The Forecast Quality of CBOE Implied Volatility Indexes

Charles J. Corrado
University of Auckland
New Zealand

Thomas W. Miller, Jr.
Washington University
St. Louis, MO

June 2003

Abstract

We examine the forecast quality of Chicago Board Options Exchange (CBOE) implied volatility indexes based on the Standard and Poor’s 100 and Nasdaq 100 stock indexes. We find that the forecast quality of CBOE implied volatilities for the S&P 100 (VIX) has significantly improved in recent years, and implied volatilities for the Nasdaq 100 (VXN) provide even higher quality forecasts of future realized volatility. CBOE implied volatilities appear to contain significant forecast errors in the period 1988-94, but we find no indication of significant forecast errors in the period 1995-2002.

JEL classification: C13, C22, C53, G13, G14

Keywords: options; implied volatility; volatility forecasting

We wish to acknowledge the helpful comments and suggestions of seminar participants at the University of Melbourne, University of Technology – Sydney, and University of Verona - Italy. Please direct correspondence to: Charles Corrado, Dept. of Accounting and Finance, Private Bag 92019, University of Auckland, Auckland, New Zealand.
(c.corrado@auckland.ac.nz)
The Forecast Quality of CBOE Implied Volatility Indexes

1. Introduction

Most investors might agree that stock prices follow a path of maximum aggravation—such is the nettlesome character of market volatility. Indeed, volatility is widely regarded as a measure of investor sentiment. Because volatility implied by option prices represent a market-based estimate of future price volatility, implied volatility is often regarded as a fear gauge (Whaley, 2000). Implied volatilities are routinely reported by financial news services and widely followed by investors and other finance professionals. Consistent with this widespread acceptance, the information content and forecast quality of implied volatility is an important topic in financial markets research.

Latane and Rendleman (1976), Chiras and Manaster (1978), and Beckers (1981) provide early assessments of the forecast quality of implied volatility. They found that implied volatilities offer better estimates of future return variability than ex post standard deviations calculated from historical returns data. More recently, Jorion (1995) finds that implied volatilities from currency options outperform volatility forecasts from historical price data.

In marked contrast to the first studies cited above, several later studies have found weaknesses in implied volatility as a predictor of future realized volatility; these include Day and Lewis (1988), Lamoureux and Lastrapes (1993), and Canina and Figlewski (1993). Christenson and Prabhala (1998) suggest that some of these weaknesses could be related to methodological issues, such as overlapping and mismatched sample periods. The validity of these concerns is supported by Fleming (1993) and Fleming, Ostdiek, and Whaley (1995), who find that implied volatilities from S&P 100 index options yield efficient forecasts of month-ahead S&P 100 index return volatility. Further studies of the performance of S&P 100 implied volatility by Christensen and Prabhala (1998) and Fleming (1998) find that implied volatility forecasts are upwardly biased, but dominate historical volatility in terms of ex ante forecasting power. Fleming (1999) shows that the forecast bias of S&P 100 implied volatility is not economically significant after accounting for transaction costs linked to the bid-ask spread. More recently, Blair, Poon, and Taylor (2001) conclude that implied volatilities from
S&P 100 index option prices provide more accurate volatility forecasts than those obtained from either low- or high-frequency index returns.

In this paper, we examine the forecast quality of two implied volatility indexes published by the Chicago Board Options Exchange (CBOE). These volatility indexes are reported under the ticker symbols VIX and VXN. The VIX volatility index is based on the Standard & Poor’s 100 stock index, with ticker symbol OEX. The VXN volatility index is based on the Nasdaq 100 stock index, with ticker symbol NDX. Current volatility index values for VIX and VXN are accessible in real time, as are current stock index values for OEX and NDX. Like previous studies of volatility implied by option prices, our benchmark for comparison is return volatility for the underlying index realized during the life of the option. We find that the VIX and VXN volatility indexes published by the CBOE easily outperform historical volatility as predictors of future return volatility for both the S&P 100 index (OEX) and the Nasdaq 100 index (NDX). We also find that the econometric problem of errors in variables has disappeared from CBOE volatility index data since 1995. Specifically, instrumental variable regressions do not yield assessments of forecast quality that are consistently superior to those obtained from ordinary least-squares (OLS) regressions.

This paper is organized as follows: In the next section, we present the volatility measures used in this study and summarize their basic statistical properties. A framework for analysis of volatility forecasts from realized and implied volatility measures is developed in Section 3. In Section 4, we present an empirical assessment of the forecast quality of CBOE volatility indexes using ordinary least-squares (OLS) regressions. Assessments based on an instrumental variables methodology are presented in Section 5. In Section 6 we analyze the statistical significance of volatility forecast errors embodied in CBOE implied volatilities. The summary and conclusion follow in Section 7.

2. Data sources and volatility measures

2.1 Data sources

Data for this study span the period January 1988 through December 2002, and include index returns and option-implied volatilities for the Standard and Poor’s 100 stock index and the Nasdaq 100 stock index. Index returns are computed from index data published by Reuters under the ticker symbols OEX for the S&P 100 index and NDX for the Nasdaq 100 index. Option-implied volatilities for these indexes are supplied by the Chicago Board Options Exchange (CBOE).
2.2 Volatility measures

Two volatility measures are used in this study. The first volatility measure is the sample standard deviation of daily index returns, which serves as the benchmark for this study. Annualized index return volatility within month $m$ is computed for each calendar month in the sample period as defined in equation (1).

\[
VOL_m = \sqrt{\frac{30}{22} \times \frac{252}{n_m - 1} \sum_{d=1}^{n_m} \left( \frac{r_{d,m}}{n_m} - \frac{1}{n_m} \sum_{h=1}^{n_m} r_{h,m} \right)^2}
\]

In equation (1), $r_{d,m}$ represents an index return on day $d$ in month $m$, and $n_m$ is the number of trading days in month $m$. The volatility measure $VOL_m$ is computed separately in each month and represents a series of non-overlapping monthly return standard deviations for the S&P 100 and Nasdaq 100 indexes.

The adjustment factor $\sqrt{30/22}$ embedded in equation (1) produces a volatility series that conforms to the same 22-trading day basis to which CBOE implied volatilities are calibrated. As explained in Fleming, Ostdiek, and Whaley (1995), the CBOE implied volatility calculations convert calendar days to trading-days via this function:

\[
\text{Trading days} = \text{Calendar days} - 2 \times \text{int} \left( \text{Calendar days} / 7 \right)
\]

The conversion from 30 calendar days to 22 trading days yields an adjustment of $\sqrt{30/22}$ (1.1677), which is required to restate the annualized return standard deviation to a 22-trading day basis. This calibration is necessary to achieve comparability between the realized volatility series, $VOL_m$, and the CBOE implied volatility series $VIX$ and $VXN$. As discussed in Bilson (2003), the essentially identical adjustment $\sqrt{7/5}$ (1.1832) also effectively calibrates volatility measures to the same time basis.

The second volatility measure, CBOE implied volatility, is the primary focus of this study. The Chicago Board Options Exchange (CBOE) provides two volatility series reported under the ticker symbols $VIX$ and $VXN$ derived from options traded on the S&P 100 and Nasdaq 100 indexes, respectively. The $VIX$ and $VXN$ volatility series are computed as weighted averages of individual implied volatilities for eight near-the-money call and put options from two nearby option expiration dates. Harvey and Whaley (1991) show that S&P 100 put and call implied volatilities are negatively correlated and so combining them results in a more efficient estimator. Corrado and Miller (1996) and Ederington and Guan (2002) analyze various weighting schemes and find that the weighting scheme used in the CBOE volatility indexes is expected to be as efficient as any other suggested in the literature.
A binomial tree equivalent to the option-pricing formula of Black and Scholes (1973) is used to compute CBOE implied volatilities. Authoritative references for the exact algorithm used to compute VIX and VXN are Whaley (1993) and Fleming, Ostdiek, and Whaley (1995). In this study, $VIX_m$ and $VXN_m$ denote implied volatilities for S&P 100 and Nasdaq 100 indexes observed at the close of the last trading day of month $m$, and so represent market forecasts of future return volatility in month $m+1$.

