Forecasting Future Volatility from Option Prices

Allen M. Poteshman*

December 6, 2000

* Department of Finance, College of Commerce and Business Administration, University of Illinois at Urbana-Champaign (poteshma@uiuc.edu). Helpful discussions with Narasimhan Jegadeesh, Neil Pearson, and Mike Weisbach are gratefully acknowledged. I bear full responsibility for all remaining errors.

Forecasting Future Volatility from Option Prices

Abstract

Evidence exists that option prices produce biased forecasts of future volatility across a wide variety of options markets. This paper presents two main results. First, approximately half of the forecasting bias in the S&P 500 index (SPX) options market is eliminated by constructing measures of realized volatility from five minute observations on SPX futures rather than from daily closing SPX levels. Second, much of the remaining forecasting bias is eliminated by employing an option pricing model that permits a non-zero market price of volatility risk.

It is widely believed that option prices provide the best forecasts of the future volatility of the assets which underlie them. One reason for this belief is that option prices have the ability to impound all publicly available information – including all information contained in the history of past prices – about the future volatility of the underlying assets. A second related reason is that option pricing theory maintains that if an option prices fails to embody optimal forecasts of the future volatility of the underlying asset, a profitable trading strategy should be available whose implementation would push the option price to the level that reflects the best possible forecast of future volatility.

A large empirical literature, however, suggests that option prices do a poor job of forecasting future volatility. In an important study Canina and Figlewski (1993) report that in the OEX (S&P 100 index) options market

> Implied volatility is an inefficient and biased forecast of realized future volatility that does not impound the information contained in recent historical volatility. In fact, the statistical evidence shows little or no correlation at all between implied volatility and subsequent realized volatility.

While other studies have not reached the extreme conclusion that option implied volatility contain essentially no information about future realized volatility, they almost uniformly find that option prices provide biased predictions of future realized volatility. This is the finding, for example, of Lamoureux and Lastrapes (1993) for options on individual stocks, of Fleming (1998) in a more recent study of the OEX options market, and of Jorion (1995) for options on foreign currency futures.

An influential interpretation of this failure is provided by Figlewski (1997) who maintains that investors may well be making optimal forecasts of future volatility but that these forecasts are not reflected in option prices because of frictions in the options markets. In particular, he argues that investors' optimal forecasts of future volatility may not be reflected in option prices because of either (1) the difficulty of carrying out the necessary arbitrage strategies that would force the prices to their proper levels or (2) the availability to market makers of lucrative alternative strategies in which they simply profit from the large bid-ask spreads in the options markets. Christensen and Prabhala's (1998) study of the volatility forecasting performance of OEX options is consistent with the Figlewski interpretation. They show that a statistical correction for error in the predicted volatility variable that is orthogonal to investors' true volatility predictions substantially improves the performance of option price based predictions of future realized volatility. This finding is compatible with the Figlewski interpretation insofar as the option market frictions to which he attributes the low quality of the option market forecasts may plausibly produce errors in measures of predicted volatility that bear no systematic relationship to investors' true predictions.

The present paper investigates whether either of two factors other than orthogonal error in the predicted volatility variable plays an important role in the poor quality of option price based forecasts of future volatility in the SPX (S&P 500 index) options market. The first factor is suboptimal measurement of the realized volatility of the underlying asset. The significance of this factor will be assessed by comparing volatility forecasting results when realized volatility is measured from daily closing data with those obtained when realized volatility is measured more precisely either from daily SPX futures data or from five minute interval SPX futures observations. The second factor is misspecification of the presumed option pricing model that produces systematic (i.e., non-orthogonal) error in the predicted volatility variable. The importance of this factor will be gauged by comparing the volatility forecasting results using the standard option pricing model which assumes that the market price of volatility risk is zero and that innovations to the level and volatility of the underlying asset are uncorrelated to those from an option pricing model which relaxes these restrictions.

Understanding the source of the poor volatility forecasting ability of option prices will be helpful in developing improved methods for predicting future volatility. It also has important implications for option pricing theory. It is well known that the Black-Scholes formula significantly misprices options. This pricing failure clearly results from the violation of one or more of the assumptions made by the Black-Scholes model. The Figlewski interpretation of the volatility forecasting literature suggests that the Black-Scholes assumptions of continuous trading, no bid-ask spreads, and perfect liquidity are especially problematic, because violation of these assumptions are likely to prevent the arbitrage trades that would force option prices to contain investors' optimal forecasts of future volatility and to encourage market makers to pursue strategies that profit from the bid-ask spread even if they believe options are mispriced. By contrast, much of the theoretical work on option pricing focuses on relaxing the assumption that the Black-Scholes model makes about the risk-neutral dynamics of the underlying asset. Understanding the source of the poor quality of volatility forecasts from option prices has the potential to guide theorists in the most useful direction for advances in option pricing theory.

This paper's volatility forecasting tests in the SPX market yields two main empirical findings. First, about half of the forecasting bias that is observed when standard testing procedures are implemented is eliminated when more precise measures of realized volatility are obtained from five-minute futures data. Second, most of the remaining bias is eliminated by extracting volatility predictions from option prices using an option pricing model that permits a non-zero market price of volatility risk and a non-zero correlation between innovations to the level and volatility of the underlying asset. The remaining bias is not significant. In addition to the empirical results, Monte Carlo simulations indicate that orthogonal error in the predicted volatility variable does not produce appreciable forecasting bias. The simulations also suggest that failure to account for a non-zero market price of volatility risk has the potential to produce forecasting bias of the size eliminated in the empirical work by using an option pricing model which does not restrict the price of volatility risk to be zero.

The remainder of the paper is organized as follows. Section I describes the standard approach to testing the volatility forecasting performance of option prices, reviews some of the important empirical findings, and discusses possible sources of the bias that has been reported in the literature. Section II describes the data. Section III reports the results of volatility forecasting tests under the conventional option pricing model both when standard data processing procedures are followed and when further error is removed from both the predicted volatility and the realized volatility variables. Section IV tests the volatility forecasting ability of option prices under an option pricing model that permits a non-zero market price of volatility risk and a non-zero correlation between innovations to the level and the volatility of the underlying asset. Section V performs a Monte Carlo simulation to study the impact of error in the data and of model misspecification on the results. Section VI concludes.

I. Volatility Forecasting from Option Prices: Standard Procedures and Previous Results

Studies of the volatility forecasting performance of option prices typically run the following two regressions

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \varepsilon(t) \tag{1}$$

and

$$Vol_{Realized}\left(t\right)=\alpha+\beta Vol_{Implied}\left(t\right)+\gamma Vol_{Historical}\left(t\right)+\varepsilon\left(t\right) \tag{2}$$

where $Vol_{Realized}\left(t\right)$ is a measure of the *ex-post* average realized volatility of the underlying asset over the period *t* to $t+T$, $Vol_{Implied}\left(t\right)$ is the average volatility forecast over the period *t* to $t+T$ implied from the price of an option observed at time *t* that expires at time $t+T$, $Vol_{Historical}\left(t\right)$ is the average volatility forecast over the period *t* to $t+T$ from some subset of the information available to the market at time *t* (usually from the history of past returns on the underlying asset), and $\varepsilon\left(t\right)$ is a zero mean forecast error that is uncorrelated with the forecasting variables.

The previous literature makes three major claims about the interpretation of these regressions. First, it maintains that option prices are *informative* about future volatility if and only if the estimate of β is significantly greater than zero. Second, it argues that option prices contain *biased* forecasts of future volatility if and only if the joint hypothesis of α being equal to zero and β being equal to one is rejected. Finally, it claims that option prices contain *informationally inefficient* estimates of future volatility if and only if the estimate of γ is statistically distinguishable from zero.

The extant studies on forecasting volatility from option prices are usefully broken into two periods. In the earlier period options had not been trading on exchanges for a long enough period of time to conduct time-series studies. Consequently, these earlier studies (Black and Scholes (1972), Latané and Rendleman (1976), Chiras and Manaster (1978), Beckers (1981)) examine a cross-section of options written on different underlying securities. More recent studies (Canina and Figlewski (1993), Christensen and Prabhala (1998), Day and Lewis (1992), Fleming (1998), Jorion (1995), Lamoureux and Lastrapes (1993), Day and Lewis (1993)), on the other hand, run the above regressions in a time-series framework using options written on only one underlying asset. Both types of studies generally find that option-implied forecasts of future volatility are *informative* but also that they are *biased*. The results on *informational efficiency* are mixed. The estimates from regression (1) and (2) (sometimes with the variables transformed by squaring or taking the natural logarithm) for some of the more recent time-series studies are presented in Table I. On the whole, the bias increases with the difficulty of performing arbitrage trades in a market to take advantage of mispricing with the bias greatest for options on stock indices and least for options on crude oil futures which trade side by side with the underlying futures in the same market. This observation provides some evidence in favor of the Figlewski interpretation discussed above. Note, however, that the relationship between the degree of bias and the difficulty of performing the arbitrage trades is somewhat loose.

The bias that is consistently found in studies of the volatility forecasting ability of option prices constitutes a rejection of the joint hypothesis of market efficiency and the option pricing model that is used to extract volatility forecasts from option prices. This rejection can be explained in four ways: (1) the market under study is inefficient; (2) improbable events are observed; (3) the econometric method utilized is flawed; and/or (4) the presumed option pricing model is misspecified. These possibilities will be discussed in turn.

According to the market inefficiency explanation for the bias in the volatility forecasting studies, the finding that β is statistically less than one indicates that options market forecasts of future volatility are more variable than actual realized future volatility. Consequently, in order to get the best fit in the regressions of realized volatility onto forecast volatility, the forecasts need to be dampened with a coefficient of less than one. Such forecasts are irrational since the time series of *ex ante* expected values of a quantity should always be less variable than the time series of *ex post* realized values of the quantity. It is possible that investors systematically make these types of irrational forecasts across a wide variety of options markets. However, given the high liquidity and large number of professional traders present in some of the markets that have been studied, a market inefficiency interpretation should be entertained only after every possible alternative explanation has been exhausted.

Another possible explanation for the observed bias in option price based forecasts of future volatility is simply that improbable events are observed in the various markets. According to this interpretation of the findings the *ex ante* volatility forecasts are unbiased and efficient, but the *ex post* observed volatilities are unusual (relative to their true distribution) so that it appears that the *ex ante* forecasts are biased or inefficient. Although this possibility cannot be excluded, it is unlikely that improbable events have been observed in each of the different markets that have been studied.

A third possible explanation for the forecasting bias documented in the literature is econometric problems. Most of the studies use daily data and options with horizons of one to three months. As a result, there is overlap in the error term which biases downward the usual OLS standard errors. Beginning with Canina and Figlewski (1993), researchers have typically used variants of the Hansen (1982) GMM method to correct the standard errors. Simulations in Canina and Figlewski (1993) and Jorion (1995) indicate that in the absence of error in the variables, the GMM corrections make it unlikely that the literature's conclusion that option-based forecasts of future volatility are biased is caused by faulty standard error calculations.

Another econometric problem that could spuriously produce the forecasting bias seen in the literature is orthogonal error in the forecasting variable in regressions (1) and (2). For simplicity, focus on specification (1). If there is orthogonal measurement error in the $Vol_{Implied}(t)$ variable, then it is well-known that the β estimate will be biased downward and that the bias can be corrected by an instrumental variable procedure (Johnston (1984), Chapter10). Christensen and Prabhala (1998) argue that orthogonal error could be present in the $Vol_{Implied}(t)$ variable, because of nonsynchronous measurement of option prices and index levels, early exercise and dividends, bid-ask spread, the wild-card option in the OEX market, and/or misspecification of the process governing index returns. When Christensen and Prabhala (1998) estimate equation (1) using data from the OEX option market and a sampling procedure in which the forecast intervals do not overlap, they get estimates of $\alpha = -0.56$ $(t_{\alpha=0} = -3.47)$ and $\beta = 0.76$ $(t_{1-\beta=0} = 2.99)$ (the latter *t*-statistic is derived from the reported *t*-statistic for $\beta = 0$.) When past implied volatility is used as an instrument, the instrumental variable estimates are $\alpha = -0.15$ $(t_{\alpha=0} = -0.63)$ and $\beta = 0.97$ $(t_{1-\beta=0} = 0.25)$. Similar results holds for regression equation (2).

Despite the Figlewski (1997) arguments and the Christensen and Prabhala (1998) results, there are a number of reasons to be uncertain as to whether orthogonal error in the forecasting variable is a major source of the biased forecasts of future volatility from option prices. First of all, the error in the $Vol_{Implied}(t)$ variable may well be correlated with its true value. By setting $Vol_{Implied}(t)$ equal to the Black-Scholes implied volatility of a close to at-the-money (henceforth, ATM) option, Christensen and Prabhala (1998) assume that the dynamics of the OEX index obey a stochastic volatility model in which innovations to the level of the stock price and innovations to the level of the volatility are uncorrelated and the market price of volatility risk is zero. It is widely accepted, however, that there is a substantial negative correlation between the level and volatility of the OEX index. In addition, there is evidence that the market price of volatility risk is nontrivial in the SPX market (Pan (1999) and Poteshman (2001)) which most likely has similar properties to the OEX market. Furthermore, the true process for the OEX index may have jumps or be non-Markovian. Any of these types of misspecification of the dynamics of the underlying asset are likely to produce error in the $Vol_{Implied}(t)$ variable which is correlated with its true value. In the presence of such errors the assumptions of instrumental variable estimation are violated. In particular, the instrument chosen by Christensen and Prabhala (1998) (last period's implied volatility) will probably be correlated with the error in the current period's implied volatility. Consequently, the fact that the instrumental variable procedure appears to remove the bias from the regression coefficients does not demonstrate that orthogonal error in the $Vol_{Implied}(t)$ variable causes the forecasting bias documented in the literature. Second of all, Christensen and Prabhala (1998) take the natural logarithm of the variables in equations (1) and (2) before running the regressions. It will be seen below that this transformation of the variables has a large impact on the coefficient estimates. Consequently, their findings may in part be driven by their decision to make this transformation rather than from orthogonal error in the $Vol_{Implied}(t)$ variable. Finally, simulations presented in Jorion (1995) suggest that plausible orthogonal error in the $Vol_{Implied}(t)$ variable does not produce a large enough downward bias in the estimate of the β coefficient to explain the empirical results.

The fourth and final possible explanation for the biased and sometimes inefficient nature of volatility forecasts from option prices that will be considered is that the presumed option pricing model is misspecified. It will be helpful to break model misspecification into two categories. The first category is misspecification of the risk-neutral dynamics of the underlying asset, and the second category is violation of any of the model's other assumptions (e.g. continuous trading, frictionless markets). Misspecification in the second category will tend to introduce orthogonal error into the forecasting variable which has already been considered in the discussion of econometric explanations. Consequently, the focus here will be on misspecification of the risk-neutral dynamics of the underlying asset.

The early studies and the Canina and Figlewski (1993) study use the Black-Scholes model to extract volatility forecasts from option prices. Accordingly, these studies assume that the underlying asset follows a geometric Brownian motion with constant volatility. The more recent time-series studies, however, use a more sophisticated stochastic volatility model. The model that these studies employ is described by the following equations:

$$\frac{dS_t}{S_t} = \mu(S_t, V_t, t)\,dt + \sqrt{V_t}\,dW_t^S \tag{3}$$

$$dV_t = \phi(V_t, t)\,dt + \xi(V_t, t)\,dW_t^V \tag{4}$$

$$Corr\left(dW_t^S, dW_t^V\right) = 0 \tag{5}$$

$$\lambda(S_t, V_t, t) = 0 \tag{6}$$

$$r = constant \tag{7}$$

where S_t is the level of the underlying asset, V_t is the instantaneous *variance* of the underlying asset, $\lambda(S_t, V_t, t)$ is the market price of volatility risk, and r is the instantaneous risk-free rate. As will be shown below, under this model the Black-Scholes implied volatility of an ATM option is almost exactly equal to the average volatility expected over the life of the option. The more recent time-series papers exploit this fact to extract volatility forecasts from model (3)-(7) by using the Black-Scholes implied volatility of close to ATM options as the forecast for the average volatility over the life of the option.

