STOCK RETURN DYNAMICS, OPTION VOLUME, AND THE INFORMATION CONTENT OF IMPLIED VOLATILITY

STEWART MAYHEW
CHRIS STIVERS*

This article reports new empirical results on the information content of implied volatility, with respect to modeling and forecasting the volatility of individual firm returns. The 50 firms with the highest option volume on the Chicago Board Options Exchange between 1988 and 1995 are examined.

We would like to thank Bill Lastrapes, Mark Rubinstein, Scott Mixon, and seminar participants at the University of Alabama, the University of Georgia, the Federal Reserve Bank of Atlanta, Arizona State University, and Brigham Young University. Chris Stivers acknowledges financial support from a Sanford-Terry Research Grant from the University of Georgia. Stivers is also a visiting scholar at the Federal Reserve Bank of Atlanta. The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Finally, the U.S. Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This study expresses the authors' views and does not necessarily reflect those of the Commission, the Commissioners, or other members of the SEC staff.
*Correspondence author, Terry College of Business, University of Georgia, Brooks Hall, Athens, GA; e-mail: cstivers@terry.uga.edu

Received May 2002; Accepted October 2002

- *Stewart Mayhew is in the Banking and Finance Department at Terry College of Business at the University of Georgia in Athens, GA, and in the Office of Economic Analysis at the U.S. Securities and Exchange Commission in Washington, DC.*
- *Chris Stivers is an assistant professor in the Banking and Finance Department at Terry College of Business at the University of Georgia in Athens, GA.*

The Journal of Futures Markets, Vol. 23, No. 7, 615–646 (2003)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fut.10084

First, the results indicate that the ability of implied volatility to subsume all relevant information about conditional variance depends on option trading volume. For the most active options in the sample, implied volatility reliably outperforms GARCH and subsumes all information in return shocks beyond the first lag. For these active options, implied volatility performs substantially better than indicated by the prior results of Lamoureux and Lastrapes (1993), despite significant methodological improvements in the time-series volatility models in this study including the use of high-frequency intraday return shocks. For the lower option-volume firms in the sample, the performance of implied volatility deteriorates relative to time-series volatility models. Finally, compared to a time-series approach, the implied volatility of equity index options provides reliable incremental information about future firm-level volatility. Jrl Fut Mark 23:615–646, 2003

INTRODUCTION

Over the past few decades, there has been a vast quantity of research dedicated to modeling the joint dynamics of stock returns and volatility. It is now widely recognized that volatility varies substantially over time, and that volatility changes are persistent and somewhat predictable (Bollerslev et al., 1992; Harvey & Whaley, 1992; Schwert, 1989). It is also widely accepted that option prices incorporate forward-looking information that helps forecast future volatility (Christensen & Prabhala, 1998; Day & Lewis, 1992; Fleming, 1998; Lamoureux & Lastrapes, 1993).

An important question is whether option prices subsume all relevant information about future volatility, or whether additional information can be gleaned by applying time-series models to historical data. There is now a sizable literature addressing this topic in the context of market-level volatility and S&P 500 or S&P 100 Index options. Some of the first authors to look at index options, including Day and Lewis (1992) and Canina and Figlewski (1993), suggested that historical data contains important information that is not incorporated into index option prices. However, most subsequent studies have concluded that the implied volatility from index options captures most, if not all, of the relevant information in the historical data (Blair et al., 2001; Christensen & Prabhala, 1998; Fleming, 1998; and the survey article by Poon & Granger, 2001).

In contrast, individual stock options have received scant attention. A few early studies examined individual stock options (e.g., Chiras & Manaster, 1978; Latané & Rendleman, 1976). These early results tended to favor implied volatility, but the results are less than compelling because these articles predate the development of conditional heteroskedasticity models and employed naive models of historical volatility. Since then, only one major article examines the forecasting power of implied volatility for individual stock options in a modern, time-series framework—that of Lamoureux and Lastrapes (1993), whose analysis is based on the options of 10 stocks from April 1982 to March 1984. In a "horse race" comparison, these authors found that GARCH provided a better model of conditional variance than implied volatility for all 10 of their stocks. Moreover, they found that when implied volatility was added as a state variable in the GARCH conditional volatility equation, historical return shocks still provided important additional information beyond that reflected in option prices.

Further research in this area seems warranted, for several reasons. First, option markets have matured and expanded considerably since 1984, when the Lamoureux and Lastrapes (1993) sample ends. Option trading volume has increased, modeling and trading technologies have become more sophisticated, and, according to Rubinstein (1994), there appears to have been a fundamental structural change in option markets after the crash in October 1987. Second, the Lamoureux and Lastrapes (1993) methodology is now somewhat obsolete. Their methodology does not account for asymmetric GARCH effects (Glosten et al., 1993) or the additional information in intraday realized volatility (Andersen & Bollerslev, 1998), and it inappropriately allows implied volatility to feed back into the lagged conditional volatility term in the GARCH equation (Amin & Ng, 1997). Third, there are some interesting questions that can only be addressed by looking at a cross-section of individual stock options. For example, by looking at a sample of stocks with cross-sectional variation in option volume, one can examine whether high option trading volume is a necessary prerequisite for implied volatility to incorporate all relevant information. In addition, one can use equity index options to examine whether a firm's implied volatility captures all information about systematic risk. If index implied volatility helps explain residual volatility after accounting for own implied volatility, this would indicate an informational inefficiency in firm option prices. Finally, as argued by Campbell et al. (2001), modeling and understanding firm-level volatility may have practical applications in the areas of risk management, diversification, arbitrage, event studies, and option pricing. For these reasons, an updated study of individual firm options seems long overdue.

All of these issues are addressed in this article. The relation between stock returns, conditional variance, and implied volatility are examined for 50 actively-traded individual stock options listed on the Chicago Board Option Exchange over the 1988–1995 period. This study contributes to the existing literature in several dimensions.

First, the analysis of Lamoureux and Lastrapes (1993) is updated and expanded. Compared to their study, this study examines five times as many firms, uses more recent data (postdating the crash of 1987), and examines the robustness of the conclusions over a wider variety of empirical specifications. Compared to their study, this article employs a considerably longer time series—8 years instead of 2. This is important, as 2 years of daily data are probably insufficient to obtain accurate GARCH estimates (see Engle & Mezrich, 1995). In contrast to the results of Lamoureux and Lastrapes, this article's results are considerably more favorable toward implied volatility. In a horse race, this article finds that implied volatility outperforms GARCH specifications. Further, when implied volatility is included in the conditional variance equation, it captures most or all of the relevant information in past return shocks, at least for stocks with actively-traded options. Finally, this article finds that return shocks from period $t - 2$ and older provide reliable incremental volatility information for only a few firms in the sample.

Second, this study explores cross-sectional variation in the relative forecasting power of implied volatility and time-series models. By evaluating a cross-section of firms with varying levels of option volume, one can examine whether the quality of implied-volatility information depends on option trading volume. Indeed, the results indicate that even within the sample of 50 stocks with the highest option volume, implied volatility performs significantly better for those firms having higher trading volume. Outside of the 10 highest option-volume firms, past return shocks generally provide incremental information about future volatility. Thus, the evidence in this study suggests that a highly liquid option market is necessary for implied volatility to incorporate all relevant information about volatility.

Third, Andersen and Bollerslev (1998), Andersen et al. (2001), and others have demonstrated that the performance of time-series volatility models can improve significantly with the inclusion of information impounded into the variance of high frequency (intraday) returns. Following this literature, this article focuses on daily returns, but explicitly incorporates a measure of intraday realized volatility. The results indicate that a volatility metric formed from 5-minute returns does provide useful information about the conditional volatility of daily returns. However, for firms with high option volume, the implied volatility still reliably dominates the information from past return shocks.

Fourth, this article evaluates other methodological improvements beyond Lamoureux and Lastrapes. For example, the well-documented asymmetric response of conditional volatility to positive and negative shocks is accounted for, using the approach of Glosten et al. (1993). Moreover, following Amin and Ng (1997), the specification for conditional volatility does not permit lagged implied volatility to bleed into the lagged conditional variance term in the GARCH equation. These methodological improvements enhance the ability of the time-series model to capture effects that are not captured by implied volatility.

Finally, this study finds that the implied volatility from equity index options provides incremental information about firm-level conditional volatility. For nearly all firms, index implied volatility contains information beyond that in past return shocks. This suggests an alternative method for modeling volatility for stocks without traded options. For a few firms with less actively-traded individual options, the index implied volatility provides incremental information beyond the own firm's implied volatility. Thus, equity index options appear to impound systematic volatility information that is not available from less liquid stock options.

In the following section of this article, this study is placed in the context of the existing literature. Subsequent sections describe the data and variable construction, present the methodology, and then report the empirical results. The final section provides conclusions.

BACKGROUND ON TIME-VARYING VOLATILITY

Time-varying volatility is ubiquitous in asset return data. In an effort to parameterize the episodic nature of volatility, many dynamic models of asset returns have been proposed. However, there appears to be little consensus as to which model or family of models best describes returns. More fundamentally, financial economists do not fully understand what underlying economic forces drive the pervasive phenomena of volatility "clustering" or volatility "persistence."