2.3 Data summary statistics

Data for this study are divided between two sample periods, the 84-month period January 1988 through December 1994, and the 96-month period January 1995 through December 2002. The latter period coincides with the availability of VXN implied volatility data for the Nasdaq 100 index. In total, there are 180 monthly observations for the VIX S&P 100 volatility series and 96 monthly observations for the VXN Nasdaq 100 volatility series. Summary descriptive statistics for these volatility data are provided in Table 1.

Table 1 reveals noticeable differences between the realized and implied volatility series. For the S&P 100 index, the average implied volatility, $VIX_m$, for the S&P 100 index is greater than average realized volatility, $VOL_m$, by $3.09\% = 17.84\% - 14.75\%$ over the period 1988-94, and $2.90\% = 24.01\% - 21.11\%$ over the period 1995-2002. For the Nasdaq 100 index over the period 1995-2002, the average implied volatility, $VXN_m$, is greater than average realized volatility, $VOL_m$, by $1.47\% = 41.82\% - 40.35\%$.

Figures 1 and 2 provide a graphical display of the time series of CBOE implied volatilities and corresponding realized volatilities. Figure 1 plots implied and realized volatilities, $VIX_{m-1}$ and $VOL_m$, respectively, for the S&P 100 index over the 15-year period 1988-2002. Figure 2 plots implied and realized volatilities, $VXN_{m-1}$ and $VOL_m$, respectively, for the Nasdaq 100 index over the 8-year period 1995-2002. Implied volatilities are plotted with solid lines and realized volatilities are plotted with dashed lines. These volatility series are synchronized so that the realized volatility in month $m$, i.e., $VOL_m$, is aligned with implied volatility observed on the last trading day of month $m-1$, i.e., $VIX_{m-1}$ or $VXN_{m-1}$. Differences between realized volatility in month $m$ and implied volatility observed on the last trading day of month $m-1$ represent observed forecast errors.

The $T$-statistic immediately below assesses the statistical significance of average forecast errors between realized and implied volatility.
\[ T_{M-1} = \frac{1}{\sqrt{M}} \sum_{m=1}^{M} (VOL_m - IVOL_{m-1}) \]

This statistic yields \( T \)-values of -7.07 and -3.96 for the S&P 100 index in the periods 1988-94 and 1995-2002, respectively, and -1.145 for the Nasdaq 100 index in the period 1995-2002. These \( T \)-values indicate a statistically significant forecast bias for S&P 100 index volatility forecasts, but an insignificant bias for Nasdaq 100 index volatility forecasts.

At least part of the observed forecast bias might be attributed to the algorithm used to compute implied volatility. For example, Fleming, Ostdiek, and Whaley (1995) show that 35 basis points of the difference between S&P 100 implied and realized volatilities is explained by intraday effects associated with the algorithm used to compute CBOE implied volatilities. This type of bias would also affect Nasdaq 100 implied volatility, because it is computed with the same algorithm. Fleming and Whaley (1994) report that an additional bias of about 60 basis points is attributable to the wildcard option embedded in the American-style exercise of S&P 100 index options. The Nasdaq 100 index options have European-style exercise, so they do not have a wildcard feature.

### 2.4 Naive volatility forecasts

Visible co-movement between the volatility time series displayed in Figures 1 and 2 suggest that naive forecasts might be used to forecast future return volatility successfully. For example, consider naive volatility forecasts based on a weighted average of lagged realized volatility \( VOL_{m-1} \) and implied volatility \( IVOL_{m-1} \). We evaluate three cases using the mean square error criteria stated immediately below: \( \alpha = 1, \alpha = 0, \) and \( \alpha = \alpha^* \), where \( \alpha^* \) is chosen to minimize mean square error subject to \( 0 \leq \alpha \leq 1 \).

\[ MSE(\alpha) = \frac{1}{M} \sum_{m=1}^{M} (VOL_m - \alpha \times VOL_{m-1} - (1 - \alpha) \times IVOL_{m-1})^2 \] \hspace{1cm} (2)

Mean square errors for these naive forecasts are reported in Table 2.

As shown in Table 2, the \( \alpha = 0 \) case yields consistently smaller mean square errors than the \( \alpha = 1 \) case, indicating that the implied volatilities \( VIX_{m-1} \) and \( VXN_{m-1} \) for the S&P 100 and Nasdaq 100 indexes, respectively, dominate lagged volatility \( VOL_{m-1} \) in forming naive forecasts of realized volatility \( VOL_m \). The \( \alpha = \alpha^* \) case yields only a modest improvement for the S&P 100 index, while \( \alpha^* = 0 \) is optimal for the Nasdaq 100 index.
3. A model framework for assessing volatility forecasts

In this section, we develop a framework for a further analysis of monthly volatility forecasts. This framework may be interpreted as a null hypothesis for empirical testing, or as a model for interpretation of empirical results.

3.1 Specification of model variables

Volatility realized in month \( m \) is denoted by \( VOL_m \) as specified in equation (1). We assume that realized volatility has two components: a latent volatility, \( \sigma_m \), evolving according to an underlying economic model, and a random deviation, \( \chi_m \), of realized volatility from latent volatility. We further assume that the deviations, \( \chi_m \), are mean zero, independently distributed random variables.

\[
\begin{align*}
VOL_m &= \sigma_m + \chi_m \\
E(\chi_m) &= 0 \\
E(\sigma_m \chi_m) &= 0 \\
\end{align*}
\]  

(3)

Consequently, the total variance of realized volatility is a sum of component variances.

\[
\text{Var}(VOL_m) = \text{Var}(\sigma_m) + \text{Var}(\chi_m) \\
\]  

(4)

The assumptions underlying equations (3) and (4) are essentially those made by Andersen and Bollerslev (1998) in the context of daily volatility forecasts. They argue that rational volatility forecasts represent predictions of latent volatility \( \sigma_m \), not realized volatility \( VOL_m \), because the deviations \( \chi_m \) are an unpredictable noise component.

Equations (3) and (4) imply that a regression of current volatility \( VOL_m \) on lagged volatility \( VOL_{m-1} \) yields a regression slope coefficient attenuated by an errors-in-variables bias.

\[
\lim_{M \to \infty} \frac{\sum_{m=2}^{M} (VOL_m VOL_{m-1} - \overline{Vol}^2) }{\sum_{m=2}^{M} (Vol^2_{m-1} - \overline{Vol}^2) } = \frac{\text{Cov}(\sigma_m, \sigma_{m-1})}{\text{Var}(\sigma_{m-1}) + \text{Var}(\chi_{m-1})} \\
\]  

(5)

Nota bene, \( \text{Var}(\sigma_{m-1}) \), \( \text{Var}(\chi_{m-1}) \), and \( \text{Var}(\xi_{m-1}) \) (introduced immediately below) are asymptotically equivalent to \( \text{Var}(\sigma_m) \), \( \text{Var}(\chi_m) \), and \( \text{Var}(\xi_m) \), respectively. We retain the finite sample index \( m-1 \) for convenience in referring to the originating finite sample formula.