The question then is whether misspecification contained in equations (3)-(7) produces systematic errors in the predictions of future volatility from option prices that results in the observed forecasting bias. Some researchers have dismissed this possibility. Christensen and Prabhala (1998) write that

> Following Cox, Ingersoll and Ross (1985) and Lamoureux and Lastrapes (1993) the Black-Scholes implied volatility is approximately equal to expected future return volatility for at-the-money options (used in our study) even when returns follow the (non-Black-Scholes) stochastic volatility model of Hull and White (1987). Thus, misspecification error should be small and there is little reason for it to be correlated across time.

Fleming (1998), on the other hand, states that his results "should not be viewed as a test of market efficiency" largely because of a concern that model (3)-(7) might be misspecified. Finally, Lamoureux and Lastrapes (1993) clearly recognize that misspecification in model (3)-(7) may be influencing their results. In particular, they believe that the zero market price of volatility risk assumption contained in equation (6) could play an important role in explaining their findings. Indeed, they write that

> One possible reason for the rejection of the null is that volatility risk is priced. Therefore, further attempts to learn from the data should explicitly model a risk premium on the volatility process.

There are a number of ways that the dynamics contained in equations (3)-(7) might be misspecified that may have an important impact on the ability to forecast future volatility from option prices. Unreported simulation results indicate that the zero correlation assumption in equation (5) is not an important factor in producing the forecasting bias. An important open question, however, is whether the zero market price of volatility risk assumption contained in equation (6) plays an important role.

This section of the paper has argued that market inefficiency and unusual *ex post* draws out of the realized volatility distribution are not promising candidates for explaining the observed bias in option price based forecasts of future realized volatility. It has also suggested that there is not clear evidence that the forecasting bias comes from orthogonal error in the forecasting variable which is the preferred explanation of Christensen and Prabhala (1998) and is consistent with the Figlewski hypothesis. The remainder of this paper will investigate the extent to which the bias can be explained by errors in the variables that can be removed through more careful empirical procedures and an option pricing model that allows for a more general specification of the dynamics of the underlying asset.

II. Data

The SPX options market is studied, because it is one in which the arbitrage trades envisioned by standard option pricing models are likely to be difficult to execute. Indeed, Figlewski (1997) claims that

> Stock index options represent a polar case, where the arbitrage trade is complicated to execute at the outset, and the resulting position is both costly and risky to hedge over time. (p. 78)

Consequently, if more careful empirical procedures and a more general specification of the dynamics of the underlying asset produce unbiased forecasts in the SPX market, the difficulty of executing arbitrage trades probably does not prevent unbiased volatility forecasts in other options markets.

It is not obvious whether hedging SPX options is more or less difficult than hedging OEX options. On the one hand, the SPX index contains five times as many stocks as the OEX options, and its average stock is smaller. As a result, hedging the SPX option with its underlying asset is surely more difficult than hedging the OEX option with its underlying asset. On the other hand, both of these types of options can be hedged with SPX futures, and SPX options clearly can be better hedged than OEX options with SPX futures. SPX futures, however, only have expirations in March, June, September, and December. Accordingly, on most months neither SPX nor OEX options can be hedged with futures that have the same expiration date. Despite the fact that no definitive statement can be made about whether SPX or OEX options are more difficult to hedge, both markets should be among those where it is most difficult to implement option pricing theory's arbitrage strategy.

SPX options trade with expiration dates in the three near term months along with the following three months from the March expiration cycle (March, June, September, December). The options expire on the third Friday of the contract month. Strike price intervals are 5 points for near months and 25 points for far months. The minimum tick for options trading below $3.00 is 1/16 and 1/8 for options trading at higher prices. SPX options – unlike OEX options – are European and do not have a wildcard feature. Consequently, there is no possibility that the inability to properly account for early exercise or the wildcard feature biases the volatility forecasts from SPX options.

Bid-ask price quotes that are time-stamped to the nearest second were obtained directly from the CBOE for the period June 1, 1988 through August 29, 1997. Data are available from October 2, 1985 through August 29, 1997. The data from October 2, 1985 through May 31, 1988 is not used because of the evidence presented in Jackwerth and Rubinstein (1996) that there is a structural break in the SPX market at the time of the October 1987 stock market crash. In addition, the market was considerably less liquid during its earlier years, so the data may not be as reliable. Consequently, tests that use the Black-Scholes model to extract predictions of future volatility from option prices will use data from the period June 1, 1988 through August 29, 1997. The Heston (1993) model is a standard model for pricing options that includes a non-zero market price of volatility risk and a constant correlation between innovations to the level and instantaneous variance of the underlying asset. This model will be used to extract volatility forecasts from option prices in Section IV below. When the Heston model is employed, the data period will be June 1, 1993 through August 29, 1997 in order to avoid the assumption that investors were forecasting volatility with a model that had not yet been developed.

The SPX options trade on the CBOE from 8:30 AM to 3:15 PM Central Standard Time. As in Bakshi, Cao and Chen (1997), on each trade date, for each strike price and time-to-expiration information is extracted on the last bid-ask quote reported prior to 3:00 PM Central Standard Time. For each of these bid-ask quotes, the bid-ask midpoint, the time-to-expiration, the exercise price, and the type of option (i.e. call or put) are recorded. Before August 24, 1992 the time-to-expiration is measured as the number of calendar days from the trade date to the Friday before the third Saturday of the expiration month. After August 23, 1992 the time-to-expiration is measured as the number of calendar days from the trade date to the Thursday immediately preceding the Friday before the third Saturday of the expiration month. This method of computing time-to-expiration is followed, because prior to August 24, 1992 SPX options expired at the close of trading and since August 24, 1992 SPX options have expired at the open of trading (see Dumas, Fleming and Whaley (1998).) The risk-free rate of interest associated with each option is the one month LIBOR rate on the day the option bid-ask quote is observed. Table II provides descriptive statistics on ATM and OTM (i.e., out-of-the-money) options over the period June 1993 through August 1997 which are the options of greatest interest in this paper. In particular, the descriptive statistics cover those options from June 1, 1993 through August 29, 1997 that (1) have time-to-expiration of greater than or equal to six calendar days and less than or equal to 7/12 of a year, (2) have a bid price of greater than or equal to $3/8 and a bid-ask spread of less than or equal to $1, (3) have a Black-Scholes implied volatility greater than zero and less than or equal to 0.70, and (4) are ATM or OTM where ATM is defined as the strike price being equal to the CRSP closing value of the SPX index. ATM is defined this way only for purposes of inclusion or exclusion in Table II and to select options for estimating the Heston model in Section IV below. Throughout the rest of the paper ATM is defined as the strike price being equal to the futures price.

The fact that SPX options are European and do not have a wildcard feature considerably simplifies empirical work with them. Nonetheless, two serious challenges remain. The first challenge is matching observations on option prices to observations on the underlying index level. Even if a quote on the underlying index can be exactly matched temporally to an option price, the quoted index level will not be the proper underlying value for the option, since all 500 underlying stock prices will not correspond to trades that occurred at the quoted time. The second challenge is determining the *expected* future rate of dividend payments by the stocks that compose the index until the expiration of an option. Of course, dividend rates can always be calculated after the fact from the actual dividends paid out by the SPX stocks. The *ex-post* rate, however, may not match the *ex-ante* expectation at the time the option is priced.

When addressing these two challenges, it is not necessary to determine the contemporaneous underlying value and dividend rate separately. It is sufficient to determine the quantity $Se^{-\delta T}$, where δ is the dividend rate paid by the index and T is the time to expiration of the option. If an accurate value for $Se^{-\delta T}$ can be determined, then the option can be priced under the assumption that the underlying value is equal to $Se^{-\delta T}$ and that the dividend rate is zero. (This is because a stock with a continuous dividend yield δ which grows from S at time 0 to S_T at time T will grow from $Se^{-\delta T}$ at time 0 to S_T at time T if it pays no dividend.) The quantity

$Se^{-\delta T}$ will be determined from transactions data on SPX futures which trade on the Chicago Mercantile Exchange (CME).

Data on SPX futures were obtained from the Futures Industry Institute which is the official data supplier for the CME. The SPX futures trade with delivery dates in March, June, September, and December. On a given trade date there is typically one delivery date which is extremely liquid (i.e. thousands of transactions per day) and a second delivery date which is moderately liquid (i.e. on the order of thirty transactions per day). Consequently, the quantity $Se^{-\delta T}$ associated with an option bid-ask quote will be determined from the futures data in one of two ways depending on whether for a given trade date the delivery date of the most liquid futures contract is the same as the expiration date of the option.

The SPX futures contracts expire at the open of the trading day throughout the time period studied in this paper. Taking this into account, when a given option bid-ask quote observation corresponds to an option that expires at the time of delivery of the most liquid futures contract, the transaction price of the futures contract which trades at the time closest to the observation of the option bid-ask quote is used in conjunction with spot-futures parity to determine the quantity $Se^{-\delta T}$. In particular, spot-futures parity is used to determine $Se^{-\delta T}$ from

$$Se^{-\delta T} = Fe^{-rT} \tag{8}$$

where F is the futures prices, r is the one month LIBOR rate, and T is the time to expiration of the option (or, equivalently, the time to delivery of the futures) computed as specified above.

When, on the other hand, an option corresponding to a bid-ask quote expires at a time other than that of the most liquid futures contract, the following procedure is used to compute $Se^{-\delta T}$. First, three futures are identified. Let F_1 be the transaction price of the futures contract with the most liquid delivery date that transacts closest in time to the option bid-ask quote. Let

F_2 be the transaction price of the futures contract with the second most liquid delivery date that transacts closest in time to the option bid-ask quote. Let F_3 be the transaction price of the futures contract with the most liquid delivery date that transacts closest in time to F_2. Finally, let T_1, T_2, and T_3 be the time to delivery of, respectively, F_1, F_2, and F_3. Then the dividend rate is computed via spot-futures parity from F_2 and F_3 (which are typically observed within a few seconds of each other) by

$$\delta = r - \frac{\log\left(F_3/F_2\right)}{\left(T_3 - T_2\right)} \tag{9}$$

This dividend rate is then used in conjunction with F_1 (which is typically observed within a few seconds of the option bid-ask quote) and spot-futures parity to compute

$$Se^{-\delta T} = F_1 e^{\left(-(r-\delta)*T_1 - \delta * T\right)} \tag{10}$$

which is associated with the bid-ask quote for the option that expires at a time T in the future.

The price of a European option that expires at a time T in the future which is written on an asset with price S that pays a continuous dividend stream δ is the same as the price of the corresponding European option on an asset with price $Se^{-\delta T}$ that pays no dividends. Hence, all of the volatility forecasts from option prices used in this paper's tests will be formed by first determining the quantity $Se^{-\delta T}$ from the appropriate futures price(s) as described above, and, then extracting a volatility forecast from the option price by assuming that the level of the underlying asset is the derived value of $Se^{-\delta T}$ and that the underlying asset pays no dividends.

III. Forecasting Volatility from Option Prices using the Standard Option Pricing Model

Most of the recent literature on the quality of volatility forecasts from option prices uses an option pricing model in which the dynamics of the underlying asset are described by equations (3)-(7). This section of the paper will report the results of forecasting volatility from SPX options under this model. Different aspects of the regression specification and data processing procedure will be varied one at a time in order to determine their impact on the volatility forecast regression results.

The sampling procedure of Christensen and Prabhala (1998) is adopted throughout the paper. SPX options expire on the Friday before the third Saturday of each month. Following Christensen and Prabhala (1998), prices of options that expire the next calendar month are observed on the Wednesday after each expiration date. The September 1992 observation date is missing, because short term option prices are missing from the CBOE data for the required Wednesday date and for several trade dates on either side of the required date. The option prices are used to predict the future volatility until the options expire – approximately three and a half to four and a half weeks later. Since this sampling procedure results in no overlap at all in the data, there is no need to account for overlapping data when making statistical inferences.

The realized volatility of the SPX index over some time period from t to $t+T$ is computed as follows. Let $\Delta = T/N$ be a fixed interval of trading time (measured in years) and let $\left\{S_t, S_{t+\Delta,} S_{t+2\Delta}, ..., S_{t+N\Delta}\right\}$ be the $N+1$ evenly spaced (in trading time) observed levels of the SPX index over the time period from t to $t+T$. Define the return from time $t+(i-1)\Delta$ to time

$t + i\Delta$ to be $R_{t+i\Delta} \equiv \ln\left(S_{t+i\Delta}/S_{t+(i-1)\Delta}\right)$ for $i = 1,\ldots,N$. The annualized realized volatility over the period from t to $t+T$ is then defined as

$$Vol_{Realized}\left(t\right) \equiv \sqrt{\frac{1}{\Delta}\left[\frac{1}{N}\sum_{i=1}^{N}\left(R_{t+i\Delta}\right)^2\right]} = \sqrt{E\left[Var_{Realized}\left(t\right)\right]}. \tag{11}$$

The quantity underneath the square root symbol is an unbiased estimate of the annualized variance of the returns over the period t to $t+T$ when the mean return is constrained to be zero. (The denominator of the inner fraction is N rather than the more usual $N-1$ because no information is used up calculating the mean return.) Since the square root function is concave, Jensen's inequality implies that the quantity $Vol_{Realized}\left(t\right)$ is an upwardly biased estimate of the annualized volatility of the returns over the period t to $t+T$. Back of the envelope type calculations like those provided in footnote 10 of Fleming (1998) or in the text associated with footnote 13 of Green and Figlewski (1999) suggest that this bias is not large in the present context. This issue will be addressed in the Monte Carlo simulations in Section V below.

In order to determine the volatility forecast over the period t to $t+T$ implied from option prices under model (3)-(7), note that Hull and White (1987) show that under model (3)-(7)

$$O_t = \int_0^{\infty} O^{BS}\left(\sigma = \sqrt{\overline{V}}\right) f\left(\sqrt{\overline{V}}\middle|V_t\right) d\sqrt{\overline{V}} = E\left[O^{BS}\left(\sigma = \sqrt{\overline{V}}\right)\middle|V_t\right] \tag{12}$$

where

$$\overline{V} \equiv \frac{1}{T}\int_t^{t+T} V_t dt, \tag{13}$$

O_t is the time t price of a European option, $O^{BS}\left(\sigma = \sqrt{\overline{V}}\right)$ is the Black-Scholes price of the option when volatility is equal to $\sqrt{\overline{V}}$, and $f\left(\sqrt{\overline{V}}\middle|V_t\right)$ is the probability density of $\sqrt{\overline{V}}$

conditional on V_t, the level of instantaneous variance at time *t*. Now Feinstein (1988) shows that the Black-Scholes formula is nearly linear in volatility for close to ATM options (where ATM is defined as the strike price being equal to the futures price) which implies that

$$E\left[O_{ATM}^{BS}\left(\sigma=\sqrt{\overline{V}}\right)\middle|V_t\right]\approx O_{ATM}^{BS}\left(\sigma=E\left[\sqrt{\overline{V}}\middle|V_t\right]\right). \tag{14}$$

Combining equations (12) and (14) gives

$$O_t^{ATM}\approx O_{ATM}^{BS}\left(\sigma=E\left[\sqrt{\overline{V}}\middle|V_t\right]\right) \tag{15}$$

or

$$\sigma_{ATM}^{BSImpl}\approx E\left[\sqrt{\overline{V}}\middle|V_t\right] \tag{16}$$

where O_t^{ATM} is the time *t* price of an ATM option under model (3)-(7) and σ_{ATM}^{BSImpl} is the Black-Scholes implied volatility of an ATM option. Equation (16) indicates that under model (3)-(7) the Black-Scholes implied volatility of an ATM option is approximately equal to the expected square root of the average variance over the life of the option. The quality of this approximation is governed by the quality of the approximation in equation (14). In fact, the Black-Scholes function is strictly concave in volatility so that Jensen's inequality biases the expectations derived from Black-Scholes implied volatilities. This bias, however, is small, and it too will be investigated in the Monte Carlo simulations in Section V.