One view is that volatility episodes occur because news events naturally cluster in time, or because it takes time for the market to process news. This is one simple interpretation of the volatility process parameterized by discrete-time conditional heteroskedasticity models, such as the ARCH model of Engle (1982), the GARCH model of Bollerslev (1986), and many other permutations.[1] The high volatility observed at time t is presumed to be associated with a news event, that is likely to be reinterpreted or followed by more news events at time $t + 1$ and later.

[1]See Bollerslev et al. (1992) for a survey.

Another view is that changes in volatility are associated with changes in the economic, political, or regulatory environment, industry factors, or the firm's life cycle. This view is consistent with the "regime-switching" family of models. Both GARCH and regime-switching approaches are consistent with observed periods of high volatility and periods of low volatility. The main difference is that GARCH models impose more structure on the day-to-day volatility changes, with recent return shocks playing a driving role in determining conditional volatility. Regime switching models, in contrast, assume that volatility is generally constant from day to day, but periodically jumps to a new level. Lamoureux and Lastrapes (1990) have demonstrated that there appear to be structural changes in GARCH parameters across subperiods.

Models that incorporate both regime-switching and ARCH/GARCH effects have been developed by Cai (1994), Hamilton and Susmel (1994), Gray (1996), and others. These models are fairly flexible in their ability to fit the data, but there are practical problems associated with implementing them. There is little theory to guide one in selecting from among many equally plausible empirical specifications. For any particular specification, a large number of parameters must be estimated from historical data. For example, in Gray's (1996) article, moving from standard GARCH to a two-regime switching GARCH model increases the number of parameters from 5 to 12, and it is by no means clear that two regimes are sufficient to capture year-to-year variation in unconditional variance. Using implied volatility as a state variable in the conditional volatility equation is an alternative, more parsimonious approach to capturing the effects of changing regimes.

DATA AND VARIABLE CONSTRUCTION

This article analyzes the dynamics of 50 individual stocks with options traded on the Chicago Board Options Exchange (CBOE). Among those stocks with options listed on the CBOE over the entire 1988–1995 period, the firms having the highest option trading volume are selected for the sample. Daily dividend-adjusted returns on the large-stock portfolio and the 50 stocks were obtained from CRSP for the period January 4, 1988 through December 31, 1995.

For implied volatility at the market level, this study uses the daily time series of the Volatility Index (VIX) from the Chicago Board Options Exchange. This index represents the implied volatility of an at-the-money option on the S&P 100 Index with 22 trading days to expiration. It is constructed by taking a weighted average of the implied volatilities for eight options, including a call and a put at the two strike prices closest to the money and the nearest two expirations (excluding options within 1 week of expiration). Each of the eight component implied volatilities is calculated from the bid-ask quote midpoint using a binomial tree that accounts for early exercise and dividends. This procedure was designed to reduce noise and mitigate measurement errors. The procedure is described in detail by Whaley (1993).

For the implied volatility of the individual stock options, an implied volatility measure is constructed using the same procedure as used for the VIX. This procedure uses option price data from the Berkeley Options Data Base, realized dividends from CRSP, and T-bill rates from Datastream.[2] For each stock on every trading day, the procedure uses eight designated implied volatilities that are calculated using midpoints of the final quote of the day, matched with a contemporaneous observation of the underlying stock price. Each of the eight implied volatilities is calculated using a binomial tree that explicitly accounts for early exercise and discrete dividends. The eight estimates are then aggregated using the VIX weighting procedure. On certain days, prices for one or more of the eight options were unavailable, or else a reported option price was below the lower arbitrage bound, in which case implied volatility does not exist. On these days, the previous day's implied volatility estimate is used. Using stale implied volatility data will create a small bias against finding an important role for implied volatility in the subsequent volatility modeling.

Table I lists the sample of 50 firms, along with volatility and option volume statistics for each firm. The sample includes the 50 firms that had the highest total option trading volume over the period 1988–1995 and that met the following additional screens. First, the firm had to have the complete set of daily returns from CRSP over the 1988–1995 period and have the same ticker over this period. Second, each firm had to have valid implied volatility observations in at least 1,800 of the 2,022 daily periods (approximately 90%). These additional screens removed only 8 of the top 58 option-volume firms. Table I also reports how many daily observations use a stale IV in place of the current IV for each firm (as discussed in the preceding paragraph). Note that the median (mean) proportion of the observations with a current IV are 92.5% (92.5%) for the top 10 option-volume firms versus 92.1% (91.7%) for the lowest option-volume firms. This tiny difference in the proportion of stale IVs

[2]The choice of which dividend and interest rate inputs to use does not have a significant impact on the implied volatility estimates, as the VIX averaging procedure eliminates most of the bias that would be introduced by using an interest rate or dividend with error.

TABLE I

Descriptive Statistics

Firm	*St. Dev. of Return (Annual %)*	*Mean Impl. Vol. (Annual %)*	*St. Dev. of Impl. Vol. (Annual %)*	*% Observ. with Impl. Vol.*	*Total Volume 1988–1995*	*Median Volume 1988–1995*
IBM	24.57	23.73	5.31	95.4	37008649	17093.0
Merck	23.58	24.00	4.21	92.0	8278922	3154.0
General Motors	28.63	26.91	4.87	93.7	8206743	3337.0
Chrysler	37.05	36.18	9.87	95.6	6975108	2509.5
General Electric	20.65	20.61	4.31	90.3	6534711	2560.5
AT&T	21.80	21.28	4.55	89.1	5575220	2080.5
Ford	27.81	27.26	4.80	92.2	5233991	1997.5
Eastman Kodak	25.86	24.12	4.30	90.8	5000864	1745.0
Wal-Mart	27.04	27.97	6.07	92.9	4758688	1663.0
Hewlett-Packard	33.68	30.52	5.27	92.7	4596612	1797.5
Boeing	25.15	25.04	4.57	90.5	3955964	1415.0
Bristol Meyers	20.58	20.07	3.79	89.7	3609440	1417.0
Texas Instruments	37.78	36.36	6.90	90.3	3454342	1030.0
Coca-Cola	22.81	23.33	4.26	92.8	3347979	1298.5
Johnson&Johnson	23.67	22.70	4.05	90.7	3220516	1257.0
Pepsico	25.83	24.23	5.02	89.8	3035505	1128.0
Sears Roebuck	26.68	26.00	4.78	91.7	2877494	943.5
Bank America	31.32	30.96	8.44	93.4	2847173	1025.0
Avon	29.03	27.77	9.58	90.8	2386215	328.0
Dow Chemical	23.97	22.94	4.22	91.7	2339912	825.5
Polaroid	32.91	31.89	6.55	92.6	2292825	211.0
Homestake Mining	39.36	39.02	7.09	98.7	2285327	827.5
K Mart	31.45	30.18	10.13	91.8	2001961	600.0
Occidental Petroleum	25.97	25.29	6.02	89.7	1732599	512.5
Limited	38.14	36.28	6.55	96.1	1699887	563.0
McDonald's	23.66	24.46	4.56	91.5	1647557	576.0
Gap	38.46	37.27	6.33	92.8	1536138	461.0
Xerox	23.44	23.95	4.58	91.9	1486025	434.0
Computer Associates	48.22	44.83	11.79	93.9	1355592	430.0
Delta Airlines	27.76	28.97	4.64	90.9	1349471	472.5
Toys R Us	30.38	31.38	5.60	91.5	1291677	449.0
3M	19.22	19.34	3.40	91.3	1262817	441.0
Baxter International	26.81	27.32	4.22	93.2	1227256	322.5
Heinz	23.09	24.82	4.47	92.2	1223315	312.5
International Paper	23.47	24.29	3.72	92.7	1192588	420.0
Mobil	19.10	18.28	3.42	90.6	1190552	402.0
Battle Mountain Gold	45.08	46.47	12.27	94.9	1168160	369.0
ARCO	18.25	17.93	3.13	90.5	1096445	385.0
Schlumberger	25.82	24.02	4.20	92.0	1060152	343.0
Bethlehem Steel	39.36	38.43	6.12	98.1	991581	298.0
Alcoa	26.01	25.70	3.61	92.2	944056	341.0
Honeywell	23.98	27.78	3.90	92.1	913414	211.0
Amoco	20.18	18.74	3.32	91.3	896638	281.5
ITT	20.92	21.64	3.48	91.1	847406	237.0
First Interstate	35.84	32.48	12.32	90.0	838402	237.0
Fluor	30.69	29.32	6.68	92.5	707369	211.0
Halliburton	31.60	29.93	4.93	89.1	697708	206.0
Black & Decker	36.55	35.04	9.64	97.2	652372	168.0
Mead	27.54	28.47	6.85	92.1	639748	77.0
Bell Atlantic	20.77	18.89	3.66	89.7	634737	171.0

for the high option-volume firms versus the low option-volume firms seems unlikely to be the primary factor behind the option-volume-related results that are later reported.

This study also uses a realized volatility metric that is constructed from high-frequency intraday returns. This is accomplished by computing, for each day, the sum of squared 5-minute returns, where returns are calculated from midpoints of bid-ask spreads quoted on the primary exchange. Intraday quotes were obtained from the New York Stock Exchange TAQ database.

METHODOLOGY

This section describes the empirical specifications used in this article's analysis. A progression of different empirical specifications are examined that are chosen to address the empirical questions outlined in the introduction. The results for each specification are reported in following sections.