A similar framework is applied to implied volatility. Specifically, let \( IVOL_{m-1} \) denote an option-implied volatility observed at the end of month \( m-1 \). We assume that \( IVOL_{m-1} \) represents a market forecast of the latent volatility \( \sigma_m \), with a random forecast error \( \xi_{m-1} \).
\[ IVOL_{m-1} = E_{m-1}(\sigma_m) = \sigma_m + \xi_{m-1} \quad (6) \]

\[ E(\xi_{m-1}\chi_m) = 0 \quad \text{Cov}(\xi_{m-1}\sigma_m) = 0 \]

The assumed zero covariance between forecast errors \( \xi_{m-1} \) and latent volatility \( \sigma_m \) allows a positive forecast bias, i.e., \( E(\xi_{m-1}) > 0 \), while still implying that the total variance of implied volatility is a sum of component variances.

\[ Var(IVOL_{m-1}) = Var(\sigma_m) + Var(\xi_{m-1}) \quad (7) \]

Equations (3), (6), and (7) imply that a regression of realized volatility \( VOL_m \) on implied volatility \( IVOL_{m-1} \) yields a regression slope coefficient attenuated by an errors-in-variables bias.

\[
\lim_{M \to \infty} \frac{\sum_{m=2}^{M} (Vol_m IVOL_{m-1} - \overline{Vol IVol})}{\sum_{m=2}^{M} (IVOL_{m-1}^2 - IVol^2)} = \frac{Var(\sigma_m)}{Var(\sigma_m) + Var(\xi_{m-1})} \quad (8)
\]

Equation (8) suggests that a one-sided test for a slope coefficient significantly less than one is equivalent to a test for a significant forecast error variance \( Var(\xi_{m-1}) \).

The errors-in-variables bias also affects multivariate regressions of current volatility, \( VOL_m \), on implied volatility, \( IVOL_{m-1} \), and lagged volatility, \( VOL_{m-1} \).

\[ VOL_m = b_0 + b_1 \times IVOL_{m-1} + b_2 \times VOL_{m-1} \quad (9) \]

Asymptotic values for the slope coefficients \( b_1 \) and \( b_2 \) in equation (9) within the framework developed above reflect biases induced by the variances, \( Var(\chi_m) \) and \( Var(\xi_{m-1}) \).

\[ \lim_{M \to \infty} b_1 = 1 - \frac{Var(\xi_{m-1})[Var(\sigma_m) + Var(\chi_{m-1})]}{D} \]

\[ \lim_{M \to \infty} b_2 = \frac{Cov(\sigma_m, \xi_{m-1}) Var(\xi_{m-1})}{D} \]

\[ D = [Var(\sigma_m) + Var(\xi_{m-1})][Var(\sigma_{m-1}) + Var(\chi_{m-1})] - Cov^2(\sigma_m, \xi_{m-1}) \]

Equation (10) reveals that the slope coefficient \( b_1 \) is biased downward below one and the slope coefficient \( b_2 \) is biased upward above zero. However, as the forecast error variance \( Var(\xi_{m-1}) \) diminishes, the slope coefficient \( b_1 \) approaches unity and \( b_2 \) approaches zero.
4. OLS volatility forecast regressions

Christenson and Prabhala (1998) and Fleming, Ostdiek, and Whaley (1995) point out that implied volatilities may contain observation errors that could affect regressions using implied volatility as an independent variable. However, Fleming, Ostdiek, and Whaley (1995) argue that observation error is minimized in CBOE volatility indexes because an equal number of call and put options are used to compute volatility index values. Nevertheless, as discussed in the previous section, CBOE implied volatilities may still contain forecast errors that could affect regressions using implied volatility as an independent variable.

Christenson and Prabhala (1998) use log-transformed data in their regressions, i.e., \( \ln V_{OLm} \) and \( \ln IV_{OLm} \), as the skewness and kurtosis of log-transformed volatility is closer to that of a normal distribution. Table 1 reveals that this is true for data used in this study; particularly for volatility data from the period 1988-94. However, Fleming (1998), Fleming, Ostdiek, and Whaley (1995) and other studies use untransformed volatility data. There are reasons to support the use of both log-transformed and untransformed volatility data. Consequently, we perform parallel regressions using both the original volatility measures \( V_{OLm} \) and \( IV_{OLm} \), and the log-transformed volatility measures \( \ln V_{OLm} \) and \( \ln IV_{OLm} \).

4.1 Univariate forecast regressions

Table 3 contains empirical results from both univariate and multivariate forecast regressions. We first focus on univariate regressions comparing the ability of historical and implied volatility to forecast future realized volatility. In Table 3, regression parameter estimates are reported in columns two through four, with Newey and West (1987) standard errors shown in parentheses below each regression coefficient. We found no significant differences affecting our conclusions using either ordinary least squares or White (1980) standard errors. Column five lists adjusted \( R^2 \)-squared statistics for each regression. Column six reports chi-square statistics based on the Newey and West (1987) covariance matrix testing the joint null hypothesis of a zero intercept and unit slope, i.e., \( \alpha = 0, \beta = 1 \). The corresponding \( p \)-values appear in parentheses under each chi-square statistic. The last column lists Breusch (1978) and Godfrey (1978) statistics testing for serial dependencies in regression residuals, with corresponding \( p \)-values in parentheses below each Breusch-Godfrey statistic. We first discuss results obtained from S&P 100 volatility measures and then follow with results from Nasdaq 100 volatility measures.
4.1.a S&P 100 univariate regressions

Panels A and B of Table 3 report regression results for the S&P 100 index over the periods 1988-94 and 1995-2002, respectively. In the period 1988-94, univariate regressions of current volatility on lagged volatility, \( VOL_m \) on \( VOL_{m-1} \) and \( \ln VOL_m \) on \( \ln VOL_{m-1} \), yield slope coefficients of 0.334 and 0.468, respectively, that are significantly less than one. Slope coefficients for regressions of current volatility on implied volatility, \( VOL_m \) on \( VIX_{m-1} \) and \( \ln VOL_m \) on \( \ln VIX_{m-1} \), yield slightly higher slope coefficients of 0.623 and 0.805, respectively, but they are still significantly less than one.

For the period 1995-2002, Panel B reveals that univariate regressions of current volatility on lagged volatility, \( VOL_m \) on \( VOL_{m-1} \) and \( \ln VOL_m \) on \( \ln VOL_{m-1} \), yield slope coefficients of 0.612 and 0.694, respectively, that are significantly less than one. By contrast, regressions of current volatility on implied volatility, \( VOL_m \) on \( VIX_{m-1} \) and \( \ln VOL_m \) on \( \ln VIX_{m-1} \), yield slope coefficients of 0.916 and 1.147, respectively, that are insignificantly different from one. However, Newey-West chi-square statistics of 19.41 and 32.55 for these regressions reject the joint null hypothesis of a zero intercept and unit slope, i.e., \( \alpha = 0, \beta = 1 \).

Figures 3 and 4 provide scatter plots of S&P 100 realized volatility \( VOL_m \) against implied volatility \( VIX_{m-1} \) for the periods 1988-94 and 1995-2002, respectively. For reference, both figures contain a solid line with zero intercept and unit slope (\( \alpha = 0, \beta = 1 \)) along with a dashed line representing an OLS fit to the data. In Figure 3, which corresponds to the period 1988-94, the significant bias of the OLS slope coefficient is visually obvious. However, in Figure 4, which represents the period 1995-2002, the dashed OLS line is nearly parallel to the solid unit slope reference line.

4.1.b Nasdaq 100 univariate regressions

Panel C of Table 3 reports regression results for the Nasdaq 100 index over the period 1995-2002. For this period, univariate regressions of current volatility on lagged volatility, \( VOL_m \) on \( VOL_{m-1} \) and \( \ln VOL_m \) on \( \ln VOL_{m-1} \), yield slope coefficients of 0.662 and 0.744, respectively, that are significantly less than one. However, slope coefficients for regressions of current volatility on implied volatility, \( VOL_m \) on \( VXN_{m-1} \) and \( \ln VOL_m \) on \( \ln VXN_{m-1} \), yield slope coefficients of 1.051 and 1.067, respectively, which are not statistically different from one. For these implied volatility regressions, the Newey-West chi-square statistic of 4.82 is insignificant at the 5-percent significance level, but the chi-square statistic of 9.21 is
significant at the 1-percent level, thereby rejecting the joint null hypothesis of a zero intercept and unit slope.