For the first set of tests the realized volatility variable, $Vol_{Implied}\left(t\right)$, is set equal to the Black-Scholes implied volatility of the closest to ATM call at observation time *t* where ATM is defined as the strike price being equal to the futures price. For purposes of finding the closest to ATM call the futures price is taken to be $Se^{(r-\delta)T}$ which is obtained by multiplying the quantity $Se^{-\delta T}$, which is derived from the futures prices as described in the previous section, by e^{rT}

where r is the one month LIBOR rate. When implying the Black-Scholes volatility from the option price, the level of the SPX index is set equal to $Se^{-\delta T}$, and it is assumed that the underlying index pays no dividends.

The initial set of tests uses three different time-series of SPX levels to compute the realized volatility variable via equation (11). The first time-series of SPX levels for computing $Vol_{Realized}\left(t\right)$ is the daily closing values for the SPX index from the CRSP indices files. The second times-series is the daily 3:00 PM SPX level implied from the SPX futures transaction prices. The third time-series is the five minute trading time (i.e., on each trade the 8:35 AM CST, 8:40AM CST, …, 3:15 PM CST) SPX index levels implied from SPX futures transaction prices. For the second and third time-series, the SPX level for a particular target time is inferred from the SPX futures prices as follows. First a dividend rate, δ, for the SPX index is computed from equation (9) above where as before F_2 is the transaction price of the futures contract with the second most liquid delivery date that transacts closest in time to the target time, and F_3 is the transaction price of the futures contract with the most liquid delivery date that transacts closest in time to F_2. Next F_1 is chosen to be the transaction price of the SPX futures with the most liquid delivery date that transacts closest in time to the target time. Then spot-futures parity is used to infer the SPX level at the target time from the futures transaction data by

$$S = F_1 e^{-(r-\delta)T_1} \tag{17}$$

where T_1 is the time to delivery of F_1 and r is the one month LIBOR rate.

The results of estimating equation (1) using the three different time-series of SPX levels to construct $Vol_{Realized}\left(t\right)$ are presented in Table III . For these tests $Vol_{Implied}\left(t\right)$ is set equal to the Black-Scholes implied volatility of the closest to ATM call. Panel A contains the results for the June 1993 through August 1997 data period which will be compared with the results in the next section of the paper where the Heston (1993) model is employed. Panel B contains the results for the June 1988 through August 1997 data period.

Focusing on Panel A which covers the June 1993 through August 1997 data period, it can be seen that as the time-series of SPX index levels used to construct the $Vol_{Realized}(t)$ variable changes from (1) CRSP daily SPX closing levels to (2) daily 3:00 PM SPX levels implied from futures prices to (3) five minute interval SPX levels implied from futures prices the α estimates move monotonically toward zero and the β estimates move monotonically toward 1. Altogether the α estimate moves 70% of the way toward zero and the β estimate moves about 50% of the way toward one when the realized volatility is computed from five minute futures implied SPX index levels rather than from SPX daily closing levels. In addition, the adjusted R^2 increases from 0.40 to 0.51 and the *F*-statistic for testing the joint hypothesis that $\alpha = 0$ and $\beta = 1$ decreases from 22.29 to 7.37. The *F*-statistic, nonetheless, still rejects the joint hypothesis at conventional levels. The reported *t*-statistics indicate that none of the α estimates are significantly different than zero while the β estimate is statistically different than one only when the realized volatility is computed from CRSP daily closing levels. Panel B of Table III shows that when the data period is June 1988 through August 1997 similar – although a bit less marked – improvement in the volatility forecasting performance of option prices also occurs as the realized volatility is computed in cleaner ways.

The improvement in forecasting performance when moving from daily SPX closing prices to daily 3:00 PM SPX levels inferred from futures prices is to be expected, because the daily closing prices contain positive serial correlation from infrequent trading of some components of the index which is not present in the SPX levels inferred from futures prices. The improvement in the forecasting performance when moving from daily 3:00 PM SPX levels inferred from futures prices to five minute trading time SPX levels inferred from futures prices is also not surprising, since it is well known that more frequent sampling of a time-series improves estimates of its volatility (but not its mean). Bevan, Poon and Taylor (1999) also report improvements in a forecasting context from using high frequency index returns to compute realized volatility.

If the remaining bias in the volatility forecasts from option prices is produced by orthogonal error in the forecast (i.e. independent) variable, then constructing the $Vol_{Implied}(t)$ variable in a cleaner manner should reduce the bias. Although there are a number of possible sources of orthogonal error in $Vol_{Implied}(t)$, the large bid-ask spreads in the SPX options market are most likely an important contributor. (For a one month close to ATM SPX option it is not unusual for the bid-ask spread to be more than five percent of its value.) Consequently, the next set of tests will examine the impact of reducing the orthogonal error in $Vol_{Implied}(t)$ produced by the bid-ask spread by computing $Vol_{Implied}(t)$ from several rather than one close to ATM options.

Table IV presents the results of estimating equation (1) when the $Vol_{Implied}(t)$ variable is constructed from the closest to ATM call, the closest to ATM put, the two closest to ATM calls, the two closest to ATM puts, and the four closest to ATM options (i.e., the two closest to ATM calls and the two closest to ATM puts.) In all cases, the $Vol_{Implied}(t)$ variable is set equal to the Black-Scholes implied volatility of the option(s). When there is more than one option, this volatility is defined as the real number that minimizes the sum of the squared pricing errors of the options (i.e., it is *not* the average of the individual Black-Scholes implied volatilities.) Here and for the remainder of the paper, the $Vol_{Realized}(t)$ variable is constructed from five minute trading time observations on the SPX level inferred from futures prices. From Panel A of Table IV, which covers the period June 1993 through August 1997, it can be seen that there is very little difference in the estimates when puts are used instead of calls or when the implied volatility is extracted from multiple option prices rather than just one. For example, when $Vol_{Implied}(t)$ is computed from the four closest to ATM options instead of from just the closest to ATM call, the α estimate changes from 0.003 to 0.001, the β estimate changes from 0.87 to 0.89, the R^2 increases from 0.51 to 0.52, and the *F*-statistic decreases from 7.37 to 6.14. These limited improvements in the forecasting regression estimates suggest that orthogonal error in the $Vol_{Implied}(t)$ variable plays at most a secondary role in producing the bias in the regression estimates. There is some indication in Panel B of Table IV that over the entire data period from June 1988 through August 1997 call prices provide better forecasts of future volatility than put prices. However, there is no improvement in the regression results in Panel B when two rather than one ATM calls are used to construct $Vol_{Implied}(t)$ or when two rather than one ATM puts are used to construct $Vol_{Implied}(t)$. This lack of improvement also suggests that orthogonal error in $Vol_{Implied}(t)$ does not play a substantive role in biasing the option price based forecasts of future volatility.

Next the regression specification in equation (1) will be more carefully considered. It can be viewed as a special case of the regression

$$f\left(Vol_{Realized}(t)\right) = \alpha + \beta f\left(Vol_{Implied}(t)\right) + \varepsilon(t) \tag{18}$$

where $f(\bullet) = I(\bullet)$. Table I indicates that the previous literature has also employed the specifications $f(\bullet) = (\bullet)^2$ and $f(\bullet) = \ln(\bullet)$. In order to assess the impact of these choices on the regressions, Table V presents the results of estimating equation (18) for all of these specifications. Here and for the remainder of the paper, $Vol_{Implied}(t)$ is set equal to the volatility value that minimizes the sum of the Black-Scholes pricing errors for the two closest to ATM calls and two closest to ATM puts. The results are presented in order of increasing concavity of the $f(\bullet)$ function: $f(\bullet)=(\bullet)^2$, then $f(\bullet)=I(\bullet)$, and, finally, $f(\bullet)=\ln(\bullet)$. It is apparent from Table V that the specification chosen has an appreciable effect on all of the statistics that are presented. For example, in Panel A which covers the June 1993 through August 1997 data period, using the specification $f(\bullet)=\ln(\bullet)$ rather than $f(\bullet)=(\bullet)^2$ changes the α estimate from 0.001 to –0.068, the β estimate from 0.78 to 1.03, the adjusted R^2 from 0.47 to 0.55, and the *F*-statistic from 4.88 with a *p*-value of 0.01 to 9.36 with a *p*-value of 0.00 . Similar differences are observed in Panel B for the June 1988 through August 1997 time period.

The results in Table V show that the choice of $f(\bullet)$ makes a difference. Furthermore, since the Black-Scholes formula is nearly linear in volatility (rather than in variance or the logarithm of volatility) the theory outlined above requires the specification $f(\bullet)=I(\bullet)$ in order to make use of the approximation

$$E\left[O_{ATM}^{BS}\left(\sigma=\sqrt{\overline{V}}\right)\middle|V_t\right]\approx O_{ATM}^{BS}\left(\sigma=E\left[\sqrt{\overline{V}}\middle|V_t\right]\right) \tag{19}$$

in equation (14). The fact that the theory calls for the $f(\bullet)=I(\bullet)$ specification does not, of course, entail that the $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables meet the assumptions for linear regression. The most important of these assumptions are that there is a linear relationship between the variables and that the error term in the regression has a spherical distribution. Accordingly, the appropriateness of linear regression for the $f(\bullet)=I(\bullet)$ specification will be checked by testing for a linear relationship between the $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables and by testing the regression residuals for serial correlation and heteroskedasticity.

A RESET(2) test is used to check whether there is a linear relationship between the $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables. This test augments the specification (1) by adding a regressor which is the square of the predicted value from the original specification (1). The test involves determining the significance of the coefficient of the new regressor. The RESET(2) test is used because of Monte Carlo evidence that the test outperforms a number of alternatives in small samples and when the error term has a non-normal distribution. See Godfrey, McAleer and McKenzie (1988) for details on the test and for this Monte Carlo evidence. The distribution of the RESET(2) statistic is $F(1, n-3)$ where n is the number of observations in the time-series. The RESET(2) statistics for the shorter data period June 1993 through August 1997 and for the longer data period June 1988 through August 1997 are, respectively, 0.8120 which corresponds to a *p*-value of 0.63 and 0.4937 which corresponds to a *p*-value of 0.52. Consequently, a linear relationship between the $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables is not rejected at conventional levels. In order to test for serial correlation in the residuals from equation (1), Durbin-Watson statistics are computed for both the shorter and longer data period. The value of the statistics are, respectively, 2.12 and 2.04. Since neither is statistically different from 2, there is no evidence that the residuals are autocorrelated. Finally, Breusch-Pagan tests are conducted to check for heteroskedasticity in the residuals. For the shorter data period, the Breusch-Pagan statistic is 0.7187 which corresponds to a *p*-value of 0.40 while for the longer data period, the Breusch-Pagan statistic is 2.2636 which corresponds to a *p*-value of 0.13. Hence, there is no evidence that the residuals are heteroskedastic. This finding is confirmed by the fact that running GLS to correct for any heteroskedasticity does not change the coefficient estimates to the precision to which they are reported in Table V. Since the underlying option pricing theory requires that $f(\bullet) = I(\bullet)$ and the resulting variables conform to the main assumptions for linear regression, the rest of the analysis in this paper uses this specification.

The next issue that will be addressed is the efficiency of the volatility forecasts extracted from option prices via model (3)-(7). The efficiency tests will be carried out by running the "encompassing" regression described by equation (2). Before running these regressions, however, it will be useful to establish the forecasting ability of volatility predictions obtained from historical data on the SPX index by estimating the following equation:

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Historical}(t) + \varepsilon(t) \tag{20}$$

As before, $Vol_{Realized}(t)$ is the realized volatility from time *t* to time $t+T$ where *T* is the time to expiration of the option observed at time *t* according to the sampling scheme described above. $Vol_{Historical}(t)$ is a measure of the historical volatility of the SPX index for some period of time leading up to time *t*. For example, if the historical data period is one month, then $Vol_{Historical}(t)$ is calculated via equation (11) from the five minute interval SPX index levels inferred from SPX futures prices from twenty-two trade dates before time *t* to one trade date before time *t*.

The results of estimating equation (20) for historical data periods of one to six months are presented in Table VI. Overall, historical volatility's ability to forecast future volatility appears to be comparable to that of option price forecasts based on the model described by equations (3)-(7). Indeed, for the shorter data period of June 1993 through August 1997, the *F*-statistics do not reject the joint hypothesis $\alpha = 0$ and $\beta = 1$ at conventional levels. A GARCH(1,1) model was also estimated for the past 500 trade dates of past returns and used to predict future volatility. In unreported results, the volatility forecasts from the GARCH(1,1) models performed very poorly.

However, no conclusion about the forecasting ability of GARCH models is drawn from this finding, because no effort was made to determine either the optimal GARCH specification or the optimal amount of past data to estimate the GARCH model.

Table VII contains the estimates of equation (2) when the historical volatility, $Vol_{Historical}(t)$, is computed from 1, 2, 3, or 6 months of past data. The coefficient estimates on the $Vol_{Implied}(t)$ variable drop slightly compared to its value in the univariate regression while the coefficient estimates on $Vol_{Historical}(t)$ drop dramatically from their values in the univariate regressions. In fact, for the data period June 1993 through August 1997 all of the regressions have an intercept that is not statistically different from zero, a coefficient on $Vol_{Implied}(t)$ that is not statistically different from one, and a coefficient on $Vol_{Historical}(t)$ that is not statistically different from zero. Hence, the individual coefficient estimates are consistent with market efficiency given the model of market equilibrium described by equations (3)-(7). Nonetheless, the joint hypothesis of market efficiency and model (3)-(7) is rejected at conventional levels, because the probability that the *F*-statistics are as large as observed is only one percent if $\alpha = 0, \beta = 1,$ and $\gamma = 0.$ Furthermore, in the longer data period from June 1988 through August 1997 reported in Panel B, most of the β estimates are significantly different from one.

The tests presented in this section of the paper indicate that approximately half of the bias found when using standard procedures to forecast the future volatility of the SPX index from SPX option prices under model (3)-(7) can be eliminated by more carefully constructing the realized volatility variable. A statistically significant bias remains, however, even when the cleaner empirical techniques are used in the volatility forecasting exercise. The next section investigates the impact on the forecasting regressions of using the Heston option pricing model to construct the implied volatility variable.

IV. Forecasting Volatility from Option Prices using the Heston Model

The existing literature on forecasting future volatility from option prices uses either the Black-Scholes model or model (3)-(7) to extract forecasts of future volatility from option prices. This section of the paper will examine the impact of extracting future volatility forecasts from the Heston (1993) model instead. The Heston model is chosen because it generalizes equations (3)-(7) in two important ways and has been available to option market participants since the summer of 1993.

The dynamics for the underlying asset presumed by the Heston model are described by the following set of equations:

$$\frac{dS_t}{S_t} = \mu\left(S_t, V_t, t\right)dt + \sqrt{V_t}\,dW_t^S \tag{21}$$

$$dV_t = k\left(\theta - V_t\right)dt + \eta\sqrt{V_t}\,dW_t^V \tag{22}$$

$$Corr\left(dW_t^S, dW_t^V\right) = \rho \tag{23}$$

$$\lambda\left(S_t, V_t, t\right) = \lambda V_t \tag{24}$$

$$r = \delta = constant \tag{25}$$

where $k, \theta, \eta, \rho, \lambda, r,$ and δ are constants. As before, S_t and V_t are, respectively, the time t level and instantaneous variance of the underlying asset. The differentials dW_t^S and dW_t^V are increments of standard Weiner processes with constant correlation ρ. The market price of volatility risk is denoted $\lambda(S_t, V_t, t)$, r is the constant rate for risk free borrowing and lending, and δ is the constant dividend payout rate of the underlying asset.