First, the following model for stock return dynamics is estimated for all 50 firms over the entire 1988 to 1995 sample period.

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \tag{1}$$

$$\varepsilon_t \sim N(0, h_t) \tag{2}$$

$$h_t = \frac{\beta_0 + (\beta_1 + \beta_2 D_{t-1}^{-})\varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3\, IV_{t-1}^2 \tag{3}$$

where R_t is the simple return from time $t - 1$ to t with dividends, ε_t is the return residual, IV_t^2 is the daily implied variance from options on the respective stock at the end of day t, D_{t-1}^{-} is a dummy variable that takes on a value of 1 if ε_{t-1} is negative, L indicates the lag operator, and the αs, βs, and γs are parameters. $N(\ )$ designates the normal distribution. Because daily stock returns are essentially unpredictable, the volatility results should not be sensitive to variation in the specification of (1) (see the arguments and approach in Fleming et al., 2001). Nevertheless, later specifications do consider different conditional mean equations. The system (1)–(3) is estimated simultaneously by maximum likelihood estimation using the conditional normal density.

The question of interest is how best to model h_t, the conditional variance of the stock return. This article's empirical investigation evaluates five variations of the above conditional variance specification to compare the information in lagged return shocks and implied volatility.

Note that, by restricting the appropriate coefficients to be zero, this specification nests the widely used GARCH(1,1) and asymmetric GJR-GARCH(1,1) model. This specification allows for the asymmetric volatility phenomenon and does not force implied volatility at time $t - 2$ and older to feed back into the conditional variance equation.[3]

For inference, two alternate methods are examined. First, the statistical significance of the estimated coefficients are examined using a Wald test. p-Values for each estimated coefficient are calculated using standard errors that are robust to departures from conditional normality, in accordance with Bollerslev and Wooldridge (1992). Second, the empirical investigation compares the importance of different explanatory terms by evaluating restrictions on the full model. Likelihood ratio tests (LRT) are performed that compare the unrestricted and restricted log-likelihood function values from the maximum likelihood estimation on each variation of the model. The LRT yields a chi-squared (X) test statistic, where X is the number of restrictions. The null hypothesis is that the additional terms in the unrestricted model do not provide reliable incremental information beyond the restricted model.

Out-of-sample results are also examined. Following Noh et al. (1994), an out-of-sample procedure is conducted which repeats the primary analysis shown later with parameters estimated from a rolling window of 1,000 past observations. This procedure yields 1,022 daily volatility forecasts for comparison across the models.

Because the implied volatility horizon in this study is 22 trading days, the implied volatility will display substantial persistence from day to day. Another way to examine the volatility information from implied volatility that avoids this persistence issue is to analyze how much incremental volatility information is contained in the daily change of implied volatility. To evaluate this issue, three variations of the following conditional variance equation are estimated, with the conditional mean as given in (1):

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_3 IV_{t-2}^2 + \beta_4 \Delta_1 IV_{t-1}^2 \tag{4}$$

where $\Delta_1 IV_t^2$ is the 1-day change in the respective firm's daily implied variance from $t - 1$ to t and the other terms are as defined for (3). The

[3]In contrast, Lamoureux and Lastrapes (1993) examine a symmetric GARCH(1,1) model. In their model that combines both the implied volatility and past return shocks, their specification forces the same lagged decay structure on both the $t - 2$ and older return shocks and the $t - 2$ and older implied volatilities.

choice of this specification relies on earlier findings that lagged return shocks older than $t - 1$ add almost no volatility information.

A second change-in-IV model is estimated to further evaluate whether the $t - 2$ return shock provides incremental volatility information. The following conditional variance equation is estimated for each firm:

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \varepsilon_{t-2}^2 + \beta_3 IV_{t-3}^2 + \beta_4 \Delta_2 IV_{t-1}^2 \tag{5}$$

where $\Delta_2 IV_t^2$ is the 2-day change in the respective firm's daily implied variance from $t - 2$ to t and the other terms are as defined for (4). For this model, the coefficient of interest is β_2.

Next, to investigate whether the implied volatility of index options is useful in predicting the volatility of individual stocks, four variations of the following system are estimated:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \tag{6}$$

$$h_t = \frac{\beta_0 + \beta_1 \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{t-1}^2 + \beta_5 VIX_{t-1}^2 \tag{7}$$

where VIX_t^2 is the daily implied variance from OEX index options, the α, β, and γ terms are parameters, and all other terms are defined as before.

Comparing the values of β_3 and β_5 can help assess the efficiency of the firm's own implied volatility as a volatility forecast. If the firm's own implied volatility fully reflects all information, one would expect β_5 to be insignificant. This model is also estimated with the restriction $\beta_3 = 0$, to test whether *VIX* on its own is useful for predicting firm-level volatility. If so, this technique could be used to help predict volatility for stocks with no traded options.

Next, the empirical investigation examines whether the implied volatility might have significant explanatory power in the mean equation. The motivation has two dimensions. First, for robustness, adding the lagged implied volatility to the mean equation ensures that effects in the mean equation do not change the results for the volatility equation. Second, and perhaps more interestingly, adding implied volatility to the mean equation allows one to check for a positive risk-return relation between the conditional mean and the level of implied volatility. There is an ongoing debate in the literature about whether a higher conditional volatility is associated with variation in the mean equation. See Scruggs (1998) for recent evidence and a summary of prior results.

To test whether a firm's implied volatility helps explain its conditional mean return, the following specification is estimated:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \alpha_2 RM_{t-1} + \alpha_3 IV_{t-1} + \varepsilon_t \tag{8}$$

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_3 IV_{t-1} + \beta_5 VIX_{t-1} \tag{9}$$

where RM_t is the stock market return from the S&P 500 Index and all other variables are defined as before. Here, the coefficient of interest is α_3. The results are presented later.

Finally, in regard to modeling future daily volatility, the information from the prior day's sum of squared 5-minute returns is also examined. The information from the 5-minute returns is compared to the volatility information from lagged daily returns and lagged implied volatility. Due to data constraints, this aspect of the empirical investigation only examines data from 1993 through 1995, $n = 757$ trading days.

For this investigation, the following model on daily stock returns is estimated:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \tag{10}$$

$$h_t = \frac{\beta_0 + \beta_1 \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{t-1}^2 + \frac{\beta_6 V5_{t-1}}{1 - \gamma_2 L} \tag{11}$$

where $V5_t$ is the firm's intraday sum of squared 5-minute returns for day t, and all other terms are defined as before.

In addition to the features discussed for the system in (1) and (3), the specification of (10) and (11) allows for the lagged daily return shocks and the lagged 5-minute sum of squared returns to have a different lagged decay structure (through the γ_1 and γ_2 coefficients). Thus, in addition to this study's main interest in implied volatility, this investigation also enables one to compare volatility information from the lagged return shocks at the daily horizon versus the 5-minute horizon. This specification does not model the GJR asymmetry because results reported in the next section indicate that the GJR asymmetry is relatively unimportant in this sample and for parsimony.

EMPIRICAL RESULTS

Standard GARCH Versus Implied Volatility

This section reports the results from estimating the empirical specifications presented in the previous section. The tables report the mean and median of the estimated coefficients across the 50 stocks. The number of coefficients that are positive (negative) and significant at the 5% level are also reported. To illustrate how the results vary with the level of option trading volume across firms, the tables first present summary results in Panel (a) for the top volume-quintile of the firms. For contrast, Panel (b) in the tables presents the results for the bottom volume-quintile of firms. Finally, Panel (c) in the tables summarizes results for the entire sample of 50 stocks.

The tables do not report the estimated α_1 coefficients on the lagged return in the conditional mean Equation (1). This choice is made for parsimony and because the estimated α_1 values are near 0 for the sample. In the sample of 50 firms, the mean (median) α_1 is 0.016 (0.020). Only nine (four) firms have an estimated α_1 that is positive (negative) and statistically significant at the 5% level.

Table II reports parameter estimates for five variations of the conditional volatility model (3). Within each panel of Table II, the first line reports estimates of the GARCH(1,1) model. As expected, the estimated β_1 coefficients are positive for all 50 stocks and positive and significant for 37 of the stocks. Similarly, the estimated γ_1 coefficients are positive and significant for 44 of the 50 individual stocks. Thus, the standard GARCH effects are very evident in the data.

The third line in each segment represents the model with only the implied volatility at $t - 1$ and no lagged return-shock terms. In this specification, the estimated coefficient β_3 on the *IV* term is positive and highly significant for all 50 stocks. Comparing likelihood values, the "implied volatility only" model outperforms GARCH(1,1) for 42 of the 50 firms. This contrasts with the results reported in Table I in Lamoureux and Lastrapes (1993). Using data from April 1982 through March 1984, they find that GARCH(1,1) outperforms implied volatility for all 10 of the stocks they analyzed.[4] Further, the implied volatility model in line three of Table II does substantially better than the GARCH(1,1) model for all 10 of the top volume-quintile firms. On the other hand, five of the eight stocks where GARCH(1,1) performs better than the "IV only" model are among the 11 firms with the lowest option-trading volume. Thus, option trading volume appears informative about the quality of information in a firm's IV.

