic volatility: Exchange rate processes implicit in Deutschemark options. Review of Financial Studies, 9, 69–108.

Bates, D. S. (1996b). Testing option pricing models. In G. S. Maddala & C. R. Rao (Eds.), Handbook of statistics 14: Statistical methods in finance. Amsterdam: Elsevier Science B.V.

Bates, D. S. (2000). Post-'87 crash fears in the S&P 500 futures options markets. Journal of Econometrics, 94, 181–238.

Black, F. (1976). Studies of stock market volatility changes. Proceedings of the American Statistical Association, Business and Economics Section, 177–181.

Black, F., & Scholes, M. S. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81, 637–659.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (1997). Nonlinearities in the relation between the equity risk premium and the term structure. Management Science, 43, 371–385.

Brenner, M., & Eom, Y. H. (1997). No-arbitrage option pricing: New evidence on the validity of the martingale property (Working Paper S-97-10). Salomon Center, Leonard N. Stern School of Business, New York University, June.

Buraschi, A., & Jackwerth, J. (2001). The price of a smile: Hedging and spanning in options markets. Review of Financial Studies, 14, 495–527.

Campbell, J. Y. (1987). Stock returns and the term structure. Journal of Financial Economics, 18, 373–399.

Canina, L., & Figlewski, S. (1993). The information content of implied volatility. Review of Financial Studies, 6, 659–681.

Chernov, M., & Ghysels, E. (2000). A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of option valuation. Journal of Financial Economics, 56, 407–458.

Christie, A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. Journal of Financial Economics, 10, 407–432.

Cox, J. C., & Ross, S. A. (1976). The valuation of options for alternative stochastic processes. Journal of Financial Economics, 3, 145–166.

Davidian, M., & Carroll, R. J. (1987). Variance function estimation. Journal of the American Statistical Association, 82, 1079–1091.

Day, T., & Lewis, C. (1992). Stock market volatility and the information content of stock index options. Journal of Econometrics, 52, 267–287.

Derman, E., & Kani, I. (1994). Riding the smile. Risk, 7, 32–39.

Derman, E., Kani, I., & Chriss, N. (1996). Implied trinomial trees of the volatility smile. Journal of Derivatives, 3, 7–22.

Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical test. Journal of Finance, 53, 2059–2106.

Duprire, B. (1994). Pricing with a smile. Risk, 7, 18–20.

Engle, R. F., & Mustafa, C. (1992). Implied ARCH models from options prices. Journal of Econometrics, 52, 289–311.

Feinstein, S. (1989). The Black–Scholes formula is nearly linear in s for at-the-money options; Therefore, implied volatilities from at-the-money options are virtually unbiased (Working Paper 88-9). Federal Reserve Bank of Atlanta, December.

Fenton, V. M., & Gallant, A. R. (1996). Qualitative and asymptotic performance of SNP density estimators. Journal of Econometrics, 74, 77–118.

Ferson, W. E., Heuson, A., & Su, T. (2000). How much do expected stock returns vary over time? Answers from the options markets (Working Paper). University of Washington.

Fleming, J. (1998). The quality of market forecasts implied by S&P 100 index option prices. Journal of Empirical Finance, 5, 317–345.

Fleming, J. (1999). The economic significance of the forecast bias of S&P 100 index option implied volatility. Advances in Futures and Options Research, 10, 219–251.

Fleming, J., Ostdiek, B., & Whaley, R. (1995). Predicting stock market volatility: A new measure. Journal of Futures Markets, 15, 353–387.

French, K., & Roll, R. (1986). Stock return variances. Journal of Financial Economics, 17, 5–26.

Gallant, A. R., & Nychka, D. W. (1987). Seminonparametric maximum likelihood estimation. Econometrica, 55, 363–390.

Gallant, A. R., Rossi, P. E., & Tauchen, G. (1992). Stock prices and volume. The Review of Financial Studies, 5, 199–242.

Gallant, A. R., & Tauchen, G. (1989). Seminonparametric estimation of conditionally constrained heterogeneous processes: Asset pricing applications. Econometrica, 57, 1091–1120.

Gallant, A. R., & Tauchen, G. (1996a). SNP: A program for nonparametric time series analysis, version 8.5, user's guide (Working Paper). University of North Carolina at Chapel Hill, Chapel Hill, NC.

Gallant, A. R., & Tauchen, G. (1996b). Which moments to match. Econometric Theory, 12, 657–681.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics, 42, 27–62.

Green, C. T., & Figlewski, S. (1999). Market risk and model risk for a financial institution writing options. Journal of Finance, 54, 1465–1499.

Hamilton, J. D. (1988). Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. Journal of Economic Dynamics and Control, 12, 385–423.

Harrison, J. M., & Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory, 20, 381–408.

Heston, S. L. (1993). A closed form solution for options with stochastic volatility with applications to bonds and currency options. Review of Financial Studies, 6, 327–343.

Hull, J. C., & White, A. (1987). The pricing of options with stochastic volatilities. Journal of Finance, 42, 281–300.

Jackwerth, J. C., & Rubinstein, M. (1996). Recovering probability distributions from option prices. Journal of Finance, 51, 1611–1631.

Jackwerth, J. C., & Rubinstein, M. (2001). Recovering stochastic processes from option prices (Working Paper). Haas School of Business, University of California at Berkley, Berkley, CA.

Latané, H., & Rendleman, R. (1976). Standard deviation of stock prices implied in option prices. Journal of Finance, 31, 369–381.

Longstaff, F. A. (1995). Option pricing and the martingale restriction. Review of Financial Studies, 8, 1091–1124.

Madan, D., Carr, P., & Chang, E. (1998). The variance gamma process and option pricing. European Financial Review, 2, 79–105.

Melino, A., & Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. Journal of Econometrics, 45, 239–265.

Melino, A., & Turnbull, S. M. (1995). Misspecification and the pricing and hedging of long-term foreign currency options. Journal of International Money and Finance, 14, 373–393.

Merton, R. C. (1973). Theory of rational option pricing. Bell Journal of Economics, 4, 141–183.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics, 3, 125–144.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica, 59, 347–370.

Pan, J. (2001). The jump-risk premia implicit in options: Evidence from an integrated time-series study. Journal of Financial Economics, 63, 3–50.

Poteshman, A. M. (1998). Estimating a general stochastic variance model from option prices (Working Paper). University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, October.

Rubinstein, M. (1994). Implied binomial trees. Journal of Finance, 49, 771–818.

Schwarz, G. (1978). Estimating the dimensions of a model. Annals of Statistics, 6, 461–464.

Scott, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application. Journal of Financial and Quantitative Analysis, 22, 419–438.

Scott, L. O. (1997). Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Application of Fourier inversion methods. Mathematical Finance, 7, 413–426.

Scruggs, J. T. (1998). Resolving the puzzling intertemporal relationship between the market risk premium and conditional market variance: A two-factor approach. Journal of Finance, 53, 575–604.

Stein, E. M., & Stein, J. C. (1991). Stock price distributions with stochastic volatility, an analytic approach. Review of Financial Studies, 4, 727–752.

Stutzer, M. (1996). A simple nonparametric approach to derivative security valuation. Journal of Finance, 51, 1633–1652.

Whaley, R. E. (1986). Valuation of American futures options: Theory and empirical tests. Journal of Finance, 41, 127–150.

Whitelaw, R. F. (1994). Time variation and covariations in the expectation and volatility of stock market returns. Journal of Finance, 49, 515–541.

Wiggins, J. B. (1987). Option values under stochastic volatility: Theory and empirical estimates. Journal of Financial Economics, 19, 351–372.