The Heston model is more flexible than the model described by equations (3)-(7) insofar as it allows a non-zero correlation between the Weiner processes driving the system and a non-zero market price of volatility risk. This greater flexibility is potentially significant. The non-zero correlation may be important, because there is known to be a substantial negative correlation between the level and volatility of the SPX index (i.e., the leverage effect). The non-zero market price of volatility risk may be important, because it will bias the forecasting tests in the observed direction if investors need to be compensated for bearing volatility risk and the required compensation increases as the level of volatility increases. It should be noted that despite these two advantages, the Heston model – unlike the model described by equations (3)-(7) – restricts the drift and diffusion coefficients of the variance process to particular parametric forms.

Recently, models have appeared with even more realistic dynamical assumptions (e.g., the inclusion of jumps in the return process) than the Heston model. These models are not used, however, because they were not available to investors during the data period. Indeed, the research design will follow this literature's practice (see Figlewski (1997)) of using only models and data that were known to be available to investors at a given time when forming volatility forecasts from option prices at that time. Accordingly, it will be assumed that on each monthly observation date from June 1993 through August 1997 SPX option market investors estimate the $k, \theta,$ and η parameters of the variance process given by equation (22) from the daily closing levels of the SPX index in the CRSP indices file over the period July 3, 1962 (the first date available in the CRSP file) through the end of the month preceding the observation date. Over this period, an instantaneous variance is estimated for each calendar month by the sample variance entailed by the SPX daily closing levels for that month (i.e., the square of the expression in equation (11) is used so that the mean return is still constrained to be zero.) The resulting time-series of monthly instantaneous variance estimates is then used to estimate the parameters in equation (22) via maximum likelihood. The necessary likelihood function is provided by equations (18) and (20) in Cox, Ingersoll and Ross (1985). The month of October 1987 is omitted from the time-series, because no diffusion model can reasonably accommodate the October 1987 crash. Estimating the parameters in this way yields $k = 6.87$, $\theta = 0.0170$, and $\eta = 0.44$ for the period July 1963 through May 1993. These estimates do not change by much as the months June 1993 through July 1997 are added to the estimation period.

Once the $k, \theta,$ and η parameters have been determined for a given observation date, ρ and λ are chosen to minimize the sum of the squared pricing errors of all out-of-the-money SPX options on that observation date that meet several criteria. The out-of-the money SPX options must have a Black-Scholes implied volatility greater than zero and less than 0.70, a bid price of greater than 3/8, a bid-ask spread that is less than or equal to one dollar, and a time to expiration of greater than six calendar days and less than or equal to seven months. The price of each option is taken to be the midpoint of the last bid-ask quote prior to 3:00 PM CST. When determining ρ and $\lambda,$ the instantaneous volatility is also allowed to vary in order to minimize the sum of squared pricing errors of the out-of-the money options.

After ρ and λ are calculated, however, a new instantaneous volatility is determined by finding the level of instantaneous volatility which minimizes the sum of squared pricing errors of the four closest to ATM options on the observation date (i.e., the same two closest to ATM calls and two closest to ATM puts that are used in the tests in Section III) given the estimated values of $k, \theta, \eta, \rho,$ and λ. $Var_{Instantaneous}(t)$ is then set equal to the square of the instantaneous volatility derived from the four closest to ATM options. The forecast variable $Vol_{Implied}(t)$ is defined as the expected average instantaneous volatility over the remaining life of the options given $Var_{Instantaneous}(t)$, the estimated $k, \theta,$ and η parameters, and the assumption that instantaneous variance process obeys

$$dV_t = k(\theta - V_t)dt + \eta\sqrt{V_t}dW_t^V. \tag{26}$$

It is shown in the Appendix that under these conditions, $Vol_{Implied}(t)$ is equal to

$$Vol_{Implied}(t) = \sqrt{\theta + \left[\frac{1-e^{-kT}}{kT}\right]\left[Var_{Instantaneous}(t) - \theta\right]}. \tag{27}$$

$Vol_{Realized}(t)$ is computed as before from five minute trading time observations on the SPX index level derived from futures prices.

Estimating equation (1) via OLS when $Vol_{Implied}(t)$ is computed from the Heston model using the method just described results in an α estimate of –0.006 $(t_{\alpha=0} = -0.37)$, and a β estimate of 0.96 $(t_{1-\beta=0} = 0.27)$ where OLS standard errors are used to compute the *t*-statistics. Hence, individually the estimates of α and β are not statistically different than 0 and 1 respectively. The adjusted R^2 is 0.51, and the *F*-statistic is 3.76 which corresponds to a *p*-value of 0.03 for the joint hypothesis $\alpha = 0$ and $\beta = 1$. Consequently, the hypothesis that option price based forecasts of future volatility are unbiased is marginally rejected at conventional levels. This stands in contrast to the tests that use model (3)-(7) to compute $Vol_{Implied}(t)$ in which this hypothesis is clearly rejected. Furthermore, the simulations presented in the next section suggest that in the present context there is actually a seven to 19 percent chance of observing an *F*-statistic of 3.76 or larger.

Even if it is stipulated that the Heston model provides the proper description of market equilibrium, there are still two viable interpretations of the improvement in volatility forecasting. One interpretation maintains that the λ parameter estimates represent only investor aversion to volatility risk. In this case, the use of the Heston model results in option price based forecasts of future volatility that properly reflect investors' correct volatility forecasts. Another interpretation, however, holds that investors make systematic errors in forecasting future volatility which are reflected in option prices. These forecasting error then will also be reflected in the λ parameters, since they are estimated from the option prices. In this case, the use of the Heston model still results in correct option price based forecasts of future volatility, but these forecasts no longer match investors' incorrect volatility forecasts.

Table VIII presents the results of estimating equation (2) when $Vol_{Implied}(t)$ is computed under the Heston model and the historical volatility prediction is computed from one, two, three, or six months of past data. The coefficient estimates on the $Vol_{Implied}(t)$ variable drop slightly compared to its value in the univariate regressions while the coefficient estimates on $Vol_{Historical}(t)$ drop dramatically from their values in the univariate regressions. In fact, the α estimates are never statistically different from zero, the β estimates are never statistically different from one, and the γ estimates are never statistically different from zero. In addition, the *p*-values for the *F*-statistics are between 0.05 and 0.07 which are border line for rejecting the joint hypothesis $\alpha = 0, \beta = 1,$ and $\gamma = 0$ at conventional levels. Unreported simulation similar to those that are presented in the next section suggest that in the present context there is actually a nine to 20 percent chance of observing *F*-statistics as larger or larger than those reported in Table

VIII. Hence, when the cleaner data processing techniques used in this paper are combined with the use of the Heston model to predict future volatility from option prices, there is no longer evidence that SPX option prices contain biased or inefficient forecasts of future volatility.

V. Monte Carlo Simulations

All of the point estimates and standard errors (i.e. *t*-statistics) in the tests presented above were computed using OLS. Although the non-overlapping sampling procedure eliminates the most important source of inaccuracies in the OLS standard errors, the point estimates or standard errors may, nonetheless, be biased by errors in the variables or persistence in the volatility process. The potential impact of these factors will be assessed by conducting Monte Carlo simulations in order to construct the sampling distributions of the OLS estimates under various assumptions about the dynamics of the underlying asset and the errors present in the options and the futures data that are used to generate the $Vol_{Realized}(t)$, $Vol_{Implied}(t)$, and $Vol_{Historical}(t)$ variables.

The simulations assume that the true dynamics of the underlying SPX index level, *S*, and SPX instantaneous variance, *V*, are described by equations (21)-(25). The drift of the index return is set to $\mu(S_t, V_t, t) = 0.124$ which is the continuously compounded rate of return that would increase the SPX index from its level of 209.59 at the beginning of 1986 to its level of 899.47 at the end of August 1997. In order to calibrate the variance process and the correlation between the Weiner processes, the $k, \theta, \eta, \lambda,$ and ρ parameters are estimated from the real data using the method described in the previous section on each of the 2442 trade dates from January 4, 1988 through August 29, 1997. These parameters are then set equal to the average of the resulting time-series which yields the calibration $k = 4.07$, $\theta = 0.0181$, $\eta = 0.506$, and $\rho = -0.761$. For later reference, note that the mean λ value is –3.521.

Each Monte Carlo experiment uses this calibration to simulate 2000 sample paths of index levels and instantaneous variances at five minute trading time intervals for a period of time that corresponds to the opening of trading on January 2, 1986 through the close of trading on September 18, 1997 which is the expiration time for the final option that is observed in August 1997. Each path is simulated from equations (21)-(23) with a bivariate Euler scheme from a starting index level of 209.59 and a starting variance level equal to the long-run mean of 0.0181. Taking account of holidays, late openings, and early closings, there were 244,634 five-minute trading intervals in the SPX futures market from the opening of trading on January 2, 1986 through the close of trading on September 18, 1997. Consequently, the sample paths are obtained by simulating equations (21)-(23) from the initial values for 244,634 five minute steps. In order to assess the impact of the discretization bias introduced by the use of the Euler scheme, several of the experiments were repeated simulating equations (21)-(23) for 2,446,340 steps using a bivariate Euler scheme where each step corresponded to 30 seconds of trading time. The simulated paths were then sampled at a frequency of ten steps to obtain five minute trading time paths. The results were nearly identical to those reported below.

Throughout the simulations the interest rate is set to $r = 0.059$ which is the average of daily one month LIBOR rates over the January 1988 through August 1997 period, and the dividend rate is set to $\delta = 0.025$. Various simulations are run with and without error introduced into the SPX index levels and option prices from which the realized, historical, and implied volatilities are computed. Error is introduced into the SPX index levels as follows. For each of the 244,634 steps on each simulated path, spot-futures parity is used to convert the simulated

SPX level into a futures price for each of the three delivery times that were actually used at the corresponding time during the tests on the real data. A band with width of either \$0.00, \$0.10, or \$0.20 is then centered on each of the spot-futures parity derived futures prices, and a price for each of the three futures is drawn uniformly from its band. These three futures prices are then used to derive an SPX level (with error) as in Sections III and IV above. Error is introduced into the option prices as follows. On each observation date (i.e., at each point on the simulated paths that corresponds to 3:00 PM on the Wednesday after an options expiration date during the period June 1993 through August 1997) a target ATM strike price level is set equal to the futures price computed from spot-futures parity given the current simulated level of the SPX index (without error), the time to expiration of the options on that date in the actual data, r, and δ. The strike prices of the two closest to ATM calls and two closest to ATM puts are then set to values that give them a displacement from this target ATM strike price level that is the same as the displacement of the actual option strike prices from the actual target ATM strike price level on the corresponding date in the real data. Next, the true prices of these options are set to their correct prices under the Heston model given the assumed parameters values. A band with width of either zero, the bid-ask spread of the corresponding call or put in the real data observed on the same trade date, or twice that bid-ask spread is then centered on each of the true options prices, and a price for each of the four options is drawn uniformly from its band.

This procedure is used to repeatedly simulate five minute interval SPX index levels with error as well as the four option prices with error at the 51 time points that correspond to 3:00 PM on the Wednesdays following option expirations over the period June 1993 through August 1997. For each simulation, the $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables are calculated at the 51 time points. The $Vol_{Realized}(t)$ variable is computed from the five minute SPX index levels using equation (11). The $Vol_{Implied}(t)$ variable is set equal to the volatility level that minimizes the sum of squared Black-Scholes pricing errors of the four simulated option prices. Equation (1) is estimated from the $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables for each simulated path, and Table IX reports fractiles, means, standard deviations, and empirical *p*-values for the associated statistics when the simulation and estimation is conducted 2000 times. The different panels of the table correspond to various assumptions about the errors in the futures prices, the errors in the options prices, and the market price of volatility risk. The empirical percentile is the percentage of times the observed statistic exceeds the simulated statistic under the panel's assumptions. The empirical 1-sided *p*-value is the percentage of times that the simulated statistic is both on the same side of the statistic's null value as the observed statistic and further away from the statistic's null value than the observed statistic. The empirical 2-sided *p*-value is the percentage of times that the simulated statistic is further away (in absolute value) from the statistic's null value than the observed statistic. The null values of the statistics are $\alpha = 0$, $\beta = 1$, $R^2 = 1$, and $F = 0$.

Panels A-E of Table IX report the simulation results when the market price of volatility risk is set to zero under various assumptions about the errors in the futures and option prices. Since the SPX futures prices move in increments of \$0.05 and the market is very liquid, it is assumed that observed SPX futures prices for the most part have errors in the range -\$0.05 to \$0.05 which corresponds in the simulations to a futures error band of \$0.10. In Panels A-E the futures error band is set to a width of \$0.00, \$0.10, or \$0.20. In the SPX options market the bid-ask spreads are wide. For example, an ATM short maturity option that trades for around \$5.00 typically has a bid-ask spread of about \$0.50. It will be assumed that for the most part observed SPX option prices have errors within the range of their bid-ask spreads. (In other words, in the empirical work when the observed option price is set to the bid-ask midpoint, it is assumed that the true option price without error usually falls somewhere within the bid-ask spread.) This corresponds in the simulations to an options error band of 1*x* the actual bid-ask spread. In Panels A-E the options error band is set either to 0, 1*x*, or 2*x* the actual bid-ask spread.

In Panel A the futures and options error bands are both set equal to zero. This results in simulated mean values of $\alpha = 0.007,$ $\beta = 0.95,$ $R^2 = 0.75,$ and $F = 1.26.$ When there is no error in the data and the market price of volatility risk is set equal to zero, the simulated forecasts outperform the real volatility forecasts (whose results are repeated in the observed statistic column) on each of these statistics except for the intercept. The deviation of the mean α and β values from their null values of, respectively, 0 and 1 presumably results from a combination of the non-linearity of the Black-Scholes formula in volatility for ATM options, the fact that the options used are not exactly ATM, and bias in the estimator of the realized volatility. The empirical percentiles for the observed α and β values of 0.24 and 0.21, respectively, do not indicate rejection of their null values at conventional levels. The simulated distribution of the *F*-statistic, however, indicates that there is a less than 1 percent chance that the observed statistic would be as large as it is under the assumptions of the simulation in Panel A. The standard deviations for the α and β statistics are, respectively, 0.008 and 0.08, while the mean OLS standard errors are 0.009 and 0.07. The closeness of these numbers suggests that the use of OLS standard errors are appropriate in this context. This conclusion continues to be true when there is error in the futures or options prices and when the market price of volatility risk is non-zero.

Panel B adds orthogonal error into the options data equal to a uniform draw out of the bid-ask spread of the options used in the actual empirical work. Adding this error into the option prices has almost no impact on the distributions of the simulated statistics. The small impact that it does have, however, is in the direction that would be expected. For example, the standard deviation of the *F*-statistic estimate increases from 1.25 to 1.26. These simulation results make it seem unlikely that orthogonal error in the independent variable is an important source of the volatility forecasting bias observed in the literature as suggested by Christensen and Prabhala (1998). They also raise questions for the Figlewski interpretation of the volatility forecasting bias to the extent that the option pricing errors produced either by the difficulty of executing an arbitrage strategy or by market makers being focused on making a profit from the bid-ask spread are unsystematic. Panel C presents the results of simulations where there is a uniform $0.10 error in the futures prices and no error in the options prices. Although the impact on the distributions of the statistics is a bit larger than for the case where there is error only in the option prices, it is still quite small. Panel D presents the results of simulations where there is error in both the options and futures prices, and Panel E presents the results when the error band for the options prices is doubled to twice the actual bid-ask spread and the error band for the futures prices is doubled to $0.20. The main results are much the same. These errors in the variables do not make a substantial difference in the distributions of the statistics. It appears unlikely that they can account for the bias in the estimates of equation (1) observed in this paper or in the literature.