The fourth line in each segment of Table II shows the results when both GARCH and implied volatility terms are included. For the top volume quintile of firms, the estimated β_1 on the lagged return shock is reliably positive for only 5 of the 10 firms. Further, the estimated γ_1 coefficients decrease dramatically from a mean of 0.885 in the line-one model to 0.016 in the line-four model.

[4]The only firm that overlaps with this article's sample and Lamoureux and Lastrapes' (1993) sample is Toys R Us. For the 1988–1995 sample period with Toys R Us, the "IV only" model outperforms both a standard GARCH(1,1) and the GJR-GARCH model.

In contrast, the standard GARCH terms are much more important for the bottom volume-quintile of firms. For these 10 firms, the estimated γ_1 coefficients only decrease from a mean of 0.801 in the line-one model to a mean of 0.434 in the line-four model. The γ_1 coefficients remain positive and statistically significant for 5 of the 10 firms.

Further analysis is conducted using a likelihood ratio test (LRT). The incremental information in the GARCH terms is evaluated by comparing the unrestricted model in line four versus the restricted model in line three. A p-value of 1% is chosen to reject the null in all of the LRT

TABLE II

Conditional Volatility of Individual Stock Returns and Own-Firm Implied Volatility

Model		$\beta_0 \times 10^5$	β_1	β_2	β_3	γ_1	$\mathcal{L}$
(a) Top Quintile of Firms by Option Volume							
1. GARCH	Mean	2.09	0.050			0.885	7303.6
	Median	0.745	0.050			0.919	
	++/−−	4/0	9/0			10/0	
2. GJR-GARCH	Mean	1.08	0.024	0.037		0.916	7309.5
	Median	4.69	0.013	0.027		0.954	
	++/−−	0/0	2/0	5/0		10/0	
3. IV	Mean	3.70			0.999		7337.6
	Median	1.25			0.982		
	++/−−	1/1			10/0		
4. GARCH+IV	Mean	1.75	0.055		0.881	0.016	7338.3
	Median	0.79	0.052		0.948	−0.047	
	++/−−	1/0	5/0		10/0	3/2	
5. GJR-GARCH+IV	Mean	−0.94	0.046	0.027	0.963	−0.123	7340.5
	Median	0.16	0.050	0.009	0.950	−0.196	
	++/−−	0/1	2/0	0/0	9/0	2/1	
(b) Bottom Quintile of Firms by Option Volume							
1. GARCH	Mean	4.40	0.064			0.801	7310.6
	Median	1.22	0.064			0.916	
	++/−−	2/0	6/0			9/0	
2. GJR-GARCH	Mean	1.63	0.042	0.021		0.886	7309.8
	Median	1.12	0.030	0.023		0.911	
	++/−−	4/0	4/0	2/0		10/0	
3. IV	Mean	5.20			0.812		7322.9
	Median	4.45			0.786		
	++/−−	4/0			10/0		
4. GARCH+IV	Mean	2.90	0.082		0.623	0.434	7334.6
	Median	1.02	0.081		0.599	0.409	
	++/−−	1/0	6/0		8/0	5/0	
5. GJR-GARCH+IV	Mean	2.77	0.074	0.005	0.600	0.476	7335.4
	Median	0.98	0.076	0.007	0.597	0.444	
	++/−−	1/0	2/0	0/0	9/0	5/0	

TABLE II (*Continued*)

Model		$\beta_0 \times 10^5$	β_1	β_2	β_3	γ_1	$\mathcal{L}$
(c) Full Sample							
1. GARCH	Mean	3.21	0.055			0.856	7253.7
	Median	0.78	0.044			0.925	
	++/−−	11/0	37/0			44/0	
2. GJR-GARCH	Mean	1.71	0.038	0.027		0.892	7252.3
	Median	0.71	0.028	0.022		0.935	
	++/−−	10/0	16/0	10/0		48/0	
3. IV	Mean	3.24			0.897		7275.6
	Median	2.42			0.914		
	++/−−	7/3			50/0		
4. GARCH+IV	Mean	2.56	0.073		0.778	0.129	7282.1
	Median	1.27	0.076		0.798	0.119	
	++/−−	3/1	28/0		48/0	11/5	
5. GJR-GARCH+IV	Mean	1.91	0.066	0.013	0.777	0.164	7283.8
	Median	1.17	0.064	0.007	0.783	0.083	
	++/−−	1/3	18/0	5/0	48/0	15/3	

The table reports mean and median parameter estimates across 50 stocks for the following model of daily returns:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \qquad h_t = \frac{\beta_0 + (\beta_1 + \beta_2 D_{t-1}^{-})\varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{t-1}^2$$

Note. R_t is the firm's simple discrete return (with dividends), ε_t is the return residual, h_t is the conditional variance of ε_t, $D_{t-1}^{-} = 1$ if ε_{t-1} is negative and equals 0 otherwise, IV_t^2 is the daily implied variance from the options of the respective stock, L indicates the lag operator, and the αs, βs, and γ are estimated coefficients. The sample period is January 1988 through December 1995. The system is estimated simultaneously by maximum likelihood estimation using the conditional normal density. The table also reports the number of coefficients that are positive and significant (++) and the number that are negative and significant (−−) at the 5% level, calculated with quasi-maximum likelihood standard errors in accordance with Bollerslev and Wooldridge (1992). Panels (a) and (b) present the results for the top and bottom quintiles of the firms, sorted by option trading volume. Panel (c) presents the results for all 50 firms.

testing. None of the top five option-volume firms reject a null that the GARCH terms provide no incremental information beyond the IV term. In contrast, the four lowest option-volume firms all reject this null. Overall, 33 of the 50 firms accept the null that the traditional GARCH terms add no incremental information beyond the "IV only" model in line three.

Next, a similar LRT is conducted that tests the incremental information in time $t - 2$ and older return shocks. For this LRT, a variation of the model that restricts the γ_1 and β_2 to be 0 is compared to the model in line four. The results of this LRT indicate that including the information from $t - 2$ and older lagged return shocks is reliably important for only 3 of the 50 firms and none of the top quintile of option-volume firms.

Next, a LRT is conducted that tests the incremental information in IV beyond traditional GARCH by comparing the unrestricted model in line four versus the restricted model in line one. This LRT indicates that IV provides reliable incremental information beyond the GARCH model for 43 of the 50 firms and for all 10 of the top quintile of option-volume firms. Thus, the results from both the Wald test and the LRT indicate that there is little reason to include lagged return shocks beyond lag-one. Further, for the highest option-volume firms, implied volatility subsumes nearly all of the information in past lagged return shocks.

Several observations above suggest that the quality of the information in IV is positively related to option trading volume. Although the quality of IV information does not fall monotonically with a firm's option trading volume, there does appear to be an overall positive relation between the quality of IV information and option trading volume. Here, several additional observations are made to support this contention.

First, Figure 1 displays the estimated coefficients from the GARCH+IV model in Table II for all 50 firms. For this model, the conditional variance may place weight on an estimated constant, the lagged return shocks, and IV^2_{t-1}. Panel A reports the estimated β_3s on IV^2_{t-1} across the 50 firms, in the order of high-to-low average option-trading volume. This figure depicts a noticeable decrease in β_3 as the option-trading volume decreases. For example, for the top 20 option-volume firms, 16 of the estimated β_3s are greater than 0.7 and the average β_3 is 0.885. However, for the bottom 20 option-volume firms, only six of the estimated β_3s are greater than 0.7 and the average β_3 is only 0.645. This difference-in-means is statistically significant at a 1% p-value. Thus, these results indicate a lower weight placed on IV^2_{t-1} for the lower option-volume firms.

Figure 1 (B) reports the weight that each firm's conditional variance places on the lagged return shocks for the GARCH+IV model. This panel reports the sum of the coefficients on the first five lagged return shocks, based on each firm's estimated β_1 and γ_1 coefficients. This panel reflects an increase in the weight on the lagged return shocks for the lower option-volume firms. For example, for the highest 20 option-volume firms, the mean of the summed coefficients is 0.084 and only 5 of the 20 firms have a value for the summed coefficients that exceeds 0.14. In contrast, for the lowest 20 option-volume firms, the mean of the summed coefficients is 0.137 and 10 of the 20 firms have a value for the summed coefficients that exceeds 0.14. This difference-in-means is statistically significant at a 1% p-value.