Figure 5 provides a scatter plot of Nasdaq 100 realized volatility, $VOL_m$, against implied volatility, $VXN_{m-1}$, for the period 1995-2002. The dashed OLS regression line in this figure is approximately congruent with the solid reference line with zero intercept and unit slope. The Nasdaq 100 regression results reported in Table 3 and displayed in Figure 5 yield strong graphic support for the Nasdaq 100 volatility index, $VXN$, as an efficient predictor of future realized volatility.

4.2 Multivariate forecast regressions

4.2.a S&P 100 multivariate regressions

Panel A of Table 3 reports multivariate regression results using both log-transformed and untransformed volatility data for the S&P 100 index for the period 1988-94. The regression of realized volatility, $VOL_m$, on implied volatility, $VIX_{m-1}$, and lagged volatility, $VOL_{m-1}$, yields slope coefficients of 0.830 and -0.234, respectively. The chi-square statistic of 4.53 for this regression does not reject the joint null hypothesis of a zero intercept and slope coefficient of one for $VIX_{m-1}$ at conventional significance levels. The regression of $\ln VOL_m$ on $\ln VIX_{m-1}$ and $\ln VOL_{m-1}$ yields slope coefficients of 0.978 and -0.164, respectively, and the chi-square statistic of 1.18 does not reject the joint null of a zero intercept and slope coefficient of one for $\ln VIX_{m-1}$. The insignificant Breusch-Godfrey statistics for these multivariate regressions do not indicate the presence of significant serial dependence in regression residuals.

Multivariate regression results for S&P 100 volatility for the period 1995-2002 are reported in Panel B of Table 3. Regressing $VOL_m$ on $VIX_{m-1}$ and $VOL_{m-1}$ yields slope coefficients of 0.840 and 0.069, respectively. Again, the chi-square statistic of 3.00 does not reject the joint null hypothesis of a zero intercept and unit slope coefficient for $VIX_{m-1}$. The regression of $\ln VOL_m$ on $\ln VIX_{m-1}$ and $\ln VOL_{m-1}$ yields slope coefficients of 0.951 and 0.145, respectively. However, the chi-square statistic of 8.40 rejects the joint null of a zero intercept and unit slope coefficient for $\ln VIX_{m-1}$.

An important aspect of the regressions based on S&P 100 volatility data is the fact that adjusted $R$-squared values from multivariate regressions do not exhibit substantial differences from adjusted $R$-squared values obtained from univariate regressions using only implied volatility measures $VIX_{m-1}$ or $\ln VIX_{m-1}$ as independent variables. Thus, on the basis of
adjusted $R^2$-squared values, adding independent variables beyond implied volatility does not appear to improve the explanatory power of the regressions.

4.2.2 Nasdaq 100 multivariate regressions

Results from multivariate regressions for the Nasdaq 100 for the period 1995-2002 are reported in Panel C of Table 3. The regression of $VOL_m$ on $VXN_{m-1}$ and $VOL_{m-1}$ yields slope coefficients of 1.278 and -0.185, respectively, with corresponding standard errors of 0.127 and 0.083. For this regression, the chi-square statistic of 4.81 rejects the joint null hypothesis of a zero intercept and unit slope coefficient for $VXN_{m-1}$. The regression of $\ln VOL_m$ on $\ln VXN_{m-1}$ and $\ln VOL_{m-1}$ yields slope coefficients of 1.242 and -0.151, respectively. However, the chi-square statistic of 3.76 for this regression does not reject the joint null hypothesis of a zero intercept and unit slope for $\ln VXN_{m-1}$. Breusch-Godfrey statistics for these multivariate regressions are not significant.

Nasdaq 100 volatility regressions share a similar characteristic with the S&P 100 volatility regressions in that adjusted $R^2$-squared values from multivariate regressions do not exhibit substantial differences from adjusted $R^2$-squared values obtained from univariate regressions using only implied volatility, $VIX_{m-1}$ or $\ln VIX_{m-1}$, as an independent variable. Consequently, it does not appear that additional independent variables beyond implied volatility improve the explanatory power of these regressions.

5. Instrumental variable regressions

The econometric problem of errors in explanatory variables is widely accepted as an impediment to assessing the forecast quality of implied volatility. The standard econometric approach to dealing with this problem is the use of instrumental variables [see, for example, Greene (1993), Johnston (1984), or Maddala (1977)]. Drawing on the analysis in Greene (1993), Christensen and Prabhala (1998) propose using lagged implied volatility as an instrument for implied volatility.

5.1 Asymptotic coefficients with instrumental variables

A framework for analysing volatility forecast regressions was developed in Section 3, and is here applied to analyzing instrumental variable regressions. In the univariate instrumental variable procedure, the first- and second-step regressions are stated immediately below. A hat indicates a fitted value from the first-step regression.
The slope coefficient, $c_1$, of the first-step regression has this asymptotic value:

$$
\lim_{M \to \infty} c_1 = \frac{\sum_{m=3}^{M} \left( IVOL_{m-3} IVOL_{m-2} - \overline{IVOL}^2 \right)}{\sum_{m=3}^{M} \left( IVOL_{m-2}^2 - \overline{IVOL}^2 \right)} = \frac{\text{Cov}(\sigma_m, \sigma_{m-1})}{\text{Var}(\sigma_{m-1}) + \text{Var}(\xi_{m-2})} \quad (12)
$$

The asymptotic value of the slope coefficient, $a_1$, for the second-step regression has an asymptotic value of one.

$$
\lim_{M \to \infty} a_1 = \frac{\sum_{m=3}^{M} c_1 \left( IVOL_{m} IVOL_{m-2} - \overline{IVOL}^2 \right)}{\sum_{m=3}^{M} c_1^2 \left( IVOL_{m-2}^2 - \overline{IVOL}^2 \right)} = \frac{\text{Var}(\sigma_{m-1}) + \text{Var}(\xi_{m-2})}{\text{Cov}(\sigma_m, \sigma_{m-1}) \times \text{Var}(\sigma_{m-1}) + \text{Var}(\xi_{m-2})} = 1 \quad (13)
$$

In the multivariate instrumental variable procedure, the first- and second-step regressions are:

$$
\overline{IVOL}_{m-1} = c_0 + c_1 \times IVOL_{m-2} + c_2 \times VOL_{m-1}
\quad (14)
\quad VOL_m = b_0 + b_1 \times \overline{IVOL}_{m-1} + b_2 \times VOL_{m-1}
$$

The asymptotic coefficient, $c_1$, of the first-step regression is:

$$
\lim_{M \to \infty} c_1 = \frac{\text{Cov}(\sigma_m, \sigma_{m-1})}{\text{Var}(\sigma_{m-1})} \left( 1 + \text{Var}(\xi_{m-2}) \left( 1/\text{Var}(\sigma_{m-1}) + 1/\text{Var}(\sigma_{m-1}) \right) \right) \quad (15)
$$

The slope coefficient, $b_1$, in the second-step multivariate regression can be shown to have an asymptotic value of one, i.e., $\lim_{M \to \infty} b_1 = 1$. However, the slope coefficient, $b_2$, in the second-step multivariate regression is asymptotically positive.

$$
\lim_{M \to \infty} b_2 = \frac{\text{Cov}(\sigma_m, \sigma_{m-1})}{\text{Var}(\sigma_{m-1})} \left( 1 + \text{Var}(\xi_{m-2}) \left( 1/\text{Var}(\sigma_{m-1}) + 1/\text{Var}(\sigma_{m-1}) \right) \right) \quad (16)
$$

Table 4 reports the results of instrumental variable regressions specified in equations (11) and (14) above. Separate results are reported for the S&P 100 for the periods 1988-94 and 1995-2002, and for the Nasdaq 100 for the period 1995-2002.
5.2 Instrumental variable regression results

5.2.a S&P 100 instrumental variable regressions

Panel A of Table 4 reports results from univariate and multivariate instrumental variable regressions based on the S&P 100 index for the period 1988-94. The univariate regression of realized volatility, $VOL_m$, on the implied volatility instrument, $\hat{VIX}_{m-1}$, yields a slope coefficient of 0.569, which is clearly significantly less than one. By contrast, the univariate regression of log-volatility $\ln VOL_m$ on the log-implied volatility instrument, $\ln \hat{VIX}_{m-1}$, yields a slope coefficient of 0.996, which is not significantly different from one. Nevertheless, the chi-square statistics of 81.47 and 81.95 reject the null hypothesis of a zero intercept and unit slope in both univariate regressions.