Panels F-I report the results of simulations under the various assumptions about the options and futures errors when the market price of volatility risk is assumed to be $\lambda = -3.521$. When the options and future errors are zero in Panel F, the non-zero market price of volatility risk has very little impact on the distribution of the α estimate. The mean value of the simulated β coefficient, on the other hand, drops from 0.95 to 0.89 which is the value of the observed β estimate. The mean simulated *F*-statistic increases from 1.26 to 2.55. Although this is still less than the observed value of 6.14, seven percent of the simulated *F*-statistics are now greater than 6.14. Hence, if investors' price of volatility risk is approximately $\lambda = -3.521$, the observed regression estimates are not unusual. The simulation results are consistent with the β coefficient decreasing from its null value of 1 to 0.95 because of non-linearity of the Black-Scholes formula in volatility and biases in the estimation of the realized volatility and then decreasing from 0.95 to 0.89 because of the failure of the Black-Scholes formula to include a non-zero market price of volatility risk. Panels G-I add orthogonal error into the futures prices, the options prices, or both. These errors do not have an important impact on the distribution of the statistics.

Table X repeats the simulation experiments from Table IX with the single difference that the $Vol_{Implied}(t)$ variable is now computed from the Heston model using the four options on each of the 51 observation dates. Estimating the Heston model parameters on each observation date in the manner in which it was done in the empirical work would be of interest. This is not computationally feasible in a Monte Carlo setting, however, because it requires repeatedly minimizing over several variables an objective function that sums the squared pricing errors under the Heston model of typically over 100 options per observation date. Instead, $Vol_{Implied}(t)$ is computed from the same four option prices as in the simulations reported in Table IX conditional on the Heston model parameters used to generate the data. In particular, on each of the observation dates, the level of instantaneous variance is found which minimizes under the Heston model with the data generating parameters the sum of the squared pricing errors of the four option prices (which, in general, contain error). This level of instantaneous variance is then plugged into equation (27) to compute $Vol_{Implied}(t)$. By eliminating uncertainty about the parameters of the Heston model, this procedure reduces the variability of the distribution of the regression statistics. This reduced variability makes the simulations more conservative insofar as it biases them toward rejecting investor rationality.

Panels A-E of Table X report the simulation results for various levels of orthogonal error in the futures and options prices when the market price of volatility risk is zero. The α and β values both move closer to their null values of 0 and 1. For example, comparing Panels A of Tables IX and X, the mean α value goes from 0.007 to –0.001 and the mean β value goes from 0.95 to 0.97 when the $Vol_{Implied}(t)$ variable is computed from the Heston model rather than the Black-Scholes model. One source of the improvement is that the $Vol_{Implied}(t)$ variable computed from the Heston model is not biased by any non-linearity of the Heston option pricing formula in the instantaneous volatility. The mean *F*-statistic is 1.70 which is smaller than the observed value of 3.76. Nonetheless, 11 percent of the simulated *F*-statistic values are greater than 3.76. This suggests that the *p*-value of 0.03 computed in Section IV is too small. Indeed, if uncertainty as to the value of the parameters of the Heston model were added into the simulation, it is likely that the simulated *F*-statistic would have a more variable distribution and that even more than 11 percent of the simulated values would be greater than 3.76. The same conclusion follows from all of the panels in Table X regardless of whether there is orthogonal error in the futures or options prices (or both) or whether the market price of volatility risk is zero or non-zero. Consequently, these simulations suggest that the α and β estimates from the real data under the Heston model do not reject the joint hypothesis of $\alpha = 0$ and $\beta = 1$ at conventional levels.

As in Table IX, adding orthogonal error into the futures or options prices in Panels B-E does not lead to appreciable changes in the distributions of the regression statistics. Once again, it seems implausible that orthogonal error in the variables leads to appreciable bias in the estimation of the forecast regressions. It should also be noted that a comparison of the standard deviations of the α and β distributions and the means of the OLS α and β standard errors indicates that OLS provides appropriate standard errors in this setting regardless of whether there is orthogonal error in the variables.

Panels F-J repeat the analysis of Panels A-E with the market price of volatility risk set to $\lambda = -3.521$. The distributions of the statistics in Panels F-J is nearly identical to that in Panels A-E. This fact indicates that the use of the Heston model satisfactorily accounts for the non-zero market price of volatility risk.

The simulation experiments reported in this section of the paper indicate that orthogonal error in the futures or options prices that are used to construct the variables for the forecast regressions has very little impact on the regression statistics. In particular, it seems unlikely that the biases introduced by such error in the variables can account for the forecasting regression results in this paper or that have been reported elsewhere in the literature. It does appear that a small amount of the bias from the forecast regressions results from the procedure that is used to construct the realized volatility variable as well as from non-linearity of the Black-Scholes formula in the volatility argument. When these factors are accounted for in addition to a non-zero market price of variance risk, the distributions of the regression coefficients no longer reject the hypothesis that option market participants make unbiased forecasts of future volatility. Simulations like those presented in this section were also conducted for regression equation (2) which includes the $Vol_{Historical}(t)$ variable on the right hand side. These simulations produce the same general conclusions as those that follow from Tables IX and X.

VI. Conclusion

A large literature studying the ability of option prices to forecast the future volatility of the underlying asset concludes that such forecasts are systematically biased, and the bias tends to be most severe for options written on equity indexes. The prevailing interpretation of these findings is that option investors may be forming unbiased forecasts of the future volatility of underlying assets but that these unbiased forecasts fail to get impounded into option prices because of either (1) the difficulty in carrying out the necessary arbitrage strategies that would force the prices to their proper level, or (2) the availability of lucrative alternative strategies for market makers which profit from the large bid-ask spreads in the option markets. A recent study bolsters this interpretation by arguing that orthogonal error in the forecasts of future volatility are a substantial source of the standard bias. This interpretation has significant consequences for the entire range of research into option pricing, because it implies that non-continuous trading, bid-ask spreads, and other market imperfections substantially influence option prices. This implication is important because incorporating these types of market imperfections into option pricing models is much more difficult than, for example, altering the risk-neutral dynamics of the underlying asset.

The present paper studies the volatility forecasting ability of SPX option prices against the backdrop of the standard interpretation of the findings in the volatility forecasting literature. There are two main findings. First, approximately half of the forecasting bias in the SPX market is eliminated by constructing measures of realized volatility from five minute observations on SPX futures rather than from daily closing SPX levels. Second, much of the remaining forecasting bias is eliminated by employing an option pricing model that permits a non-zero market price of volatility risk. In addition to the empirical findings, a Monte Carlo study that accounts for errors in the futures and option prices that are used to construct the measures of realized and forecast volatility indicates that the regression coefficients obtained in the standard forecasting regressions are not unlikely to occur if option prices are set by investors who optimally forecast future volatility and subscribe to a standard option pricing model that allows a non-zero market price of volatility risk.

Appendix A. Derivation of Equation (27)

This appendix derives equation (27) in the text. Given the instantaneous variance at time t, V_t, the problem is to find the expected average volatility over the time interval t to $t+T$ under the assumption that the instantaneous variance follows the stochastic differential equation (22). Let $\bar{V}$ be the average instantaneous variance over the time interval t to $t+T$

$$\bar{V} = \frac{1}{T}\int_t^{t+T} V_\tau d\tau. \tag{A1}$$

Then the expected average instantaneous variance at time t over the interval t to $t+T$ is given by

$$E_t\left[\bar{V}\right] = \frac{1}{T} E_t\left[\int_t^{t+T} V_\tau d\tau\right] \tag{A2}$$

or

$$E_t\left[\bar{V}\right] = \frac{1}{T}\left[\int_t^{t+T} E_t\left[V_\tau\right] d\tau\right] \tag{A3}$$

where the last equation follows from an application of Lebesques's Dominated Convergence Theorem. Now equation (19) in Cox, Ingersoll and Ross (1985) states that for $\tau > t$

$$E_t\left[V_\tau\right] = V_t e^{-k(\tau-t)} + \theta\left(1-e^{-k(\tau-t)}\right). \qquad \text{(A4)}$$

Substituting into the previous equation gives

$$E_t\left[\bar{V}\right] = \frac{1}{T}\left[\int_t^{t+T} V_t e^{-k(\tau-t)} + \theta\left(1-e^{-k(\tau-t)}\right) d\tau\right]. \qquad \text{(A5)}$$

Integrating and simplifying yields

$$E_t\left[\bar{V}\right] = \theta + \left[\frac{1-e^{-kT}}{kT}\right]\left[V_t - \theta\right]. \qquad \text{(A6)}$$

Finally, if $Var_{Instantaneous}(t)$ is the instantaneous variance implied at time *t* from option prices, then taking the square root of both sides gives

$$Vol_{Implied}(t) = \sqrt{E_t\left[\bar{V}\right]} = \sqrt{\theta + \left[\frac{1-e^{-kT}}{kT}\right]\left[Var_{Instantaneous}(t) - \theta\right]} \qquad \text{(A7)}$$

which is equation (27) in the text.

References

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, 2003-2049.

Beckers, S., 1981, Standard deviations implied in option prices as predictors of future stock price variability, *Journal of Banking and Finance* 5, 363-381.

Bevan, Blair, Ser-Huang Poon, and Stephen J. Taylor, 1999, Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high frequency index returns, Working Paper, Lancaster University.

Black, Fischer, and Myron Scholes, 1972, The valuation of option contracts in a test of market efficiency, *Journal of Finance* 27, 399-417.

Canina, L., and S. Figlewski, 1993, The informational content of implied volatility, *Review of Financial Studies* 6, 659-681.

Chiras, D. P., and S. Manaster, 1978, The information content of option prices and a test of market efficiency, *Journal of Financial Economics* 6, 213-234.

Christensen, B. J., and N. R. Prabhala, 1998, The relation between implied and realized volatility, *Journal of Financial Economics* 50, 125-150.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385-407.

Day, T., and C. Lewis, 1992, The information content of option prices and a test of market efficiency, *Journal of Econometrics* 52, 267-287.

Day, Theodore, and Craig Lewis, 1993, Forecasting futures market volatility, *Journal of Derivatives* 1, 33-50.

Dumas, Bernard, Jeff Fleming, and Robert E. Whaley, 1998, Implied volatility functions: Empirical tests, *Journal of Finance* 53, 2059-2106.

Feinstein, S., 1988, A source of unbiased implied volatility forecasts, Working Paper, Federal Reserve Bank of Atlanta.

Figlewski, Stephen, 1997, Forecasting volatility, *Financial Markets, Institutions & Instruments* 6, 1-88.

Fleming, Jeff, 1998, The quality of market volatility forecasts implied by S&P 100 index option prices, *Journal of Empirical Finance* 5, 317-345.

Godfrey, L. G., Michael McAleer, and C. R. McKenzie, 1988, Variable addition and Lagrange multiplier tests for linear and logarithmic regression models, *Review of Economics and Statistics* 70, 492-503.

Green, Clifton T., and Stephen Figlewski, 1999, Market risk and model risk for a financial institution writing options, *Journal of Finance* 54, 1465-1499.

Hansen, Lars P., 1982, Large sample properties of generalized method of moment estimators, *Econometrica* 50, 1029-1054.

Heston, S. L., 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327-344.

Hull, John, and Alan White, 1987, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42, 281-300.

Jackwerth, Jens Carsten, and Mark Rubinstein, 1996, Recovering probability distributions from option prices, *Journal of Finance* 51, 1611-1631.

Johnston, J., 1984, *Econometric methods, 3rd edition* (McGraw-Hill, Inc., New York).

Jorion, Philippe, 1995, Predicting volatility in the foreign exchange market, *Journal of Finance* 50, 507-528.

Lamoureux, D., and W. Lastrapes, 1993, Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities, *Review of Financial Studies* 6, 293-326.

Latané, H. A., and R. J. Rendleman, 1976, Standard deviations of stock price ratios implied in option prices, *Journal of Finance* 31, 369-381.

Pan, Jun, 1999, "Integrated" time-series analysis of spot and option prices, Working Paper, Graduate School of Business, Stanford University.

Poteshman, Allen, 2001, Underreaction, overreaction and increasing misreaction to information in the options market, *Journal of Finance* 56, ???-???.

Willard, Gregory A., 1997, Calculating prices and sensitivities for path-independent derivative securities in multifactor models, *Journal of Derivatives* 5, 45-61.

Table I
Summary of Previous Volatility Forecasting Studies

$$f\left(Vol_{Realized}(t)\right)=\alpha+\beta f\left(Vol_{Implied}(t)\right)+\varepsilon(t)$$

$$f\left(Vol_{Realized}(t)\right)=\alpha+\beta f\left(Vol_{Implied}(t)\right)+\gamma f\left(Vol_{Historical}(t)\right)+\varepsilon(t)$$

This table reports the results of option price based volatility forecasting regressions from a selection of previous studies. The $Vol_{Realized}(t)$ variable is a measure of the *ex-post* average realized volatility of the underlying asset over the period *t* to $t+T$. The $Vol_{Implied}(t)$ variable is the volatility that is forecast over the period *t* to $t+T$ from an option price observed at time *t* that expires at time $t+T$. The $Vol_{Historical}(t)$ variable is the volatility forecast over the period *t* to $t+T$ from some subset of the information available to the market at time *t* – usually from the history of past returns on the underlying asset. The error term $\varepsilon(t)$ has zero mean and is uncorrelated with the forecasting variables. The *t*-statistics (reported in parentheses) for $\hat{\alpha}$ and $\hat{\gamma}$ test for equality to zero, and the *t*-statistics for $\hat{\beta}$ test for equality to one. (The *t*-statistics provided here are sometimes derived from standard errors or *t*-statistics that test a different hypothesis.) Some of the reported values are averages over different data sets examined in the paper. Fleming (1998) runs orthogonality (i.e. market efficiency) regressions. They are not reported here because they do not fit conveniently into the format of the table. The function $I(\bullet)$ is the identity function, i.e., $I(x)=x$.

Study	Underlying Asset	$f(\bullet)$	Observations	Forecast Horizon	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	R^2
Canina and Figlewski (1993)	OEX	$I(\bullet)$	Daily Calls, 3/83-3/87	7-127 Cal. Days	0.14 (11.3)	0.02 (19.6)		0.00
					0.08 (2.91)	-0.06 (12.47)	0.49 (2.21)	0.17

Table I – Continued

Study	Underlying Asset	$f(\bullet)$	Observations	Forecast Horizon	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	R^2
Day and Lewis (1992)	OEX	$(\bullet)^2$	Weekly Calls, 3/83-12/89	Nearby > 7 Days	0.0005 (0.00)	0.72 (0.14)		0.03
					-0.0001 (-0.00)	0.60 (0.39)	0.30 (0.71)	0.03
Fleming (1998)	OEX	$I(\bullet)$	Daily Calls & Puts, 10/85-4/92	Nearby > 15 Days	-0.0191 (-3.63)	0.57 (9.30)		0.03
					NA	NA	NA	NA
Christensen and Prabhala (1998)	OEX	$\ln(\bullet)$	Monthly Calls, 11/83-5/95	One Month	-0.56 (-3.47)	0.76 (2.99)		0.39
					-0.49 (-3.01)	0.56 (3.76)	0.23 (2.38)	0.41
Lamoureux and Lastrapes (1993)	Individual Stocks	$(\bullet)^2$	Daily Calls, 4/82-3/84	90-180 Cal. Days	NA	NA		NA
					1627.258 (4.20)	0.67 (0.90)	-1.75 (-3.13)	0.29
Jorion (1995)	Foreign Ex. Futures	$I(\bullet)$	Daily Calls & Puts, 1/85 or 7/86 –2/92	3-100 Cal. Days	0.35 (2.7)	0.52 (2.9)		0.13
					0.33 (2.65)	0.63 (1.99)	-0.09 (-0.89)	0.14
Day and Lewis (1993)	Crude Oil Futures	$(\bullet)^2$	Daily Closing Prices, 11/86-3/91	2 and 4 Months	0.03 (1.0)	0.88 (0.8)		0.72
					0.004 (1.14)	0.97 (0.19)	-0.01 (-0.01)	0.61

Table II
Descriptive Statistics for CBOE Traded S&P 500 Index Option Calls and Puts, June 1, 1993 through August 29, 1997

Descriptive statistics for daily observations on CBOE traded S&P 500 index options for the period June 1, 1993 through August 28, 1997 that meet four conditions: (1) The time-to-expiration is greater than or equal to 6 calendar days and less than or equal to 7/12 year; (2) The bid price for the option is greater than or equal to $ 3/8, and the bid-ask spread is less than or equal to $1; (3) The Black-Scholes implied volatility (BSIVol) is greater than zero and less than or equal to 0.7; and (4) The option is out-of-the-money which is determined by comparing the strike price of the option to the CRSP closing value for the S&P 500 index on the trade date. The price information for each option on each trade date is the last bid-ask quote prior to 3:00 PM CST. The call and put prices are defined as the bid-ask midpoint. The interest rate, *r*, is the one month LIBOR rate on the day that the option is observed. The moneyness of an option is defined as its strike price divided by its futures price, $Moneyness \equiv X/F$.