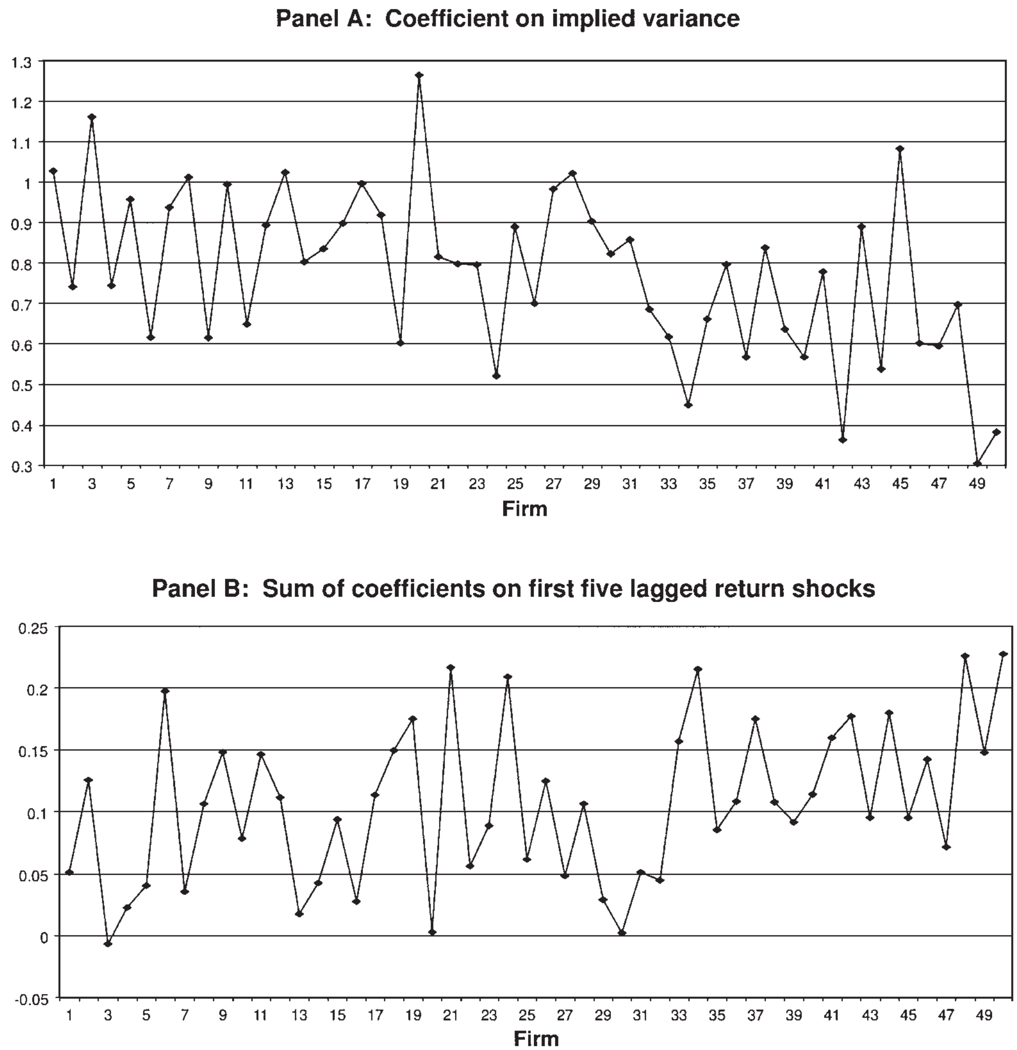

FIGURE 1
This figure displays the results from estimating the GARCH+IV model in Table II across the sample of 50 stocks. For the conditional variance equation, Panel A reports the estimated β_3 coefficient on the lagged implied variance from option prices for each firm. Panel B reports the sum of the coefficients on the first five lagged return shocks, based on the estimated β_1 and γ_1 for each firm. The results are displayed from the highest to the lowest option-volume firm, left to right.

Second, the log-likelihood function value for the GARCH+IV model is compared to the standard GARCH model (analogous to the excess log-likelihood reported in Table I of Blair et al., 2001). For the top 15 option-volume firms, the median increase in the log-likelihood function value is 25.9 when adding IV to the model. In contrast, for the bottom 15 option-volume firms, the median increase in the log-likelihood function value is much smaller at only 9.4.

Asymmetric GARCH Versus Implied Volatility

This subsection reports on the asymmetric GARCH model of Glosten et al. (1993) (GJR). These results are reported in line two of each panel in Table II. In this sample, allowing for the GJR asymmetry is reliably important for only 10 of the 50 firms, as indicated by the estimated β_2 coefficients. This suggests that the comparison of the GJR model to the implied volatility models should be very similar to the comparison with standard GARCH, as reported in the previous section.

The results support this conjecture. The "implied volatility only" model in line three outperforms the GJR-GARCH model in line two for 37 of the 50 firms, including 18 of the top 20 option-volume firms. In contrast, for the bottom 11 option-volume firms, the GJR-GARCH model performs better than the "IV only" model for seven firms. LRT's that compare the GJR model in line two, the implied volatility model in line three, and the unrestricted model in line five are also conducted. The results are very similar to the numbers reported in the previous section for the standard GARCH comparison.

Overall, for the top option-volume quintile of firms, the results suggest that implied volatility largely subsumes volatility information from all but the first lagged return shock. Thus, for firms with very widely traded options, the evidence suggests that an ARCH(1) model in conjunction with implied volatility should provide a good, parsimonious model for a firm's conditional volatility.

The results also suggest the following practical implication. What may be true for index options and widely traded stock options such as IBM may not generalize to firms with substantially less trading volume. Recall that the sample is composed of the 50 most actively traded CBOE options over this period, so the bottom 10 firms in the sample are still very actively traded compared to the vast majority of listed stock options. However, even within the sample of highly traded options, the implied-volatility information from the lower ranked option-volume firms degrades appreciably as compared to the highest option-volume firms.

Out-of-Sample Volatility Forecasts

To supplement the in-sample results reported above, this section investigates the out-of-sample performance of the implied-volatility model versus the GARCH model. Two out-of-sample comparisons are performed that are similar to those reported by Lamoureux and Lastrapes (1993). In the interest of brevity, only summary results of the more salient findings are reported here.

Before proceeding, recall that the GARCH process with daily returns is designed to obtain a 1-day ahead conditional volatility. In contrast, the explanatory variable in this article's implied-volatility model is an implied volatility from option prices that is weighted to reflect a 1-month volatility horizon. Because the out-of-sample comparison in this section focuses exclusively on a 1-day ahead volatility horizon, the GARCH model may have an advantage because the realized volatility horizon that is evaluated matches the volatility horizon of the GARCH model. This choice is made deliberately because if the implied-volatility forecast beats the GARCH forecast at a 1-day horizon, then it seems unlikely that a GARCH-based model would beat the implied volatility at a longer volatility horizon.

First, out-of-sample volatility estimates are calculated for both the implied-volatility model and the GARCH model. The procedure is as follows. First, the first 4 years of the sample (1,000 days) are used to obtain initial parameter estimates for the GARCH model and the implied-volatility model. Then, these estimated parameters and the current explanatory variables are used to form a 1-day-ahead conditional volatility estimate for day 1,001.

Next, the parameter estimation is repeated using days 2 through 1,001 as the estimation period. The updated parameters are then used to calculate the conditional volatility estimate for day 1,002. This procedure is repeated for each subsequent day to end up with 1,021 daily out-of-sample volatility forecasts for each firm, for both the GARCH model and the implied-volatility model.

It is important to note that the volatility forecast from the implied-volatility model is not simply the daily volatility that is directly implied by the value of the implied volatility. Rather, the procedure corrects for bias in the implied volatility through the parameters estimated in the 1,000-day estimation period. Thus, as pointed out by Figlewski (1997), such a procedure does not directly evaluate whether the current value of implied volatility by itself provides the best forecast. Rather, the procedure evaluates whether information in implied volatility can be used to provide a better forecast than models that rely only on lagged return shocks.

Next, the GARCH forecasts are compared to the implied-volatility forecasts using two different methods. First, it is determined which forecast has the lowest average deviation between the squared return shocks and the volatility forecast. The average squared deviation (ASD) is calculated for the volatility forecast from each model:

$$ASD = \frac{1}{T}\sum_{t=1}^{T}[(R_{it} - \mu_i)^2 - \hat{h}_{it}]^2 \tag{12}$$

where R_{it} is the realized firm return on day t, μ_i is the average daily return of firm i over the evaluation period, T is the number of days in the evaluation period, and h_{it} is the out-of-sample conditional variance that was estimated from each respective model.

The results are as follows. First, for all of the top 10 option-volume firms, the implied-volatility forecast has a lower ASD than the GARCH forecast. Overall, for the 50 firm sample, the implied-volatility forecast outperforms the GARCH forecast for 40 of the 50 firms. Similar to earlier results, the implied volatility from the lower option-volume firms is not as informative. For the lowest 10 option-volume firms, the GARCH forecast outperforms the implied-volatility forecast for five firms.

The relatively strong performance of the implied-volatility forecast versus the GARCH forecast is in contrast to the results reported by Lamoureux and Lastrapes (1993). They find that their implied-volatility forecast performs the best for only 2 firms in their 10-firm sample. However, as noted above, this article's results are not directly comparable to Lamoureux and Lastrapes because the implied-volatility model in this study corrects for the bias in the raw implied volatility. Additionally, Lamoureux and Lastrapes evaluate a volatility forecast over a longer volatility horizon.

Next, a joint comparison of the volatility forecasts is conducted by regressing the squared return shocks against the volatility forecasts from both the implied-volatility model and the GARCH model simultaneously. This analysis is not a direct comparison of volatility forecasts, but rather evaluates which forecast is more reliably related to the time-variation in squared return shocks. The following regression model is estimated using the out-of-sample volatility forecasts, $n = 1{,}021$ daily periods:

$$(R_{it} - \mu_i)^2 = a_0 + a_1 \hat{h}_{it}^{IV} + a_2 \hat{h}_{it}^{G} + \varepsilon_{it} \tag{13}$$

where $\hat{h}_{it}^{IV}$ is the volatility forecast from the implied-volatility model, $\hat{h}_{it}^{G}$ is the volatility forecast from the GARCH model, and other variables are defined as before.