The multivariate regression of realized volatility, $VOL_m$, on the implied volatility instrument, $\hat{VIX}_{m-1}$, and lagged volatility, $VOL_{m-1}$, yields slope coefficients of 0.785 and -0.235, respectively. The chi-square statistic of 5.70 for this regression does not reject the joint null hypothesis of a zero intercept and unit slope for the instrument $\hat{VIX}_{m-1}$ at the 5-percent significance level. The regression of log-volatility $\ln VOL_m$ on the instrument $\ln \hat{VIX}_{m-1}$ and lagged log-volatility $\ln VOL_{m-1}$ yields slope coefficients of 1.330 and -0.245. The chi-square statistic of 3.44 for this regression does not reject the null of a zero intercept and unit slope for the instrument $\ln \hat{VIX}_{m-1}$.

Panel B of Table 4 reports univariate and multivariate regression results for the S&P 100 in the period 1995-2002. The univariate regression of realized volatility $VOL_m$ on the implied volatility instrument $\hat{VIX}_{m-1}$ yields a slope coefficient of 0.896 with a standard error of 0.113, indicating a value insignificantly less than one. By contrast, the univariate regression of log-volatility $\ln VOL_m$ on the log-implied volatility instrument $\ln \hat{VIX}_{m-1}$ yields a slope coefficient of 1.389 with a standard error of 0.143, indicating a value significantly greater than one. Chi-square statistics of 94.97 and 95.31 reject the null hypothesis of a zero intercept and unit slope for both univariate regressions.

The multivariate regression of realized volatility, $VOL_m$, on the implied volatility instrument, $\hat{VIX}_{m-1}$, and lagged volatility, $VOL_{m-1}$, yields slope coefficients of 0.885 and 0.011, respectively. However, the chi-square statistic of 16.00 for this regression rejects the null of a zero intercept and unit slope for the instrument $\hat{VIX}_{m-1}$. The regression of log-volatility $\ln VOL_m$ on the instrument $\ln \hat{VIX}_{m-1}$ and lagged log-volatility, $\ln VOL_{m-1}$, yields
slope coefficients of 1.245 and 0.093, respectively. The chi-square statistic for this regression of 20.05 also rejects the null of a zero intercept and unit slope for the instrument $\ln VIX_{m-1}$.

5.2.b Nasdaq 100 instrumental variable regressions

Panel C of Table 4 reports univariate and multivariate instrumental variable regression results for the Nasdaq 100 for the period 1995-2002. The univariate regression of realized volatility, $VOL_m$, on the implied volatility instrument, $\hat{VXN}_{m-1}$, yields a slope coefficient of 1.205 with a standard error of 0.099, indicating a slope significantly greater than one. The univariate regression of log-volatility, $\ln VOL_m$, on the log-implied volatility instrument, $\ln \hat{VXN}_{m-1}$, yields a slope coefficient of 1.185 with a standard error of 0.091. So this slope is also significantly greater than one. Chi-square statistics of 82.06 and 68.18 both reject the null hypothesis of a zero intercept and unit slope for the implied volatility instruments.

The multivariate regression of realized volatility, $VOL_m$, on the implied volatility instrument, $\hat{VXN}_{m-1}$, and lagged volatility, $VOL_{m-1}$, yields slope coefficients of 1.438 and -0.185. The chi-square statistic of 7.43 for this regression rejects the joint null of a zero intercept and unit slope for the instrument $\hat{VXN}_{m-1}$. The regression of log-volatility, $\ln VOL_m$, on the instrument, $\ln \hat{VXN}_{m-1}$, and lagged log-volatility, $\ln VOL_{m-1}$, yields slope coefficients of 1.356 and -0.134. The chi-square statistic for this regression of 6.52 also rejects the null of a zero intercept and unit slope for the instrument $\ln \hat{VXN}_{m-1}$.

5.3 Comparing conventional OLS and instrumental regressions

A comparison of regression results reported in Panels A of Tables 3 and 4 for the S&P 100 index for the period 1988-94 reveals that instrumental variable regressions generally did not offer an improvement upon simple OLS regressions. Specifically, in Table 3 the regression of $VOL_m$ on $VIX_{m-1}$ yields a slope coefficient of 0.623, while in Table 4 the regression of $VOL_m$ on the instrument $\hat{VIX}_{m-1}$ yields a smaller slope coefficient of 0.569. The corresponding slope coefficients for the multivariate regressions also decrease from 0.830 to 0.785. However, in Table 3 the univariate regression of $\ln VOL_m$ on $\ln VIX_{m-1}$ yields a slope coefficient of 0.805, while in Table 4 the regression of $\ln VOL_m$ on the instrument $\ln \hat{VIX}_{m-1}$ yields a slope coefficient of 0.996. The corresponding multivariate regressions have slope coefficients that rise from 0.978 for $\ln VIX_{m-1}$ to 1.330 for $\ln \hat{VIX}_{m-1}$.
Panels B in Tables 3 and 4 report OLS and instrumental variable regression results for the S&P 100 index for the period 1995-2002. In Table 3, the regression of \( V_{OL,m} \) on \( V_{IX,m-1} \) yields a slope coefficient of 0.916, while in Table 4 the regression of \( V_{OL,m} \) on the instrument \( \hat{V}_{IX,m-1} \) yields a slope coefficient of 0.896. The corresponding multivariate regression slopes increase from 0.840 to 0.885. In comparison, the univariate regression of \( \ln V_{OL,m} \) on \( \ln V_{IX,m-1} \) yields a slope coefficient of 1.147, which rises to 1.389 for the regression of \( \ln V_{OL,m} \) on the instrument \( \hat{\ln V}_{IX,m-1} \). The corresponding multivariate regression slopes rise from 0.951 for \( \ln V_{IX,m-1} \) to 1.245 for the instrument \( \hat{\ln V}_{IX,m-1} \).

Finally, comparing Panels C in Tables 3 and 4 for the Nasdaq 100 index for the period 1995-2002, the instrumental variables approach appears to yield consistently worse results in both univariate and multivariate regressions. Specifically, in Table 3, the regression of \( V_{OL,m} \) on \( V_{XN,m-1} \) yields a slope coefficient of 1.051, while in Table 4, the regression of \( V_{OL,m} \) on the instrument \( \hat{V}_{XN,m-1} \) yields a slope coefficient of 1.205. The corresponding slope coefficients in multivariate regressions rise from 1.278 to 1.438. Similarly, the univariate regression of \( \ln V_{OL,m} \) on \( \ln V_{XN,m-1} \) yields a slope coefficient of 1.067, while the regression of \( \ln V_{OL,m} \) on the instrument \( \hat{\ln V}_{XN,m-1} \) yields a slope coefficient of 1.185. The corresponding multivariate regressions have slope coefficients that rise from 1.242 for \( \ln V_{XN,m-1} \) to 1.356 for the instrument \( \hat{\ln V}_{XN,m-1} \).