Variable	Obs.	Mean	S.D.	Min	Percentiles 1%	10%	50%	90%	99%	Max
Panel A: Calls										
Call Price ($)	56717	6.935	7.046	0.406	0.438	0.875	4.750	15.625	34.50	71.25
Spread ($)	56717	0.411	0.229	0.063	0.063	0.125	0.375	0.750	1.000	1.000
BSIVol (%)	56717	0.147	0.043	0.023	0.085	0.100	0.140	0.201	0.296	0.696
T (Cal. Days)	56717	90	55	7	9	24	77	176	209	212
X (Ind. Pnts.)	56717	503	178	245	265	305	460	785	970	995
r (%)	56717	0.059	0.018	0.029	0.030	0.032	0.057	0.084	0.099	0.103
Moneyness	56717	1.043	0.045	0.940	0.988	1.002	1.033	1.091	1.217	1.467
Panel B: Puts										
Put Price ($)	97575	5.071	4.944	0.406	0.438	0.750	3.563	11.375	23.125	55.500
Spread ($)	97575	0.347	0.200	0.000	0.063	0.125	0.250	0.625	1.000	1.000
BSIVol (%)	97575	0.199	0.058	0.080	0.110	0.137	0.189	0.271	0.397	0.700
T (Cal. Days)	97575	89	56	7	8	24	76	176	210	212
X (Ind. Pnts.)	97575	471	163	175	215	275	430	720	880	960
r (%)	97575	0.056	0.017	0.029	0.030	0.032	0.055	0.083	0.099	0.103
Moneyness	97575	0.914	0.065	0.527	0.700	0.826	0.929	0.982	0.996	1.005

Table III
Forecasting Regressions Varying Data Used to Construct Realized Volatility

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \varepsilon(t)$$

OLS point estimates with *t*-statistics for $\alpha = 0$ and $1 - \beta = 0$ (in parentheses). The *t*-statistics are computed from OLS standard errors. The $Vol_{Realized}(t)$ variable is the realized volatility of the SPX returns computed under the restriction that the mean return is equal to zero. The SPX returns are derived either from the daily CRSP SPX closing levels, the 3:00 PM SPX levels inferred from SPX futures prices, or five minute interval SPX index levels inferred from SPX futures prices. The $Vol_{Implied}(t)$ variable is the Black-Scholes implied volatility of the closest to ATM call computed by setting the level of the SPX index to the quantity $Se^{-\delta T}$ (which is computed from futures prices) and assuming that the SPX index pays no dividends. The *F*-statistic is used to test the joint hypothesis that $\alpha = 0$ and $\beta = 1$. The *p*-value gives the probability that the *F*-statistic is as large as it is if the hypothesis that $\alpha = 0$ and $\beta = 1$ is true. There is one observation for each Wednesday that follows an option expiration date, and the interval covered by each observation extends from that Wednesday to the next option expiration.

Realized Volatility Constructed From	α	β	Adj. R^2	*F*-statistic	*p*-value
Panel A: June 1993 – August 1997					
SPX Close	0.010	0.72	0.40	22.29	0.00
	(0.62)	(2.28)			
3:00 PM Futures	0.005	0.81	0.41	10.01	0.00
	(0.28)	(1.40)			
Five Minute Futures	0.003	0.87	0.51	7.37	0.00
	(0.17)	(1.13)			
Panel B: June 1988 – August 1997					
SPX Close	0.010	0.74	0.42	45.76	0.00
	(0.81)	(3.13)			
3:00 PM Futures	0.007	0.81	0.36	16.51	0.00
	(0.49)	(1.88)			
Five Minute Futures	0.006	0.83	0.51	21.47	0.00
	(0.51)	(2.10)			

Table IV
Forecasting Regressions Varying Options Used to Construct Implied Volatility

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \varepsilon(t)$$

OLS point estimates with *t*-statistics for $\alpha = 0$ and $1 - \beta = 0$ (in parentheses). The *t*-statistics are computed from OLS standard errors. The $Vol_{Realized}(t)$ variable is the realized volatility of the SPX returns derived from five minute interval SPX index levels inferred from SPX futures prices. The realized volatility is computed under the restriction that the mean return is equal to zero. The $Vol_{Implied}(t)$ variable is the Black-Scholes implied volatility from either the closest to ATM call, the closest to ATM put, the two closest to ATM calls, the two closest to ATM puts, or the two closest to ATM calls and the two closest to ATM puts. The Black-Scholes implied volatility is computed by using the quantity $Se^{-\delta T}$, inferred from futures prices, as the value of the SPX index level and then assuming that the SPX index pays no dividends. When the Black-Scholes implied volatility is computed from more than one option, it is defined as the real number that minimizes the sum of the squared pricing errors of the options (i.e., it is *not* the average of the individual Black-Scholes implied volatilities). The *F*-statistic is used to test the joint hypothesis that $\alpha = 0$ and $\beta = 1$. The *p*-value gives the probability that the *F*-statistic is as large as it is if the joint hypothesis is true. There is one observation for each Wednesday in the data period that follows an option expiration date, and the time interval covered by each observation extends from that Wednesday to the next option expiration date.

Options Used to Construct Implied Volatility	α	β	Adj. R^2	*F*-statistic	*p*-value
Panel A: June 1993 – August 1997					
ATM Call	0.003 (0.17)	0.87 (1.13)	0.51	7.37	0.00
ATM Put	0.006 (0.38)	0.86 (1.19)	0.51	5.26	0.01

Table IV – Continued

Options Used to Construct Implied Volatility	α	β	Adj. R^2	*F*-statistic	*p*-value
One ATM Call and One ATM Put	0.003 (0.21)	0.87 (1.11)	0.52	6.30	0.00
Two ATM Calls	0.007 (0.42)	0.84 (1.39)	0.50	7.40	0.00
Two ATM Puts	0.006 (0.36)	0.86 (1.16)	0.49	5.17	0.01
Two ATM Calls and Two ATM Puts	0.001 (0.06)	0.89 (0.93)	0.52	6.14	0.00
Panel B: June 1988 – August 1997					
ATM Call	0.006 (0.51)	0.83 (2.10)	0.51	21.47	0.00
ATM Put	0.015 (1.37)	0.76 (3.19)	0.49	26.50	0.00
One ATM Call and One ATM Put	0.009 (0.80)	0.81 (2.50)	0.50	23.95	0.00
Two ATM Calls	0.008 (0.66)	0.82 (2.23)	0.49	21.23	0.00
Two ATM Puts	0.017 (1.52)	0.75 (3.27)	0.46	25.57	0.00
Two ATM Calls and Two ATM Puts	0.010 (0.82)	0.80 (2.47)	0.49	23.06	0.00

Table V

Regression of Realized Volatility on Predicted Volatility for Various Specifications of $f(\bullet)$

$$f\left(Vol_{Realized}(t)\right) = \alpha + \beta f\left(Vol_{Implied}(t)\right) + \varepsilon(t)$$

OLS point estimates with *t*-statistics for $\alpha = 0$ and $1-\beta = 0$ (in parentheses). The *t*-statistics are computed from OLS standard errors. The $Vol_{Realized}(t)$ variable is the realized volatility of the SPX returns derived from five minute interval SPX index levels inferred from SPX futures prices. The realized volatility is computed under the restriction that the mean return is equal to zero. The $Vol_{Implied}(t)$ variable is the Black-Scholes implied volatility from the two closest to ATM calls and the two closest to ATM puts. The Black-Scholes implied volatility is computed by using the quantity $Se^{-\delta T}$, inferred from futures prices, as the value of the SPX index level and then assuming that the SPX index pays no dividends. The Black-Scholes implied volatility is defined as the real number that minimizes the sum of the squared pricing errors of the four options (i.e., it is *not* the average of the individual Black-Scholes implied volatilities). The *F*-statistic tests the joint hypothesis that $\alpha = 0$ and $\beta = 1$. The *p*-value gives the probability that the *F*-statistic is as large as it is if the joint hypothesis is true. There is one observation for each Wednesday in the data period that follows an option expiration date, and the time interval covered by each observation extends from that Wednesday to the next option expiration date. $I(\bullet)$ is the identity function, i.e., $I(x) = x$.

$f(\bullet)$	α	β	Adj. R^2	*F*-statistic	*p*-value
Panel A: June 1993 – August 1997					
$(\bullet)^2$	0.001 (0.51)	0.78 (1.92)	0.47	4.88	0.01
$I(\bullet)$	0.001 (0.06)	0.89 (0.93)	0.52	6.14	0.00

Table V – Continued

$f(\bullet)$	α	β	Adj. R^2	*F*-statistic	*p*-value
$\ln(\bullet)$	-0.068	1.03	0.55	9.36	0.00
	(-0.25)	(-0.25)			
Panel B: June 1988 – August 1997					
$(\bullet)^2$	0.002	0.70	0.40	17.59	0.00
	(1.03)	(3.69)			
$I(\bullet)$	0.010	0.80	0.49	23.06	0.00
	(0.82)	(2.47)			
$\ln(\bullet)$	-0.257	0.95	0.56	30.91	0.00
	(-1.58)	(0.61)			

Table VI
Forecasting Regression with Historical Predictor Variable

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Historical}(t) + \varepsilon(t)$$

OLS point estimates with *t*-statistics for $\alpha = 0$ and $1 - \beta = 0$ (in parentheses). The *t*-statistics are computed from OLS standard errors. The $Vol_{Realized}(t)$ variable is the realized volatility of the SPX returns derived from five minute interval SPX index levels inferred from SPX futures prices. The realized volatility is computed under the restriction that the mean return is equal to zero. $Vol_{Historical}(t)$ is the historical volatility computed from five minute observations on the SPX level derived from SPX futures prices also under the restriction that the mean return is equal to zero. The historical volatility is computed either from 1 month, 2 months, 3 months, or 6 months of SPX levels leading up to the observation time. The *F*-statistic tests the joint hypothesis $\alpha = 0$ and $\beta = 1$. The *p*-value gives the probability that the *F*-statistic is as large as it is if the joint hypothesis is true. There is one observation for each Wednesday in the data period that follows an option expiration date, and the time interval covered by each observation extends from that Wednesday to the next option expiration date.

Past Data for Historical Prediction	α	β	Adj. R^2	*F*-statistic	*p*-value
Panel A: June 1993 – August 1997					
1 Month	0.030 (2.20)	0.76 (2.15)	0.47	2.44	0.10
2 Months	0.025 (1.64)	0.80 (1.51)	0.43	1.37	0.26
3 Months	0.021 (1.27)	0.84 (1.12)	0.40	0.88	0.42
6 Months	0.003 (0.16)	1.01 (-0.08)	0.42	0.50	0.61

Table VI– Continued

Past Data for Historical Prediction	α	β	Adj. R^2	F-statistic	p-value
Panel B: June 1988 – August 1997					
1 Month	0.043	0.65	0.38	10.18	0.00
	(4.19)	(4.50)			
2 Months	0.035	0.71	0.40	5.86	0.00
	(3.16)	(3.41)			
3 Months	0.029	0.76	0.41	4.12	0.02
	(2.55)	(2.81)			
6 Months	0.028	0.75	0.36	3.72	0.03
	(2.26)	(2.58)			

Table VII
Forecasting Regression with Option Implied and Historical Predictor Variables

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \gamma Vol_{Historical}(t) + \varepsilon(t)$$

OLS point estimates with *t*-statistics for $\alpha = 0,$ $1 - \beta = 0,$ and $\gamma = 0$ (in parentheses). The *t*-statistics are computed from OLS standard errors. The $Vol_{Realized}(t)$ variable is the realized volatility of the SPX returns derived from five minute interval SPX index levels inferred from SPX futures prices. The realized volatility is computed under the restriction that the mean return is equal to zero. $Vol_{Implied}(t)$ is the Black-Scholes implied volatility from the two closest to ATM calls and the two closest to ATM puts. The Black-Scholes implied volatility is computed by using the quantity $Se^{-\delta T},$ inferred from futures prices, as the value of the SPX index level and then assuming that the SPX index pays no dividends. The Black-Scholes implied volatility is defined as the real number that minimizes the sum of the squared pricing errors of the four options (i.e., it is *not* the average of the individual Black-Scholes implied volatilities). $Vol_{Historical}(t)$ is the historical volatility computed from five minute observations on the SPX level derived from futures prices also under the restriction that the mean return is equal to zero. The historical volatility is computed either from 1 month, 2 months, 3 months, or 6 months of SPX levels leading up to the observation time. The *F*-statistic tests the joint hypothesis $\alpha = 0,$ $\beta = 1,$ and $\gamma = 0.$ The *p*-value gives the probability that the *F*-statistic is as large as it is if the joint hypothesis is true. There is one observation for each Wednesday in the data period that follows an option expiration date, and the time interval covered by each observation extends from that Wednesday to the next option expiration date.