The parameter estimates from (13), available on request, indicate the following. First, for the top 10 option-volume firms, the estimated coefficient on the implied-volatility forecast (a_1) is positive and significant (5% p-value) for all but one firm. In contrast, the coefficient for the GARCH forecast (a_2) is not positive and significant for any of the top 10 firms. However, for 6 of the bottom 10 option-volume firms, the GARCH term has a stronger positive explanatory relation than the implied-volatility term. For the entire 50 firm sample, the average (median) t-statistic on the a_1 coefficient is 2.55 (2.40), calculated with autocorrelation and heteroskedastic consistent standard errors. For a_2, the average (median) t-statistic is 1.14 (1.34).

To summarize, the out-of-sample evaluation results are consistent with earlier in-sample results. For firms with high option volume, information from the implied volatility can be used to form a superior in-sample fit and out-of-sample forecast compared to conventional GARCH models. However, for the lower option-volume firms, the information from the implied volatility is less useful—the GARCH model performs comparably to the IV model for the lowest 10 option-volume firms.

Incremental Volatility Information from Daily Changes in Implied Volatility

Next, this section evaluates how much incremental volatility information is contained in daily IV changes. Table III reports the results from estimating four variations of the conditional variance equation given by (4). The primary coefficient of interest is the β_4 on the change in implied variance from $t - 2$ to $t - 1$.

First, the model variation in line one indicate that IV^2_{t-2} by itself is an important explanatory variable for the conditional variance of all 50 stocks. Second, the results for the line-two model indicate that the

TABLE III

Conditional Volatility of Individual Stock Returns and Changes in Implied Volatility

Model		$\beta_0 \times 10^5$	β_1	β_3	β_4	$\mathcal{L}$
(a) Top Quintile of Firms by Option Volume						
1. Base	Mean	0.92		0.962		7328.4
	Median	1.83		0.903		
	++/−−	2/1		10/0		
2. $\Delta_1 IV^2_{t-1}$	Mean	−0.49		1.021	0.838	7336.3
	Median	0.37		0.991	0.781	
	++/−−	0/1		10/0	8/0	
3. ε^2_{t-1}	Mean	0.417	0.062	0.921		7334.4
	Median	0.179	0.060	0.868		
	++/−−	2/1	6/0	10/0		
4. $\Delta_1 IV^2_{t-1} + \varepsilon^2_{t-1}$	Mean	−0.79	0.053	0.980	0.783	7340.9
	Median	0.29	0.048	0.960	0.771	
	++/−−	0/1	4/0	10/0	5/0	

(*Continued*)

TABLE III (*Continued*)

Model		$\beta_0 \times 10^5$	β_1	β_3	β_4	$\mathcal{L}$
(b) Bottom Quintile of Firms by Option Volume						
1. Base	Mean	6.38		0.789		7315.1
	Median	5.85		0.764		
	++/−−	5/0		10/0		
2. $\Delta_1 IV^2_{t-1}$	Mean	4.55		0.840	0.548	7320.9
	Median	4.17		0.850	0.488	
	++/−−	3/0		10/0	4/0	
3. ε^2_{t-1}	Mean	4.89	0.101	0.730		7329.7
	Median	5.07	0.093	0.727		
	++/−−	4/0	8/0	10/0		
4. $\Delta_1 IV^2_{t-1} + \varepsilon^2_{t-1}$	Mean	4.12	0.091	0.764	0.397	7331.4
	Median	3.95	0.091	0.771	0.388	
	++/−−	4/0	7/0	10/0	4/0	
(c) Full Sample						
1. Base	Mean	4.60		0.854		7263.9
	Median	3.98		0.856		
	++/−−	13/3		50/0		
2. $\Delta_1 IV^2_{t-1}$	Mean	2.53		0.914	0.769	7273.6
	Median	2.49		0.897	0.746	
	++/−−	6/4		50/0	37/0	
3. ε^2_{t-1}	Mean	3.89	0.089	0.788		7274.9
	Median	3.64	0.083	0.782		
	++/−−	10/2	36/0	50/0		
4. $\Delta_1 IV^2_{t-1} + \varepsilon^2_{t-1}$	Mean	2.35	0.078	0.845	0.677	7281.6
	Median	1.58	0.076	0.843	0.629	
	++/−−	7/4	30/0	50/0	31/0	

Mean and median parameter estimates across 50 stocks for the following model of daily returns:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \qquad h_t = \beta_0 + \beta_1 \varepsilon^2_{t-1} + \beta_3 IV^2_{t-2} + \beta_4 \Delta_1 IV^2_{t-1}$$

Note. R_t is the firm's simple discrete return (with dividends), ε_t is the return residual, h_t is the conditional variance of ε_t, IV^2_t is the daily implied variance from the options of the respective stock, $\Delta_1 IV^2_t$ is the 1-day change in the implied variance from day $t-1$ to day t, and the αs and βs are estimated coefficients. The sample period is January 1988 through December 1995. The system is estimated simultaneously by maximum likelihood estimation using the conditional normal density. The table also reports the number of coefficients that are positive and significant (++) and the number that are negative and significant (−−) at the 5% level, calculated with quasi-maximum likelihood standard errors in accordance with Bollerslev and Wooldridge (1992). Panels (a) and (b) present the results for the top and bottom quintiles of the firms, sorted by option trading volume. Panel (c) presents the results for all 50 firms.

daily change in IV contains reliable volatility information. The estimated β_4 is positive and statistically significant for 37 of the 50 firms. Further, the size of β_4 is only marginally smaller than the β_3 coefficient on IV^2_{t-2}, which suggests that there is only a modest amount of noise in the daily IV change. Third, the line-three model indicates that the $t-1$ return shock also contains reliable volatility information beyond the older IV^2_{t-2}. Finally, the full model in line four indicates that both the IV change and the lagged return shocks provide incremental volatility information jointly.

Next, the estimation of the conditional variance specification (5) evaluates whether older return shocks ($t - 2$) provide incremental volatility information beyond IV^2_{t-3} when also controlling for the IV change from $t - 3$ to $t - 1$. The estimation finds that the β_2 on the $t - 2$ return shock is positive and significant for only 1 of the 50 firms. Further, the estimated β_4 on $\Delta_2 IV^2_{t-1}$ is positive and statistically significant for 40 firms. Thus, the results in this subsection reinforce the earlier conclusions: For the high-option-volume firms, IV plus the first lagged return shock encompasses essentially all information about future volatility.

Individual Stock Returns and Index Implied Volatility

Having verified that implied volatility has significant power in explaining and forecasting conditional volatility, it is an interesting question whether the implied volatility of index options can help explain the conditional volatility of individual stocks. Parameter estimates from specification (7) are reported in Table IV. The estimated β_5 coefficients for the first model indicate that the implied volatility of index options is

TABLE IV

Conditional Volatility of Individual Stock Returns, Own-Firm Volatility, and Market-Level Implied Volatility

Model		$\beta_0 \times 10^5$	β_1	β_3	β_5	γ_1	$\mathcal{L}$
(a) Top Quintile of Firms by Option Volume							
1. VIX	Mean	23.2			0.720		7278.0
	Median	32.0			0.750		
	++/−−	10/0			6/0		
2. GARCH+VIX	Mean	3.68	0.053		0.592	0.712	7307.2
	Median	1.83	0.053		0.680	0.868	
	++/−−	3/1	5/0		5/0	7/0	
3. IV+VIX	Mean	0.59		1.013	−0.091		7340.4
	Median	0.46		0.933	−0.031		
	++/−−	0/1		10/0	0/3		
4. GARCH+IV+VIX	Mean	0.13	0.050	0.954	−0.141	0.150	7342.0
	Median	0.24	0.047	0.885	−0.036	0.089	
	++/−−	1/1	5/0	10/0	0/3	4/2	

(*Continued*)

TABLE IV (*Continued*)

Model		$\beta_0 \times 10^5$	β_1	β_3	β_5	γ_1	$\mathcal{L}$
(b) Bottom Quintile of Firms by Option Volume							
1. VIX	Mean	16.8			1.561		7285.7
	Median	13.1			0.608		
	++/−−	9/0			8/0		
2. GARCH+VIX	Mean	4.90	0.071		0.736	0.688	7316.6
	Median	1.99	0.073		0.335	0.725	
	++/−−	4/0	5/0		3/0	7/0	
3. IV+VIX	Mean	4.50		0.652	0.660		7330.5
	Median	3.99		0.676	0.175		
	++/−−	5/0		8/0	2/0		
4. GARCH+IV+VIX	Mean	2.71	0.081	0.498	0.594	0.435	7340.4
	Median	1.11	0.078	0.475	0.153	0.463	
	++/−−	2/0	7/0	8/0	2/0	6/0	
(c) Full Sample							
1. VIX	Mean	22.6			1.172		7228.1
	Median	18.3			0.820		
	++/−−	48/0			38/0		
2. GARCH+VIX	Mean	6.63	0.077		0.832	0.603	7259.8
	Median	2.37	0.077		0.625	0.693	
	++/−−	19/2	36/0		27/0	31/0	
3. IV + VIX	Mean	3.22		0.837	0.175		7278.8
	Median	2.43		0.841	0.157		
	++/−−	8/3		46/0	6/4		
4. GARCH+IV+VIX	Mean	2.50	0.072	0.744	0.148	0.141	7284.9
	Median	1.15	0.075	0.756	0.134	0.120	
	++/−−	4/3	30/0	48/0	5/4	14/4	