In summary, attempts to correct a possible errors-in-variables bias via instrumental variable regressions do not appear consistently effective. Indeed, in only two cases among those examined did instrumental variable regressions yield regression slope coefficients for implied volatility closer to one than did standard OLS regressions: 1) the univariate OLS regression of \( \ln V_{OL,m} \) on \( \ln V_{IX,m-1} \) for the period 1988-94 yielded a slope coefficient of 0.805, while the regression of \( \ln V_{OL,m} \) on the instrument \( \hat{\ln V}_{IX,m-1} \) yielded a slope coefficient of 0.996; and 2) the multivariate OLS regression of \( V_{OL,m} \) on \( V_{IX,m-1} \) and \( V_{OL,m-1} \) for the period 1995-2002 yielded slope coefficients of 0.840 and 0.069, respectively, while the regression of \( V_{OL,m} \) on the instrument \( \hat{V}_{IX,m-1} \) and \( V_{OL,m-1} \) yielded corresponding slope coefficients of 0.885 and 0.011. In all other cases, instrumental variable regressions yielded slope coefficients for implied volatility farther from the desired value of one than did OLS regressions. This appears to suggest that any errors-in-variables bias in the \( V_{IX} \) and \( V_{XN} \) CBOE volatility indexes may not in fact be large.
6. Significance of the forecast errors of latent volatility

As suggested by the framework developed in Section 3, a test for a significant forecast error variance, \( \text{Var}(\xi_{m-1}) \), is obtained from a univariate regression of realized volatility, \( \text{VOL}_m \), on implied volatility, \( \text{IVOL}_{m-1} \). Equations (7) and (8) imply that one minus the slope coefficient of the regression of \( \text{VOL}_m \) on \( \text{IVOL}_{m-1} \) yields the following equality:

\[
1 - b_{\text{OLS}}(\text{VOL}_m, \text{IVOL}_{m-1}) = \frac{\text{Var}(\xi_{m-1})}{\text{Var}(\text{IVOL}_{m-1})}
\]  

(17)

In finite samples, the left-hand side of equation (17) above could be negative. Despite this, a one-sided test for a positive forecast error variance, i.e., \( \text{Var}(\xi_{m-1}) > 0 \), is equivalent to a test for an OLS slope coefficient \( b_{\text{OLS}} \) significantly less than one. The required regression statistics for this test are reported in Table 3.

From Panel A of Table 3 for the 1988-94 S&P 100 sample, we see that the slope coefficient of 0.623 with a standard error of 0.086 indicates rejection of the null hypothesis of \( b_{\text{OLS}} = 1 \), and therefore indicates a significantly positive forecast error variance. However, from Panel B of Table 3 for the 1995-2002 S&P 100 sample, the slope coefficient of 0.916 with a standard error of 0.087 does not reject the null hypothesis of \( b_{\text{OLS}} = 1 \). Also, from Panel C of Table 3 for the 1995-2002 Nasdaq 100 sample, the slope coefficient of 1.051 with a standard error of 0.081 does not reject the null hypothesis of \( b_{\text{OLS}} = 1 \). Thus, while CBOE implied volatilities appear to contain significant forecast errors in the early period 1988-94, we find no indication of significant forecast error variances in the latter period 1995-2002.

Point estimates of the variance of forecast errors, \( \text{Var}(\xi_{m-1}) \), may be obtained from the equality immediately below, which follows from equations (3) and (6) in Section 3.

\[
\text{Var}(\text{IVOL}_{m-1}) - \text{Covar}(\text{VOL}_m, \text{IVOL}_{m-1}) = \text{Var}(\xi_{m-1})
\]  

(18)

Based on equation (18) above, we obtain the following sample estimates of variances of forecast errors for the three data subsamples:

- \( \text{VIX} \) (1988-94): \( \text{Var}(\xi_{m-1}) = 10.41 \)
- \( \text{VIX} \) (1995-2002): \( \text{Var}(\xi_{m-1}) = 2.62 \)
- \( \text{VXN} \) (1995-2002): \( \text{Var}(\xi_{m-1}) = -13.72 \)

The regression tests discussed above indicate that the negative variance estimate of -13.72 taken from the Nasdaq 100 data and the positive estimate of 2.62 taken from the S&P 100
data in the 1995-2002 period are not statistically significant. However the variance estimate of 10.41 taken from the S&P 100 data in the earlier period 1988-94 is statistically significant.

7. Summary and conclusion

Assessing the information content and forecast quality of implied volatility has been an important and ongoing research issue among financial economists and econometricians. Early assessments lauded the forecast quality of option-implied volatility. Subsequent investigations found that option-implied volatility yielded biased and inefficient forecasts. More recent studies suggest that the econometric problem of errors in variables is a potential pitfall in assessing the forecast quality of implied volatility, and in turn propose the use of instrumental variable methodology to assess the forecast quality of implied volatility. This study compares the results of standard OLS regressions with those obtained from instrumental variable regressions, and finds that instrumental variable regressions typically yield inferior inferences in support of the forecast quality of implied volatility.

The empirical analysis in this study is based on the CBOE implied volatility indexes VIX and VXN, corresponding to S&P 100 and Nasdaq 100 stock indexes, respectively. These indexes provide an excellent data source for studies of implied volatility. Indeed, with the recent release of the VXN volatility index for the Nasdaq 100, the CBOE has significantly expanded a valuable data resource. In this study, we find that while the CBOE implied volatility indexes VIX and VXN yield upwardly biased volatility forecasts, they are still more efficient in terms of mean squared forecast errors than historical volatility as forecasts of future realized volatility.

Further regression analysis reveals that the highest R-squared values are obtained when implied volatility is the explanatory variable. Multivariate regressions indicate that adding historical volatility as an explanatory variable yields only trivial differences in regression R-squared values. Overall, the results reported in this paper suggest that the CBOE implied volatility indexes VIX and VXN dominate historical index volatility in providing forecasts of future price volatility for the S&P 100 and Nasdaq 100 stock indexes. While CBOE implied volatilities appear to contain significant forecast errors in the early period 1988-94, we find no indication of significant forecast errors in the latter period 1995-2002.
References


Table 1: Descriptive statistics for monthly volatility measures

Sample moments of monthly volatility measures: $VOL_m$ represents a realized volatility in month $m$ computed from daily returns within the month. $VIX_m$ and $VXN_m$ represent CBOE implied volatility indexes for the S&P 100 and Nasdaq 100 indexes, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Mean (%)</th>
<th>Std Dev (%)</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$VOL_m$</td>
<td>14.75</td>
<td>6.13</td>
<td>2.07</td>
<td>13.00</td>
</tr>
<tr>
<td>$VIX_m$</td>
<td>17.84</td>
<td>5.13</td>
<td>1.06</td>
<td>7.19</td>
</tr>
<tr>
<td>ln $VOL_m$</td>
<td>2.62</td>
<td>0.37</td>
<td>0.36</td>
<td>6.96</td>
</tr>
<tr>
<td>ln $VIX_m$</td>
<td>2.84</td>
<td>0.27</td>
<td>0.36</td>
<td>5.76</td>
</tr>
<tr>
<td>$VOL_m$</td>
<td>21.11</td>
<td>9.46</td>
<td>0.82</td>
<td>6.38</td>
</tr>
<tr>
<td>$VIX_m$</td>
<td>24.01</td>
<td>7.11</td>
<td>0.62</td>
<td>10.20</td>
</tr>
<tr>
<td>ln $VOL_m$</td>
<td>2.95</td>
<td>0.46</td>
<td>-0.28</td>
<td>5.72</td>
</tr>
<tr>
<td>ln $VIX_m$</td>
<td>3.13</td>
<td>0.31</td>
<td>-0.40</td>
<td>6.18</td>
</tr>
<tr>
<td>Panel C: Nasdaq 100 January 1995 – December 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$VOL_m$</td>
<td>40.35</td>
<td>18.65</td>
<td>1.18</td>
<td>7.39</td>
</tr>
<tr>
<td>$VXN_m$</td>
<td>41.82</td>
<td>14.74</td>
<td>0.70</td>
<td>5.87</td>
</tr>
<tr>
<td>ln $VOL_m$</td>
<td>3.60</td>
<td>0.44</td>
<td>0.10</td>
<td>5.61</td>
</tr>
<tr>
<td>ln $VIX_m$</td>
<td>3.67</td>
<td>0.35</td>
<td>0.03</td>
<td>5.23</td>
</tr>
</tbody>
</table>
Table 2: Mean squared errors of naive volatility forecasts