Table VII– Continued

Past Data for Historical Prediction	α	β	γ	Adj. R^2	F-stat	p-value
Panel A: June 1993 – August 1997						
1 Month	0.004	0.68	0.21	0.52	4.35	0.01
	(0.25)	(1.22)	(0.90)			
2 Months	0.000	0.72	0.20	0.52	4.35	0.01
	(0.02)	(1.26)	(0.90)			
3 Months	0.000	0.82	0.08	0.51	4.06	0.01
	(0.01)	(0.75)	(0.32)			
6 Months	-0.003	0.77	0.18	0.51	4.15	0.01
	(-0.17)	(0.97)	(0.58)			
Panel B: June 1988 – August 1997						
1 Month	0.010	0.81	0.00	0.48	15.23	0.00
	(0.79)	(1.10)	(-0.02)			
2 Months	0.009	0.66	0.16	0.49	15.80	0.00
	(0.80)	(2.24)	(1.10)			
3 Months	0.008	0.63	0.21	0.49	16.17	0.00
	(0.67)	(2.51)	(1.40)			
6 Months	0.007	0.70	0.14	0.49	15.67	0.00
	(0.55)	(2.23)	(0.97)			

Table VIII
Forecasting Regression with Heston Model Option Implied and Historical Predictor Variables, June 1993 through August 1997

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \gamma Vol_{Historical}(t) + \varepsilon(t)$$

OLS point estimates with *t*-statistics for $\alpha = 0,\ 1-\beta = 0,$ and $\gamma = 0$ (in parentheses). The *t*-statistics are computed from OLS standard errors. The $Vol_{Realized}(t)$ variable is the realized volatility of the SPX returns derived from five minute interval SPX index levels inferred from SPX futures prices. The realized volatility is computed under the restriction that the mean return is equal to zero. The $Vol_{Implied}(t)$ variable is the Heston model implied volatility for the time period *t* to $t+T,$ where *T* is the time to expiration of the shortest maturity options at time *t*. In order to compute $Vol_{Implied}(t)$, estimates are needed at time *t* for the Heston model parameters $k, \theta, \eta, \rho,$ and $\lambda.$ The $k, \theta,$ and η parameters are estimated by maximizing the likelihood function of the Heston model variance process using a time-series of monthly variances computed from CRSP SPX index closing level from July 1962 through the month before the observation time *t*. In the following, whenever the Heston model is used to compute option prices, the value of the SPX index is set to the quantity $Se^{-\delta T}$ inferred from futures prices, and it is then assumed that the SPX index pays no dividends. Conditional on the estimates of $k, \theta,$ and $\eta,$ ρ and λ are obtained by minimizing the sum of the Heston model squared pricing errors of all time *t* out-of-the-money SPX option 3:00 PM bid-ask midpoints that meet mild exclusionary criteria. The time *t* instantaneous volatility is allowed to vary freely while performing this minimization. However, once the ρ and λ parameters are determined, the Heston model implied instantaneous volatility is re-computed by minimizing the sum of the Heston model squared pricing error of the two closest to ATM calls and the two closest to ATM puts that expire in three and a half to four and a half weeks. The Heston model implied instantaneous variance, $Var_{Instantaneous}(t)$ is next set equal to the square of this Heston model implied instantaneous volatility and $Vol_{Implied}(t)$ is computed from

$$Vol_{Implied}(t) = \sqrt{\theta + \left[\frac{1-e^{-kT}}{kT}\right]\left[Var_{Instantaneous}(t) - \theta\right]}.$$

The $Vol_{Historical}(t)$ variable is the historical volatility computed from five minute observations on the SPX level derived from futures prices also under the restriction that the mean return is equal to zero. The historical volatility is computed either from 1 month, 2 months, 3 months, or 6 months of SPX levels leading up to the observation time. The *F*-statistic tests the joint hypothesis $\alpha = 0$, $\beta = 1$, and $\gamma = 0$. The *p*-value gives the probability that the *F*-statistic is as large as it is if the joint hypothesis is true. There is one observation for each Wednesday from June 1993 through August 1997 that follows an option expiration date, and the time interval covered by each observation extends from that Wednesday to the next option expiration date.

Past Data for Historical Prediction	α	β	γ	Adj R^2	*F*-stat	*p*-value
1 Month	0.000 (0.00)	0.69 (1.04)	0.25 (1.03)	0.51	2.86	0.05
2 Month	-0.005 (-0.31)	0.75 (1.03)	0.23 (1.04)	0.51	2.87	0.05
3 Month	-0.007 (-0.39)	0.85 (0.60)	0.13 (0.54)	0.50	2.57	0.07
6 Month	-0.011 (-0.59)	0.78 (0.86)	0.25 (0.85)	0.51	2.73	0.05

Table IX
Monte Carlo Simulation of Univariate Regression with $Vol_{Implied}$ Computed from the Black-Scholes Formula
$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \varepsilon(t)$$

The simulations assume that the true (unobserved) dynamics of the underlying SPX index level, *S*, and SPX instantaneous variance, *V*, are described by

$$\frac{dS_t}{S_t} = \mu(S_t, V_t, t)\,dt + \sqrt{V_t}\,dW_t^S$$

$$dV_t = k(\theta - V_t)\,dt + \eta\sqrt{V_t}\,dW_t^V$$

$$Corr\left(dW_t^S, dW_t^V\right) = \rho$$

$$\lambda(S_t, V_t, t) = \lambda V_t$$

$$r = \delta = constant.$$

This model is calibrated with $\mu(S_t, V_t, t) = 0.124$, $k = 4.07$, $\theta = 0.0181$, $\eta = 0.506$, $\rho = -0.761$, $r = 0.059$, and $\delta = 0.025$. Each experiment uses an Euler scheme to simulate 2000 sample paths of the index and instantaneous variance levels from the model. On January 2, 1986 the SPX index level was at 209.59 and there were 244,634 five-minute trading intervals in the SPX futures market from the opening of trading on January 2, 1986 through the close of trading on September 18, 1997. For this reason, each of the simulated sample paths begins with the index level at 209.59 and each path is 244,634 steps long where each step corresponds to five

Table IX – Continued

minutes of trading time. In addition, each simulation starts with the instantaneous variance level at its long-run mean value of 0.0181. The $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables are computed at fifty-one points along each simulated paths at times that correspond to 3:00 PM on the Wednesday following an option expiration over the period June 1993 through August 1997. These 51 points will be referred to as the observation dates. The panels below report results for various levels of error introduced into the SPX index levels and option prices and for various levels of the market price of variance risk. For each of the 244,634 steps on each simulated path, spot-futures parity was used to convert the SPX level into a futures price for each of the three delivery times that were actually used at the corresponding time during the tests on the real data. A band with width of either \$0.00, \$0.10, or \$0.20 was then centered on each of the spot-futures parity derived futures prices, and a price for each of the three futures was drawn uniformly from its band. These three futures prices were then used to derive an SPX level which reflects the error (if any) in the futures prices. Error is introduced into the option prices as follows. On each observation date, a target ATM strike price level is set equal to the futures price determined via spot-futures parity from the current simulated level of the SPX index (without error), the time to expiration of the options on that date in the actual data, r, and δ. The strike prices of the two closest to ATM calls and two closest to ATM puts are then set to values that give them a displacement from this target ATM strike price level that is the same as the displacement of the actual option strike prices from the actual target ATM strike price level on the corresponding date in the real data. Next, the true price of these options is set to their correct prices under the Heston model given the assumed parameters values. A band with width of either zero, the bid-ask

Table IX– Continued

spread of the corresponding call or put in the real data observed on the same trade date, or twice that bid-ask spread is then centered on each of the options prices, and a price for each of the four options is drawn uniformly from its band. Once the SPX levels and option prices are determined, the 51 pairs of $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables are constructed exactly as in the empirical work where the $Vol_{Implied}(t)$ is computed from the four option prices using the Black-Scholes model. These variables are then used to estimate the equation

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \varepsilon(t)$$

using OLS. The empirical percentile is the percentage of times the observed statistic exceeds the simulated statistic under the panel's assumptions. The empirical 1-sided *p*-value is the percentage of times that the simulated statistic is both on the same side of the statistic's null value as the observed statistic and further away from the statistic's null value than the observed statistic. The empirical 2-sided *p*-value is the percentage of times that the simulated statistic is further away (in absolute value) from the statistic's null value than the observed statistic. The null values of the statistics are $\alpha = 0$, $\beta = 1$, $R^2 = 1$, and $F = 0$. The α std. and β std. statistics are OLS standard errors for α and β respectively.

Table IX– Continued

Statistic	Fractiles of Statistics 0.01	0.05	0.10	0.50	0.90	0.95	0.99	Mean	Std. Dev.	Obs. Stat.	Empirical Percentile	Emp. 1-sided *p*-value	Emp. 2-sided *p*-value
Panel A: Futures Error Band = \$ 0.00, Options Error Band = 0, $\lambda = 0$													
α	-0.010	-0.006	-0.003	0.007	0.018	0.022	0.029	0.007	0.008	0.001	0.24	0.76	0.93
β	0.72	0.81	0.84	0.95	1.04	1.07	1.10	0.95	0.08	0.89	0.21	0.21	0.22
R^2	0.45	0.57	0.63	0.77	0.87	0.89	0.92	0.75	0.10	0.52	0.02	0.02	0.02
F-stat	0.01	0.07	0.14	0.89	2.83	3.63	6.02	1.26	1.25	6.14	0.99	0.01	0.01
p-Val	0.00	0.03	0.07	0.42	0.87	0.93	0.99	0.44	0.29				
α std.	0.006	0.007	0.008	0.009	0.011	0.012	0.013	0.009	0.001				
β std.	0.04	0.05	0.06	0.07	0.10	0.10	0.12	0.07	0.02				
Panel B: Futures Error Band = \$0.00, Options Error Band = 1*x* Actual Bid-Ask Spread, $\lambda = 0$													
α	-0.010	-0.006	-0.003	0.007	0.018	0.022	0.029	0.007	0.008	0.001	0.23	0.77	0.94
β	0.72	0.81	0.84	0.95	1.04	1.07	1.10	0.95	0.08	0.89	0.22	0.22	0.23
R^2	0.45	0.57	0.63	0.77	0.86	0.89	0.92	0.75	0.10	0.52	0.02	0.02	0.02
F-stat	0.01	0.07	0.14	0.91	2.84	3.65	6.05	1.27	1.26	6.14	0.99	0.01	0.01
p-Val	0.00	0.03	0.07	0.41	0.87	0.93	0.99	0.44	0.29				
α std.	0.006	0.007	0.007	0.009	0.011	0.012	0.013	0.009	0.001				
β std.	0.04	0.05	0.06	0.07	0.10	0.10	0.12	0.07	0.02				
Panel C: Futures Error Band = \$0.10, Options Error Band = 0, $\lambda = 0$													
α	-0.018	-0.011	-0.007	0.006	0.019	0.023	0.034	0.006	0.011	0.001	0.30	0.70	0.94
β	0.67	0.76	0.80	0.93	1.07	1.11	1.17	0.93	0.11	0.89	0.33	0.33	0.38
R^2	0.38	0.48	0.53	0.69	0.81	0.83	0.88	0.68	0.11	0.52	0.09	0.09	0.09
F-stat	0.02	0.07	0.15	0.97	3.52	5.07	8.31	1.51	1.71	6.14	0.98	0.02	0.02
p-Val	0.00	0.01	0.04	0.39	0.86	0.93	0.98	0.42	0.30				
α std.	0.007	0.008	0.009	0.011	0.014	0.015	0.018	0.011	0.002				
β std.	0.06	0.06	0.07	0.09	0.11	0.12	0.13	0.09	0.02				
Panel D: Futures Error Band = \$0.10, Options Error Band = 1*x* Actual Bid-Ask Spread, $\lambda = 0$													
α	-0.017	-0.011	-0.007	0.006	0.019	0.023	0.034	0.006	0.011	0.001	0.30	0.70	0.94
β	0.66	0.76	0.80	0.93	1.07	1.11	1.17	0.93	0.11	0.89	0.33	0.33	0.37
R^2	0.38	0.48	0.53	0.69	0.81	0.83	0.88	0.68	0.11	0.52	0.09	0.09	0.09
F-stat	0.02	0.06	0.14	0.96	3.53	5.12	8.36	1.51	1.72	6.14	0.98	0.02	0.02
p-Val	0.00	0.01	0.04	0.39	0.87	0.94	0.98	0.42	0.30				
α std.	0.007	0.008	0.009	0.011	0.014	0.015	0.018	0.011	0.002				
β std.	0.06	0.06	0.07	0.09	0.11	0.12	0.13	0.09	0.02				

Table IX– Continued

Statistic	Fractiles of Statistics 0.01	0.05	0.10	0.50	0.90	0.95	0.99	Mean	Std. Dev.	Obs. Stat.	Empirical Percentile	Emp. 1-sided *p*-value	Emp. 2-sided *p*-value
Panel E: Futures Error Band = \$0.20, Options Error Band = 2*x* Actual Bid-Ask Spread, $\lambda = 0$													
α	-0.007	-0.002	0.000	0.011	0.022	0.025	0.033	0.011	0.009	0.001	0.12	0.88	0.96
β	0.69	0.78	0.82	0.94	1.03	1.05	1.08	0.93	0.08	0.89	0.28	0.28	0.28
R^2	0.44	0.57	0.62	0.77	0.86	0.89	0.92	0.75	0.10	0.52	0.02	0.02	0.02
F-stat	0.03	0.13	0.23	1.39	3.92	4.87	7.21	1.78	1.55	6.14	0.98	0.02	0.02
p-Val	0.00	0.01	0.03	0.26	0.79	0.87	0.97	0.33	0.28				
α std.	0.006	0.007	0.007	0.009	0.011	0.012	0.013	0.009	0.001				
β std.	0.04	0.05	0.06	0.07	0.09	0.10	0.11	0.07	0.02				
Panel F: Futures Error Band = \$0.00, Options Error Band = 0, $\lambda = -3.521$													
α	-0.009	-0.005	-0.003	0.007	0.019	0.022	0.030	0.008	0.008	0.001	0.21	0.79	0.93
β	0.67	0.76	0.79	0.90	0.98	1.00	1.04	0.89	0.08	0.89	0.45	0.45	0.45
R^2	0.46	0.57	0.63	0.77	0.87	0.89	0.92	0.76	0.10	0.52	0.02	0.02	0.02
F-stat	0.06	0.20	0.38	1.97	5.46	6.92	10.33	2.55	2.26	6.14	0.93	0.07	0.07
p-Val	0.00	0.00	0.01	0.15	0.69	0.82	0.94	0.25	0.26				
α std.	0.006	0.007	0.007	0.009	0.011	0.012	0.013	0.009	0.001				
β std.	0.04	0.05	0.05	0.07	0.09	0.10	0.11	0.07	0.01				
Panel G: Futures Error Band = \$0.00, Options Error Band = 1*x* Actual Bid-Ask Spread, $\lambda = -3.521$													
α	-0.009	-0.005	-0.003	0.007	0.019	0.022	0.029	0.008	0.008	0.001	0.21	0.79	0.94
β	0.67	0.76	0.79	0.90	0.98	1.00	1.04	0.89	0.08	0.89	0.45	0.45	0.45
R^2	0.46	0.58	0.63	0.77	0.86	0.89	0.92	0.76	0.10	0.52	0.02	0.02	0.02
F-stat	0.06	0.20	0.39	1.96	5.44	6.89	10.32	2.55	2.26	6.14	0.93	0.07	0.07
p-Val	0.00	0.00	0.01	0.15	0.68	0.82	0.94	0.25	0.26				
α std.	0.006	0.007	0.007	0.009	0.011	0.012	0.013	0.009	0.001				
β std.	0.04	0.05	0.05	0.07	0.09	0.10	0.11	0.07	0.01				
Panel H: Futures Error Band = \$0.10, Options Error Band = 0, $\lambda = -3.521$													
α	-0.017	-0.010	-0.006	0.007	0.020	0.024	0.034	0.007	0.010	0.001	0.28	0.72	0.93
β	0.64	0.71	0.75	0.88	1.01	1.05	1.10	0.88	0.10	0.89	0.53	0.53	0.54
R^2	0.38	0.48	0.53	0.69	0.81	0.83	0.88	0.68	0.11	0.52	0.09	0.09	0.09
F-stat	0.04	0.19	0.39	2.44	7.55	9.76	15.16	3.32	3.23	6.14	0.86	0.14	0.14
p-Val	0.00	0.00	0.00	0.10	0.68	0.83	0.96	0.23	0.27				
α std.	0.007	0.008	0.009	0.011	0.014	0.015	0.018	0.011	0.002				
β std.	0.05	0.06	0.06	0.08	0.11	0.11	0.12	0.08	0.02				