Mean and median coefficients across 50 stocks for the following GARCH system on daily returns:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \qquad h_t = \frac{\beta_0 + \beta_1 \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{t-1}^2 + \beta_5 VIX_{t-1}^2$$

Note. R_t is the firm's simple discrete return (with dividends), ε_t is the return residual, h_t is the conditional variance of ε_t, IV_t^2 is the daily implied variance from the options of the respective stock, VIX_t^2 is the daily implied variance from OEX index options, L indicates the lag operator, and the αs, βs, and γ are estimated coefficients. The sample period is January 1988 through December 1995. The system is estimated simultaneously by maximum likelihood estimation using the conditional normal density. The table also reports the number of coefficients that are positive and significant (++) and the number that are negative and significant (−−) at the 5% level, calculated with quasi-maximum likelihood standard errors in accordance with Bollerslev and Wooldridge (1992). Panels (a) and (b) present the results for the top and bottom quintiles of the firms, sorted by option trading volume. Panel (c) summarizes the results for all 50 firms.

positively and reliably related to conditional volatility for 38 out of the 50 stocks. This suggests that index implied volatility may be a useful forecasting tool for stocks that do not have traded options. Results from the second specification indicate that in 27 cases, this result still holds after modeling GARCH(1,1) effects. The inclusion of index implied volatility somewhat erodes the magnitude of the GARCH γ_1 coefficient from a mean of 0.856 for the standard GARCH model to a mean of 0.603 for the model in line two of Table IV.

Next, the line-three model in Table IV jointly evaluates the index implied volatility and the own-firm implied volatility. The results suggest that the index implied volatility typically adds little over the own-firm implied volatility. For the top volume-quintile of firms, the average β_5 coefficient is not even positive. In contrast, for the bottom volume-quintile of firms, the estimated β_5 coefficient is positive for 8 of the 10 firms. However, the estimated β_5 is positive and statistically significant for only 2 of the 10 low option-volume firms.

Finally, the model in line four jointly includes the standard GARCH terms, own firm implied volatility, and the index implied volatility. For this model, the incremental information from the implied volatility's is similar to that in line three.

A potential concern for the estimation of (7) is that IV and VIX may be highly correlated. However, for the majority of the firms, the correlation is fairly modest. Across the sample of 50 firms, the mean (median) correlation between a firm's IV and the VIX is 0.52 (0.54), which translates to an R-squared of 27% (29%). These correlations range from -0.12 to $+0.88$. Further, the GARCH+IV model in Table II outperforms the GARCH+VIX model in Table IV for 48 of the 50 firms. These results suggest that multicollinearity between VIX and firms' IVs is not a material issue in this study.

To conclude, for the most actively traded options, the firm's implied volatility captures most or all of the relevant information about future expected firm volatility. The index implied volatility provides incremental information beyond a univariate GARCH approach for the majority of firms, which implies that the index implied volatility may provide useful volatility information for firms without traded options. For less actively traded options, the implied volatility of index options contains incremental information beyond that contained in the firm's own implied volatility for a few firms. These findings also suggest that the NASDAQ index implied volatility series (VXN) might prove useful in volatility models for technology firms and other firms that are more correlated with the NASDAQ index than with the large-firm S&P 100 Index.

Implied Volatility and Mean Returns

In the interest of brevity, the detailed results of the specification where implied volatility is included in the mean equation [Equations (8) and (9)] are not reported here. To summarize, adding the lagged implied volatility in the mean equation has a negligible effect on the variance equation. Second, little reliable association is found between the lagged level of implied volatility and the conditional mean. There tends to be a slight positive association between the conditional mean and implied volatility, but the relation is very marginal and not reliable. This is not surprising. Even if a positive relation truly exists, it would be hard to reliably measure with daily returns over an 8-year sample period.

Information from High Frequency 5-Minute Returns

Next, the volatility information from lagged high frequency intraday returns are examined, specifically the sum of squared 5-minute returns. In regards to modeling a firm's daily conditional volatility, the information from the 5-minute returns is compared to the information from standard GARCH terms and the firm's implied volatility. As noted earlier, data constraints limit the investigation here to the 1993 through 1995 period. Thus, the results in this section also serve to examine subperiod robustness of previous results.

Parameter estimates for the conditional volatility model (11) are reported in Table V. Within each panel of Table V, the model in line one reports estimates of the GARCH(1,1) model. The standard GARCH dynamics are evident over this 3-year sample. Although the statistical

TABLE V
Conditional Volatility of Individual Stock Returns, Implifed Volatility, and Intraday return Volatility

Model		$\beta_0 \times 10^5$	β_1	β_3	β_6	γ_1	γ_2	$\mathcal{L}$
(a) Top Quintile of Firms by Option Volume								
1. GARCH	Mean	5.09	0.061			0.764		2750.5
	Median	3.26	0.039			0.828		
	++/−−	2/0	2/0			8/0		
2. V5	Mean	7.94			0.476		0.430	2753.6
	Median	7.27			0.382		0.539	
	++/−−	1/0			5/0		5/1	
3. GARCH+V5	Mean	11.2	0.038		0.318	0.186	0.360	2752.9
	Median	13.7	0.024		0.183	0.260	0.380	
	++/−−	7/0	1/1		3/0	4/2	2/0	
4. IV	Mean	1.09		1.042				2758.2
	Median	2.14		0.938				
	++/−−	2/2		9/0				
5. ARCH+V5+IV	Mean	0.21	0.021	0.902	0.189			2761.4
	Median	0.35	0.006	0.740	0.204			
	++/−−	1/1	0/0	9/0	1/1			

TABLE V (*Continued*)

Model		$\beta_0 \times 10^5$	β_1	β_3	β_6	γ_1	γ_2	$\mathcal{L}$
(b) Bottom Quintile of Firms by Option Volume								
1. GARCH	Mean	5.16	0.057			0.716		2811.2
	Median	3.03	0.041			0.815		
	++/−−	3/0	2/0			9/0		
2. V5	Mean	6.65			0.468		0.444	2817.6
	Median	4.39			0.302		0.479	
	++/−−	2/0			2/0		7/1	
3. GARCH+V5	Mean	7.15	0.052		0.421	0.257	0.108	2816.2
	Median	4.92	0.043		0.125	0.303	0.131	
	++/−−	3/0	1/0		3/0	4/0	2/1	
4. IV	Mean	12.9		0.448				2817.0
	Median	10.6		0.569				
	++/−−	4/0		4/1				
5. ARCH+V5+IV	Mean	11.0	0.021	0.294	0.311			2821.4
	Median	8.44	0.005	0.271	0.296			
	++/−−	4/0	0/0	2/1	2/0			
(c) Full Sample								
1. GARCH	Mean	7.17	0.072			0.683		2767.8
	Median	2.94	0.052			0.846		
	++/−−	16/0	14/0			38/1		
2. V5	Mean	7.85			0.378		0.448	2770.5
	Median	4.75			0.292		0.505	
	++/−−	14/2			28/0		29/3	
3. GARCH+V5	Mean	7.20	0.049		0.321	0.161	0.305	2770.7
	Median	8.00	0.045		0.199	0.099	0.357	
	++/−−	26/3	4/6		19/0	17/4	18/2	
4. IV	Mean	8.69		0.691				2771.2
	Median	5.60		0.717				
	++/−−	16/4		33/1				
5. ARCH+V5+IV	Mean	7.42	0.042	0.528	0.244			2776.1
	Median	4.20	0.028	0.518	0.230			
	++/−−	12/2	1/3	29/1	8/1			

Mean and median coefficients across 50 stocks for the following model of daily returns

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \qquad h_t = \frac{\beta_0 + \beta_1 \varepsilon_{t-1}^2}{1 - \gamma_1 L} + \beta_3 IV_{t-1}^2 + \frac{\beta_6 V5_{t-1}}{1 - \gamma_2 L}$$

Note. R_t is the firm's simple discrete return (with dividends), ε_t is the return residual, h_t is the conditional variance of ε_t, IV_t^2 is the daily implied variance from the options of the respective firm, $V5_t$ is the firm's intraday sum of squared 5-minute returns, L indicates the lag operator, and the αs, βs, and γs are estimated coefficients. The sample period is January 1993 through December 1995. The system is estimated simultaneously by maximum likelihood estimation using the conditional normal density. The table also reports the number of coefficients that are positive and significant (++) and the number that are negative and significant (−−) at the 5% level, calculated with quasi-maximum likelihood standard errors in accordance with Bollerslev and Wooldridge (1992). Panels (a) and (b) present the results for the top and bottom quintiles of the firms, sorted by option trading volume. Panel (c) summarizes the results for all 50 firms.

significance of the individual coefficients is less compared to the Table II results, possibly due to the smaller sample, the same qualitative results are evident for the estimated β_1 and γ_1 parameters.