Mean squared errors (MSE) of naive volatility forecasts of realized volatility ($VOL_m$) based on weighted averages of lagged volatility ($VOL_{m-1}$) and implied volatility ($IVOL_{m-1}$) computed as follows:

\[ MSE(\alpha) = \frac{1}{M} \sum_{m=1}^{M} (VOL_m - \alpha \times VOL_{m-1} - (1-\alpha) \times IVOL_{m-1})^2 \]

Three cases are evaluated: $\alpha = 1$, $\alpha = 0$, and $\alpha = \alpha^*$, where $\alpha^*$ is chosen to minimize the mean square error subject to $0 \leq \alpha \leq 1$.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$MSE(1)$</td>
<td>39.67</td>
<td>66.69</td>
<td>202.07</td>
</tr>
<tr>
<td>$MSE(0)$</td>
<td>33.30</td>
<td>49.64</td>
<td>98.20</td>
</tr>
<tr>
<td>$MSE(\alpha^*)$</td>
<td>30.45</td>
<td>47.15</td>
<td>98.20</td>
</tr>
<tr>
<td>$\alpha^*$</td>
<td>0.357</td>
<td>0.263</td>
<td>0</td>
</tr>
</tbody>
</table>
OLS regressions of realized volatility ($VOL_m$) on CBOE implied volatilities for the S&P 100 index ($VIX_{m-1}$) or the Nasdaq 100 index ($VXN_{m-1}$), and lagged values of realized volatility ($VOL_{m-1}$). Multivariate regressions have this general form (with log-volatilities substituted in logarithmic regressions), where $IVOL_m$ denotes either $VIX_m$ or $VXN_m$ as appropriate.

$$VOL_m = a_0 + a_1 IVOL_{m-1} + a_2 VOL_{m-1}$$

Newey-West standard errors are reported in parentheses. Chi-square ($p$-value) corresponds to the null hypothesis of a zero intercept and unit slope ($\alpha = 0, \beta = 1$) in univariate regressions; and a null of a zero intercept and unit slope coefficient for implied volatility in multivariate regressions. B-G ($p$-value) corresponds to the Breusch-Godfrey test for autocorrelation in regression residuals.


<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>$VIX_{m-1}$</th>
<th>$VOL_{m-1}$</th>
<th>Adj. $R^2$</th>
<th>Chi-square ($p$-value)</th>
<th>B-G ($p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 100</td>
<td>2.986</td>
<td>0.623</td>
<td>0.348</td>
<td>0.348</td>
<td>107.84 (0.000)</td>
<td>2.95 (0.086)</td>
</tr>
<tr>
<td>$VOL_m$</td>
<td>(1.501)</td>
<td>(0.086)</td>
<td>(0.000)</td>
<td>(0.086)</td>
<td>(1.444) (0.0900)</td>
<td>(1.426) (0.126)</td>
</tr>
<tr>
<td></td>
<td>8.760</td>
<td>0.334</td>
<td>0.121</td>
<td>0.121</td>
<td>65.24 (0.000)</td>
<td>2.34 (0.000)</td>
</tr>
<tr>
<td></td>
<td>(1.444)</td>
<td>(0.0900)</td>
<td>(0.000)</td>
<td>(0.016)</td>
<td>(1.366) (0.107)</td>
<td>(1.366) (0.473)</td>
</tr>
<tr>
<td></td>
<td>2.898</td>
<td>-0.234</td>
<td>0.374</td>
<td>0.374</td>
<td>4.53 (0.104)</td>
<td>0.52 (0.473)</td>
</tr>
<tr>
<td></td>
<td>(1.366)</td>
<td>(0.130)</td>
<td>(0.107)</td>
<td>(0.107)</td>
<td>(0.284) (0.098)</td>
<td>(0.284) (0.161)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>$\ln VIX_{m-1}$</th>
<th>$\ln VOL_{m-1}$</th>
<th>Adj. $R^2$</th>
<th>Chi-square ($p$-value)</th>
<th>B-G ($p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 100</td>
<td>0.308</td>
<td>0.805</td>
<td>0.394</td>
<td>0.394</td>
<td>105.01 (0.000)</td>
<td>2.66 (0.103)</td>
</tr>
<tr>
<td>$\ln VOL_m$</td>
<td>(0.284)</td>
<td>(0.098)</td>
<td>(0.098)</td>
<td>(0.098)</td>
<td>(1.362) (0.089)</td>
<td>(1.362) (0.008)</td>
</tr>
<tr>
<td></td>
<td>1.362</td>
<td>0.468</td>
<td>0.186</td>
<td>0.186</td>
<td>39.82 (0.000)</td>
<td>6.96 (0.008)</td>
</tr>
<tr>
<td></td>
<td>(0.240)</td>
<td>(0.089)</td>
<td>(0.089)</td>
<td>(0.089)</td>
<td>(0.249) (0.161)</td>
<td>(0.249) (0.161)</td>
</tr>
<tr>
<td></td>
<td>0.249</td>
<td>-0.164</td>
<td>0.399</td>
<td>0.399</td>
<td>1.18 (0.554)</td>
<td>1.34 (0.248)</td>
</tr>
<tr>
<td></td>
<td>(0.292)</td>
<td>(0.122)</td>
<td>(0.122)</td>
<td>(0.122)</td>
<td>(0.292) (0.161)</td>
<td>(0.292) (0.161)</td>
</tr>
</tbody>
</table>
Table 3: continued


<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>$VIX_{m-1}$</th>
<th>$VOL_{m-1}$</th>
<th>Adj. $R^2$</th>
<th>Chi-square (p-value)</th>
<th>B-G (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 100 $VOL_m$</td>
<td>-0.801</td>
<td>0.916</td>
<td>0.508</td>
<td>19.41</td>
<td>(0.000)</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td>(1.943)</td>
<td>(0.087)</td>
<td>(0.000)</td>
<td>(0.073)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.307</td>
<td>0.612</td>
<td>0.369</td>
<td>32.18</td>
<td>(0.000)</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>(1.540)</td>
<td>(0.069)</td>
<td>(0.058)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.472</td>
<td>0.840</td>
<td>0.505</td>
<td>3.00</td>
<td>(0.223)</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>(1.977)</td>
<td>(0.162)</td>
<td>(0.067)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 100 $\ln VOL_m$</td>
<td>-0.627</td>
<td>1.147</td>
<td>0.596</td>
<td>32.55</td>
<td>(0.000)</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>(0.298)</td>
<td>(0.095)</td>
<td>(0.194)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.894</td>
<td>0.694</td>
<td>0.497</td>
<td>23.50</td>
<td>(0.000)</td>
<td>3.70</td>
</tr>
<tr>
<td></td>
<td>(0.195)</td>
<td>(0.064)</td>
<td>(0.054)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.446</td>
<td>0.951</td>
<td>0.597</td>
<td>8.40</td>
<td>(0.015)</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(0.322)</td>
<td>(0.186)</td>
<td>(0.600)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: continued