Table IX– Continued

Statistic	Fractiles of Statistics 0.01	0.05	0.10	0.50	0.90	0.95	0.99	Mean	Std. Dev.	Obs. Stat.	Empirical Percentile	Emp. 1-sided *p*-value	Emp. 2-sided *p*-value
Panel I: Futures Error Band = \$0.10, Options Error Band = 1*x* Actual Bid-Ask Spread, $\lambda = -3.521$													
α	-0.017	-0.010	-0.006	0.007	0.020	0.024	0.034	0.007	0.010	0.001	0.28	0.72	-0.017
β	0.64	0.71	0.75	0.88	1.01	1.05	1.10	0.88	0.10	0.89	0.53	0.53	0.64
R^2	0.38	0.48	0.53	0.69	0.81	0.83	0.88	0.68	0.11	0.52	0.08	0.08	0.38
F-stat	0.04	0.19	0.39	2.41	7.45	9.73	15.16	3.30	3.20	6.14	0.86	0.14	0.04
p-Val	0.00	0.00	0.00	0.10	0.68	0.83	0.96	0.23	0.27				0.00
α std.	0.007	0.008	0.009	0.011	0.014	0.015	0.018	0.011	0.002				0.007
β std.	0.05	0.06	0.06	0.08	0.11	0.11	0.12	0.08	0.02				0.05
Panel J: Futures Error Band = \$0.20, Options Error Band = 2*x* Actual Bid-Ask Spread, $\lambda = -3.521$													
α	-0.006	-0.001	0.001	0.012	0.023	0.026	0.034	0.012	0.008	0.001	0.10	0.90	0.96
β	0.65	0.74	0.78	0.88	0.97	0.99	1.02	0.88	0.08	0.89	0.53	0.53	0.53
R^2	0.44	0.57	0.63	0.77	0.86	0.89	0.92	0.75	0.10	0.52	0.02	0.02	0.02
F-stat	0.04	0.20	0.39	2.04	5.29	6.74	10.39	2.57	2.19	6.14	0.93	0.07	0.07
p-Val	0.00	0.00	0.01	0.14	0.68	0.82	0.96	0.25	0.26				
α std.	0.006	0.007	0.007	0.009	0.011	0.011	0.013	0.009	0.001				
β std.	0.04	0.05	0.05	0.07	0.09	0.09	0.11	0.07	0.01				

Table X: Monte Carlo Simulation of Univariate Regression with $Vol_{Implied}$ Computed from the Heston Model

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \varepsilon(t)$$

The simulations assume that the true (unobserved) dynamics of the underlying SPX index level, *S*, and SPX instantaneous variance, *V*, are described by

$$\frac{dS_t}{S_t} = \mu(S_t, V_t, t)\,dt + \sqrt{V_t}\,dW_t^S$$

$$dV_t = k(\theta - V_t)\,dt + \eta\sqrt{V_t}\,dW_t^V$$

$$Corr\left(dW_t^S, dW_t^V\right) = \rho$$

$$\lambda(S_t, V_t, t) = \lambda V_t$$

$$r = \delta = constant.$$

This model is calibrated with $\mu(S_t, V_t, t) = 0.124$, $k = 4.07$, $\theta = 0.0181$, $\eta = 0.506$, $\rho = -0.761$, $r = 0.059$, and $\delta = 0.025$. Each experiment uses an Euler scheme to simulate 2000 sample paths of the index and instantaneous variance levels from the model. On January 2, 1986 the SPX index level was at 209.59 and there were 244,634 five-minute trading intervals in the SPX futures market from the opening of trading on January 2, 1986 through the close of trading on September 18, 1997. For this reason, each of the simulated sample paths begins with the index level at 209.59 and each path is 244,634 steps long where each step corresponds to five minutes of trading time. In addition, each simulation starts with the instantaneous variance level at its long-run mean value of 0.0181.

Table X– Continued

The $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables are computed at fifty-one points along each simulated paths at times that correspond to 3:00 PM on the Wednesday following an option expiration over the period June 1993 through August 1997. These 51 points will be referred to as the observation dates. The panels below report results for various levels of error introduced into the SPX index levels and option prices and for various levels of the market price of variance risk. For each of the 244,634 steps on each simulated path, spot-futures parity was used to convert the SPX level into a futures price for each of the three delivery times that were actually used at the corresponding time during the tests on the real data. A band with width of either $0.00, $0.10, or $0.20 was then centered on each of the spot-futures parity derived futures prices, and a price for each of the three futures was drawn uniformly from its band. These three futures prices were then used to derive an SPX level which reflects the error (if any) in the futures prices. Error is introduced into the option prices as follows. On each observation date, a target ATM strike price level is set equal to the futures price determined via spot-futures parity from the current simulated level of the SPX index (without error), the time to expiration of the options on that date in the actual data, r, and δ. The strike prices of the two closest to ATM calls and two closest to ATM puts are then set to values that give them a displacement from this target ATM strike price level that is the same as the displacement of the actual option strike prices from the actual target ATM strike price level on the corresponding date in the real data. Next, the true price of these options is set to their correct prices under the Heston model given the assumed parameters values. A band with width of either zero, the bid-ask spread of the corresponding call or put in the real data observed on the same trade date, or twice that bid-ask spread is then centered on

Table X – Continued

each of the options prices, and a price for each of the four options is drawn uniformly from its band. Once the SPX levels and option prices are determined, the 51 pairs of $Vol_{Realized}(t)$ and $Vol_{Implied}(t)$ variables are constructed exactly as in the empirical work where the $Vol_{Implied}(t)$ is computed from the four option prices using the Heston model with parameters set to the calibrated values. These variables are then used to estimate the equation

$$Vol_{Realized}(t) = \alpha + \beta Vol_{Implied}(t) + \varepsilon(t)$$

using OLS. The empirical percentile is the percentage of times the observed statistic exceeds the simulated statistic under the panel's assumptions. The empirical 1-sided *p*-value is the percentage of times that the simulated statistic is both on the same side of the statistic's null value as the observed statistic and further away from the statistic's null value than the observed statistic. The empirical 2-sided *p*-value is the percentage of times that the simulated statistic is further away (in absolute value) from the statistic's null value than the observed statistic. The null values of the statistics are $\alpha = 0$, $\beta = 1$, $R^2 = 1$, and $F = 0$. The α std. and β std. statistics are OLS standard errors for α and β respectively.

Table X – Continued

Statistic	Fractiles of Statistics 0.01	0.05	0.10	0.50	0.90	0.95	0.99	Mean	Std. Dev.	Obs. Stat.	Empirical Percentile	Emp. 1-sided *p*-value	Emp. 2-sided *p*-value
Panel A: Futures Error Band = \$0.00, Options Error Band = 0, $\lambda = 0$													
α	-0.019	-0.015	-0.012	-0.002	0.011	0.015	0.023	-0.001	0.009	-0.006	0.28	0.28	0.48
β	0.74	0.83	0.87	0.98	1.07	1.09	1.13	0.97	0.08	0.96	0.42	0.42	0.63
R^2	0.45	0.57	0.63	0.77	0.87	0.89	0.92	0.76	0.10	0.51	0.02	0.02	0.02
F-stat	0.03	0.09	0.16	1.09	4.01	5.37	8.74	1.70	1.96	3.76	0.89	0.11	0.11
p-Val	0.00	0.01	0.02	0.34	0.86	0.92	0.98	0.40	0.30				
α std.	0.007	0.008	0.008	0.010	0.012	0.012	0.014	0.010	0.001				
β std.	0.04	0.05	0.06	0.07	0.10	0.11	0.12	0.08	0.02				
Panel B: Futures Error Band = \$0.00, Options Error Band = 1*x* Actual Bid-Ask Spread, $\lambda = 0$													
α	-0.019	-0.015	-0.012	-0.001	0.011	0.015	0.023	-0.001	0.009	-0.006	0.28	0.28	0.48
β	0.74	0.83	0.87	0.98	1.07	1.09	1.13	0.97	0.08	0.96	0.43	0.43	0.64
R^2	0.45	0.57	0.63	0.77	0.87	0.89	0.92	0.75	0.10	0.51	0.02	0.02	0.02
F-stat	0.02	0.09	0.16	1.09	4.03	5.41	8.62	1.71	1.96	3.76	0.88	0.12	0.12
p-Val	0.00	0.01	0.02	0.34	0.85	0.92	0.98	0.40	0.30				
α std.	0.007	0.008	0.008	0.010	0.012	0.012	0.014	0.010	0.001				
β std.	0.04	0.05	0.06	0.07	0.10	0.11	0.12	0.08	0.02				
Panel C: Futures Error Band = \$0.10, Options Error Band = 0, $\lambda = 0$													
α	-0.027	-0.020	-0.016	-0.002	0.012	0.016	0.028	-0.002	0.011	-0.006	0.34	0.34	0.57
β	0.69	0.78	0.82	0.96	1.10	1.14	1.19	0.96	0.11	0.96	0.53	0.53	0.74
R^2	0.39	0.48	0.53	0.69	0.81	0.83	0.88	0.68	0.11	0.51	0.07	0.07	0.07
F-stat	0.01	0.08	0.21	1.65	5.66	7.56	12.77	2.43	2.66	3.76	0.80	0.20	0.20
p-Val	0.00	0.00	0.01	0.20	0.81	0.92	0.99	0.31	0.30				
α std.	0.008	0.009	0.009	0.012	0.015	0.016	0.019	0.012	0.002				
β std.	0.06	0.06	0.07	0.09	0.12	0.12	0.13	0.09	0.02				
Panel D: Futures Error Band = \$0.10, Options Error Band = 1*x* Actual Bid-Ask Spread, $\lambda = 0$													
α	-0.027	-0.020	-0.016	-0.002	0.012	0.017	0.029	-0.002	0.011	-0.006	0.34	0.34	0.57
β	0.68	0.78	0.82	0.96	1.10	1.14	1.20	0.96	0.11	0.96	0.52	0.52	0.75
R^2	0.38	0.48	0.53	0.69	0.81	0.83	0.88	0.68	0.11	0.51	0.07	0.07	0.07
F-stat	0.02	0.09	0.21	1.65	5.60	7.56	12.62	2.43	2.66	3.76	0.80	0.20	0.20
p-Val	0.00	0.00	0.01	0.20	0.81	0.92	0.98	0.31	0.30				
α std.	0.008	0.009	0.009	0.012	0.015	0.016	0.019	0.012	0.002				
β std.	0.06	0.06	0.07	0.09	0.12	0.12	0.13	0.09	0.02				

Table X – Continued

Statistic	Fractiles of Statistics 0.01	0.05	0.10	0.50	0.90	0.95	0.99	Mean	Std. Dev.	Obs. Stat.	Empirical Percentile	Emp. 1-sided *p*-value	Emp. 2-sided *p*-value
Panel E: Futures Error Band = \$0.20, Options Error Band = 2*x* Actual Bid-Ask Spread, $\lambda = 0$													
α	-0.016	-0.011	-0.009	0.003	0.015	0.019	0.027	0.003	0.009	-0.006	0.16	0.16	0.50
β	0.72	0.81	0.85	0.96	1.05	1.07	1.12	0.96	0.08	0.96	0.51	0.51	0.67
R^2	0.44	0.57	0.62	0.77	0.86	0.89	0.92	0.75	0.10	0.51	0.02	0.02	0.02
F-stat	0.01	0.08	0.14	0.93	3.43	4.69	7.99	1.47	1.69	3.76	0.92	0.08	0.08
p-Val	0.00	0.01	0.04	0.40	0.87	0.93	0.99	0.43	0.30				
α std.	0.007	0.007	0.008	0.010	0.012	0.012	0.014	0.010	0.001				
β std.	0.04	0.05	0.06	0.07	0.10	0.10	0.12	0.08	0.02				
Panel F: Futures Error Band = \$0.00, Options Error Band = 0, $\lambda = -3.521$													
α	-0.019	-0.015	-0.012	-0.002	0.011	0.015	0.022	-0.001	0.009	-0.006	0.27	0.27	0.48
β	0.75	0.83	0.87	0.98	1.07	1.09	1.13	0.97	0.08	0.96	0.42	0.42	0.63
R^2	0.46	0.58	0.63	0.77	0.87	0.89	0.92	0.76	0.10	0.51	0.02	0.02	0.02
F-stat	0.02	0.09	0.15	1.06	4.00	5.34	8.49	1.68	1.94	3.76	0.89	0.11	0.11
p-Val	0.00	0.01	0.02	0.35	0.86	0.92	0.98	0.40	0.30				
α std.	0.007	0.008	0.008	0.010	0.012	0.012	0.014	0.010	0.001				
β std.	0.04	0.05	0.06	0.07	0.10	0.10	0.12	0.08	0.02				
Panel G: Futures Error Band = \$0.00, Options Error Band = 1*x* Actual Bid-Ask Spread, $\lambda = -3.521$													
α	-0.019	-0.015	-0.012	-0.001	0.011	0.015	0.022	-0.001	0.009	-0.006	0.28	0.28	0.48
β	0.75	0.83	0.87	0.98	1.07	1.09	1.13	0.97	0.08	0.96	0.42	0.42	0.63
R^2	0.46	0.58	0.63	0.77	0.87	0.89	0.92	0.76	0.10	0.51	0.02	0.02	0.02
F-stat	0.02	0.09	0.15	1.06	3.99	5.32	8.50	1.68	1.94	3.76	0.89	0.11	0.11
p-Val	0.00	0.01	0.02	0.35	0.86	0.92	0.98	0.40	0.30				
α std.	0.007	0.008	0.008	0.010	0.012	0.012	0.014	0.010	0.001				
β std.	0.04	0.05	0.06	0.07	0.10	0.10	0.12	0.08	0.02				
Panel H: Futures Error Band = \$0.10, Options Error Band = 0, $\lambda = -3.521$													
α	-0.027	-0.020	-0.016	-0.002	0.012	0.016	0.028	-0.002	0.011	-0.006	0.34	0.34	0.57
β	0.70	0.78	0.82	0.96	1.10	1.14	1.19	0.96	0.11	0.96	0.52	0.52	0.74
R^2	0.39	0.48	0.53	0.69	0.81	0.83	0.88	0.68	0.11	0.51	0.07	0.07	0.07
F-stat	0.01	0.08	0.20	1.63	5.66	7.51	12.77	2.42	2.66	3.76	0.80	0.20	0.20
p-Val	0.00	0.00	0.01	0.21	0.82	0.92	0.99	0.32	0.30				
α std.	0.008	0.009	0.009	0.012	0.015	0.016	0.019	0.012	0.002				
β std.	0.06	0.06	0.07	0.09	0.12	0.12	0.13	0.09	0.02				

Table X – Continued

Statistic	Fractiles of Statistics 0.01	0.05	0.10	0.50	0.90	0.95	0.99	Mean	Std. Dev.	Obs. Stat.	Empirical Percentile	Emp. 1-sided *p*-value	Emp. 2-sided *p*-value
Panel I: Futures Error Band = \$0.10, Options Error Band = 1*x* Actual Bid-Ask Spread, $\lambda = -3.521$													
α	-0.027	-0.020	-0.016	-0.002	0.012	0.017	0.028	-0.002	0.011	-0.006	0.34	0.34	0.57
β	0.70	0.78	0.82	0.96	1.10	1.14	1.20	0.96	0.11	0.96	0.52	0.52	0.74
R^2	0.38	0.48	0.53	0.69	0.81	0.83	0.88	0.68	0.11	0.51	0.07	0.07	0.07
F-stat	0.01	0.08	0.20	1.63	5.58	7.44	12.60	2.41	2.64	3.76	0.80	0.20	0.20
p-Val	0.00	0.00	0.01	0.21	0.82	0.92	0.99	0.32	0.31				
α std.	0.008	0.009	0.009	0.012	0.015	0.016	0.019	0.012	0.002				
β std.	0.06	0.06	0.07	0.09	0.12	0.12	0.13	0.09	0.02				
Panel J: Futures Error Band = \$0.20, Options Error Band = 2*x* Actual Bid-Ask Spread, $\lambda = -3.521$													
α	-0.016	-0.011	-0.008	0.003	0.015	0.019	0.027	0.003	0.009	-0.006	0.16	0.16	0.51
β	0.72	0.81	0.85	0.96	1.05	1.07	1.12	0.96	0.08	0.96	0.51	0.51	0.67
R^2	0.44	0.57	0.63	0.77	0.86	0.89	0.92	0.75	0.10	0.51	0.02	0.02	0.02
F-stat	0.01	0.07	0.14	0.92	3.37	4.58	7.85	1.45	1.67	3.76	0.92	0.08	0.08
p-Val	0.00	0.01	0.04	0.41	0.87	0.93	0.99	0.43	0.30				
α std.	0.007	0.007	0.008	0.010	0.012	0.012	0.013	0.010	0.001				
β std.	0.04	0.05	0.06	0.07	0.10	0.10	0.12	0.08	0.02				