The line-two model in Table V evaluates the information about future volatility that is contained in the sum of squared 5-minute returns, denoted as variable *V*5 in the model. In terms of the likelihood function value, the model in line two outperforms the standard GARCH model for 31 of the 50 firms. Next, the line-three model in Table V evaluates a specification that includes both standard GARCH terms and the sum of squared 5-minutes returns. For this model, overall, *V*5 is more reliably associated with future volatility than the daily return shocks. Thus, these results support recent literature that argues that there is valuable volatility information in lagged high-frequency intraday returns.

Next, the line-four model in Table V examines the performance of implied volatility, by itself, over this 3-year period. For the top volume-quintile of firms, implied volatility again outperforms the approach that relies on lagged return shocks. For 9 of these 10 firms, the estimated β_3 is positive and significant and the likelihood function value is higher than for the line-three model. In contrast, for the bottom volume-quintile of firms, the estimated β_3 is reliably positive for only four firms. Further, the line-three model outperforms the line-four model for 6 of these 10 firms. Thus, once again, the results seem to vary with option trading volume.

In line five of Table V, a volatility specification is estimated that includes standard ARCH, the *V*5 measure, and the firm's implied volatility. This specification does not include $t - 2$ and older return-shock terms as explanatory terms (through the γ coefficients) because this added complexity adds essentially no explanatory power.[5]

For the top volume-quintile of firms, the implied volatility captures nearly all of the volatility predictability. For 9 of the 10 firms, the estimated β_3 remains positive and statistically significant for the line-five model. Again, the results vary across firms with option trading volume. For the lowest option-volume quintile, the estimated β_3 is positive and significant for only two firms.

To conclude, the results in Table V reinforce the prior conclusions. For the highest option-volume firms, implied volatility captures essentially all of the relevant information about future volatility. However, even across the sample of 50 firms with actively traded options, the results vary substantially with volume. Once one drops below the top 10 or so option-volume firms, past return shocks generally provide incremental information about future volatility. In general, these results indicate that

[5]A full unrestricted version of equation (11) that includes the γ terms was also estimated. For 38 of the 50 firms, this additional complexity resulted in no increase in the likelihood function, as compared to the model in Table V, line five.

one should model conditional volatility as a function of implied volatility and past return shocks. However, return shocks from period $t - 2$ and older provide essentially no incremental information.

CONCLUSION

This article reports new empirical results on the information content of implied volatility, with respect to modeling and forecasting the volatility of individual firm returns. The empirical investigation examines 8 years of daily data on S&P 100 options and 50 individual stock options traded on the Chicago Board Options Exchange. In addition to reexamining the volatility information in implied volatility, the cross-sectional breadth of the sample enables this study to provide new evidence in several areas.

Overall, for stocks with very active option markets, the findings suggest that implied volatility subsumes nearly all information about future firm volatility. Most GARCH effects can be captured by modeling conditional volatility as a linear function of the implied variance from the firm's options. For the most liquid options, it appears that one may safely omit all ARCH/GARCH return-shock explanatory terms, or else include just ARCH(1). Out-of-sample volatility forecasting results provide similar conclusions. Thus, this article's results for the high option-volume firms are quite similar to results in Blair et al. (2001) for index options.

The results in this study are in contrast to earlier firm-level results in Lamoureux and Lastrapes (1993), who find that GARCH(1,1) outperforms implied volatility in a sample of 10 individual firms over April 1982 through March 1984. This contrast is even stronger because this article uses recent advances in time-series volatility models to incorporate the leverage asymmetry and information from high-frequency intraday returns.

However, for many of the stocks with relatively lower option trading volume, the information from implied volatility is comparable, or even inferior, to an approach that relies on the past time-series of return shocks. This degradation is evident even though the sample is composed of the 50 most actively traded CBOE firm options over the sample period. Presumably, the quality of implied volatility from the many other lower option-volume stocks would decrease further.

Next, for the majority of the firms, this article also documents that index implied volatility contains reliable incremental volatility information beyond a standard GARCH(1,1) specification of firm returns. These results suggest that implied volatility from equity index options can be useful in cases where no firm options are available, or where the information content of a firm's implied volatility is low due to an illiquid option market.

It is important to note that this study does not address recent literature that attempts to measure and explain the forecast bias of the Black-Scholes/binomial implied volatility (see, for example, Chernov, 2001; Figlewski, 1997; Poteshman, 2000). This important literature, along with the mainstream empirical option pricing literature, indicates that there appears to be stochastic volatility and/or jump risk premia impounded into option prices. This issue is beyond the scope of the current article, as this study is not interested in testing option pricing models, merely in extracting information from option prices. It is conceivable, however, that more information could be extracted from option prices using a pricing model that explicitly incorporates stochastic volatility. In this case, it seems likely that the results would even more strongly support this study's conclusion that information from actively traded options subsumes nearly all relevant information about conditional variance.

Finally, the evidence in this article suggest several areas for future research on the dynamics of individual stock prices. First, the results have direct implications for risk management, event studies, or other applications that rely on estimates of conditional firm-level volatility. Second, the findings suggest that further efforts to understand cross-sectional variation in implied volatility information may prove interesting. For example, with respect to option pricing, future work comparing individual stock options and index options may prove useful to better understand risk-based explanations for why the Black-Scholes implied volatility is biased. Is the constant volatility model biased only at the market level, or do individual stock option prices also contain volatility and jump risk premia? If they do, is the bias cross-sectionally related to the firm's systematic risk? Further analysis of these and other related issues should lead to an improved understanding of volatility in securities markets and the information impounded into option prices.

BIBLIOGRAPHY

Amin, K., & Ng, V. (1997). Inferring future volatility from the information in implied volatility in Eurodollar options: A new approach. The Review of Financial Studies, 10, 333–367.

Andersen, T., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review, 49, 885–905.

Andersen, T., Bollerslev, T., Diebold, F., & Ebens, H. (2001). The distribution of realized stock return volatility. Journal of Financial Economics, 61, 43–76.

Blair, B., Poon, S., & Taylor, S. (2001). Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns. Journal of Econometrics, 105, 5–26.
Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics, 31, 307–327.
Bollerslev, T., & Wooldridge, J. (1992). Quasi-maximum likelihood estimation inference in dynamic models with time-varying covariances. Econometric Theory, 11, 143–172.
Bollerslev, T., Chou, R., & Kroner, K. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. Journal of Econometrics, 52, 5–59.
Cai, J. (1994). A Markov model of unconditional variance in ARCH. Journal of Business and Economic Statistics, 12, 309–316.
Campbell, J., Lettau, M., Malkiel, B., & Xu, Y. (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. Journal of Finance, 56, 1–43.
Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. Review of Financial Studies, 6, 659–681.
Chernov, M. (2001). Implied volatilities as forecasts of future volatility, time-varying risk premia, and returns variability. Working Paper, Columbia University.
Chiras, D. & Manaster, S. (1978). The information content of option prices and a test of market efficiency. Journal of Financial Economics, 6, 213–214.
Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. Journal of Financial Economics, 50, 125–150.
Day, T., & Lewis, C. (1992). Stock market volatility and the information content of stock index options. Journal of Econometrics, 52, 267–287.
Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica, 50, 987–1008.
Engle, R. & Mezrich, J. (1995). Grappling with GARCH. Risk, September, 112–117.
Figlewski, S. (1997). Forecasting volatility. Financial Markets, Institutions and Instruments, 6, 1–88.
Fleming, J. (1998). The quality of market volatility forecasts implied by S&P 100 Index option prices. Journal of Empirical Finance, 5, 317–345.
Fleming, J., Kirby, C., & Ostdiek, B. (2001). The economic value of volatility timing. The Journal of Finance, 56, 329–352.
Glosten, L., Jagannathan, R., & Runkle, D. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance, 48, 1779–1801.
Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics, 42, 27–62.
Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroscedasticity and changes in regime. Journal of Econometrics, 64, 307–333.
Harvey, C. R., & Whaley, R. E. (1992). Market volatility prediction and the efficiency of the S&P 100 Index option market. Journal of Financial Economics, 31, 43–73.
Lamoureux, C. G., & Lastrapes, W. D. (1990). Persistence in variance, structural change, and the GARCH model. Journal of Business & Economic Statistics, 8, 225–234.

Lamoureux, C. G., & Lastrapes, W. D. (1993). Forecasting stock-return variances: Toward an understanding of stochastic implied volatilities. Review of Financial Studies, 6, 293–326.

Latane, H., & Rendleman, R. (1976). Standard deviations of stock price ratios implied in option prices. Journal of Finance, 31, 369–381.

Noh, J., Engle, R. F., & Kane, A. (1994). Forecasting volatility and option prices of the S&P Index. Journal of Derivatives, 2, 17–30.

Poon, S., & Granger, C. (2001). Forecasting financial market volatility: A review. Working Paper, University of Strathclyde and University of California, San Diego.

Poteshman, A. (2000). Forecasting future volatility from option prices. Working paper, University of Illinois at Urbana–Champaign.

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

Schwert, G. W. (1989). Why does stock market volatility change over time? Journal of Finance, 44, 1115–1153.

Scruggs, J. (1998). Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: A two-factor approach. Journal of Finance, 53, 575–603.

Whaley, R. E. (1993). Derivatives on market volatility: Hedging tools long overdue. Journal of Derivatives, 1, 71–84.