Panel C: Nasdaq 100 January 1995 – December 2002

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>VXN$_{m-1}$</th>
<th>VOL$_{m-1}$</th>
<th>Adj. $R^2$</th>
<th>Chi-square (p-value)</th>
<th>B-G (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasdaq 100</td>
<td>-3.466</td>
<td>1.051</td>
<td>0.716</td>
<td>4.82</td>
<td>0.29</td>
<td>(0.592)</td>
</tr>
<tr>
<td>VOL$_m$</td>
<td>(2.994)</td>
<td>(0.081)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11.735</td>
<td>0.662</td>
<td>0.479</td>
<td>18.79</td>
<td>7.56</td>
<td>(0.006)</td>
</tr>
<tr>
<td></td>
<td>(2.872)</td>
<td>(0.078)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5.079</td>
<td>-0.185</td>
<td>0.718</td>
<td>4.81</td>
<td>0.30</td>
<td>(0.584)</td>
</tr>
<tr>
<td></td>
<td>(3.104)</td>
<td>(0.083)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5.079</td>
<td>1.278</td>
<td>-0.185</td>
<td>0.718</td>
<td>4.81</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>(3.104)</td>
<td>(0.127)</td>
<td>(0.083)</td>
<td>(0.090)</td>
<td>(0.000)</td>
<td>(0.006)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>ln VXN$_{m-1}$</th>
<th>ln VOL$_{m-1}$</th>
<th>Adj. $R^2$</th>
<th>Chi-square (p-value)</th>
<th>B-G (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasdaq 100</td>
<td>-0.308</td>
<td>1.067</td>
<td>0.728</td>
<td>9.21</td>
<td>0.58</td>
<td>(0.447)</td>
</tr>
<tr>
<td>ln VOL$_m$</td>
<td>(0.254)</td>
<td>(0.069)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.906</td>
<td>0.744</td>
<td>0.545</td>
<td>17.34</td>
<td>11.62</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>(0.220)</td>
<td>(0.062)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.400</td>
<td>1.242</td>
<td>-0.151</td>
<td>0.729</td>
<td>3.76</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.266)</td>
<td>(0.126)</td>
<td>(0.092)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Instrumental variable regressions with realized and implied volatility measures

Instrumental variable regressions of realized volatility ($VOL_m$) on lagged values of volatility ($VOL_{m-1}$), and CBOE implied volatilities for the S&P 100 index ($VIX_{m-1}$) or the Nasdaq 100 index ($VXN_{m-1}$). $IVOL_m$ denotes either $VIX_m$ or $VXN_m$ as appropriate and $\hat{IVOL}$ denotes the instrumental variable for implied volatility.

Univariate regressions have this specification:

$$\hat{IVOL}_{m-1} = c_0 + c_1 \times IVOL_{m-2}$$
$$VOL_m = a_0 + a_1 \times \hat{IVOL}_{m-1}$$

Multivariate regressions have this specification:

$$\hat{IVOL}_{m-1} = c_0 + c_1 \times IVOL_{m-2} + c_2 \times VOL_{m-1}$$
$$VOL_m = b_0 + b_1 \times \hat{IVOL}_{m-1} + b_2 \times VOL_{m-1}$$

Newey-West standard errors are reported in parentheses. Chi-square ($p$-value) corresponds to the null hypothesis of a zero intercept and unit slope ($\alpha = 0$, $\beta = 1$) in univariate regressions; and a null of a zero intercept and unit slope coefficient for implied volatility in multivariate regressions. BG ($p$-value) corresponds to the Breusch-Godfrey test for autocorrelation in regression residuals.


<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>$\hat{VIX}_{m-1}$</th>
<th>$VOL_{m-1}$</th>
<th>Adj. $R^2$</th>
<th>Chi-square ($p$-value)</th>
<th>B-G ($p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 100</td>
<td>0.525</td>
<td>0.569</td>
<td>0.343</td>
<td>81.47</td>
<td>5.25</td>
<td>(0.022)</td>
</tr>
<tr>
<td>$VOL_m$</td>
<td>(1.679)</td>
<td>(0.093)</td>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.744</td>
<td>0.785</td>
<td>-0.235</td>
<td>0.363</td>
<td>5.70</td>
<td>3.81</td>
</tr>
<tr>
<td></td>
<td>(1.952)</td>
<td>(0.168)</td>
<td>(0.148)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ln $VOL_m$</td>
<td>ln $VOL_m$</td>
<td></td>
<td>81.95</td>
<td>5.25</td>
<td>(0.022)</td>
</tr>
<tr>
<td></td>
<td>-0.527</td>
<td>0.996</td>
<td>0.386</td>
<td>81.95</td>
<td>3.98</td>
<td>(0.046)</td>
</tr>
<tr>
<td></td>
<td>(0.431)</td>
<td>(0.151)</td>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.906</td>
<td>1.330</td>
<td>-0.245</td>
<td>0.387</td>
<td>3.44</td>
<td>2.90</td>
</tr>
<tr>
<td></td>
<td>(0.552)</td>
<td>(0.300)</td>
<td>(0.179)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>( \hat{VIX}_{m-1} )</th>
<th>( VOL_{m-1} )</th>
<th>Adj. ( R^2 )</th>
<th>Chi-square (p-value)</th>
<th>B-G (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 100 ( \text{VOL}_m )</td>
<td>-6.044 (2.784)</td>
<td>0.896 (0.113)</td>
<td>0.513 (0.000)</td>
<td>94.97 (0.082)</td>
<td>3.03 (0.051)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5.939 (3.312)</td>
<td>0.885 (0.221)</td>
<td>0.011 (0.178)</td>
<td>0.507 (0.051)</td>
<td>16.00 (0.082)</td>
<td>3.82 (0.040)</td>
</tr>
<tr>
<td>S&amp;P 100 ( \ln \text{VOL}_m )</td>
<td>-1.714 (0.449)</td>
<td>1.389 (0.143)</td>
<td>0.604 (0.000)</td>
<td>95.31 (0.239)</td>
<td>1.39 (0.040)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.507 (0.565)</td>
<td>1.245 (0.286)</td>
<td>0.093 (0.161)</td>
<td>0.601 (0.001)</td>
<td>20.05 (0.040)</td>
<td>4.23 (0.001)</td>
</tr>
</tbody>
</table>
Table 4: continued

Panel C: Nasdaq 100 January 1995 – December 2002

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>$\sqrt[3]{VXN_{m-1}}$</th>
<th>$VOL_{m-1}$</th>
<th>Adj. $R^2$</th>
<th>Chi-square (p-value)</th>
<th>B-G (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasdaq 100 $VOL_m$</td>
<td>-10.048</td>
<td>1.205</td>
<td>0.713</td>
<td>82.06</td>
<td>0.47</td>
<td>(0.493)</td>
</tr>
<tr>
<td></td>
<td>(4.371)</td>
<td>(0.099)</td>
<td>(0.000)</td>
<td>(0.493)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-12.355</td>
<td>1.438</td>
<td>-0.185</td>
<td>0.725</td>
<td>7.43</td>
<td>1.81</td>
</tr>
<tr>
<td></td>
<td>(5.118)</td>
<td>(0.228)</td>
<td>(0.159)</td>
<td>(0.024)</td>
<td>(0.178)</td>
<td></td>
</tr>
<tr>
<td>Interception $\ln VXN_{m-1}$</td>
<td>$\ln VOL_{m-1}$</td>
<td>Adj. $R^2$</td>
<td>Chi-square (p-value)</td>
<td>B-G (p-value)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasdaq 100 $\ln VOL_m$</td>
<td>-0.751</td>
<td>1.185</td>
<td>0.725</td>
<td>68.18</td>
<td>2.10</td>
<td>(0.148)</td>
</tr>
<tr>
<td></td>
<td>(0.336)</td>
<td>(0.091)</td>
<td>(0.000)</td>
<td>(0.148)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.897</td>
<td>1.356</td>
<td>-0.134</td>
<td>0.724</td>
<td>6.52</td>
<td>5.14</td>
</tr>
<tr>
<td></td>
<td>(0.390)</td>
<td>(0.219)</td>
<td>(0.155)</td>
<td>(0.038)</td>
<td>(0.023)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: S&P 100 index realized vs. implied volatility
Figure 2: Nasdaq 100 index realized vs. implied volatility
Figure 3: S&P 100 index (1988-94)

\[ \alpha = 0, \beta = 1 \]

OLS
Figure 4: S&P 100 index (1995-2002)

\[ \alpha = 0, \beta = 1 \]

OLS
Figure 5: Nasdaq 100 index (1995-2002)

OLS

$\alpha = 0, \beta = 1$