FEDERAL RESERVE BANK *of* ATLANTA

# Volatility Forecasts, Trading Volume, and the ARCH versus Option-Implied Volatility Trade-off

Glen Donaldson and Mark Kamstra

Working Paper 2004-6
March 2004

WORKING PAPER SERIES

FEDERAL RESERVE BANK *of* ATLANTA | WORKING PAPER SERIES

# Volatility Forecasts, Trading Volume, and the ARCH versus Option-Implied Volatility Trade-off

Glen Donaldson and Mark Kamstra

Working Paper 2004-6
March 2004

**Abstract:** Market expectations of future return volatility play a crucial role in finance; so too does our understanding of the process by which information is incorporated in security prices through the trading process. The authors seek to learn something about both of these issues by investigating empirically the role of trading volume in predicting the relative informativeness of volatility forecasts produced by ARCH models versus the volatility forecasts derived from option prices and in improving volatility forecasts produced by ARCH and option models and combinations of models. Daily and monthly data are explored. The authors find that if trading volume was low during period $t-1$ relative to the recent past, then ARCH is at least as important as options for forecasting future stock market volatility. Conversely, if volume was high during period $t-1$ relative to the recent past, then option-implied volatility is much more important than ARCH for forecasting future volatility. Considering relative trading volume as a proxy for changes in the set of information available to investors, their findings reveal an important switching role for trading volume between a volatility forecast that reflects relatively stale information (the historical ARCH estimate) and the option-implied forward-looking estimate.

JEL classification: G0

Key words: ARCH, volatility forecasting, VIX, options-implied volatility, trading volume

The authors thank Ron Giammarino, Lisa Kramer, Alan Kraus, and participants in various seminars and conferences for helpful discussions and the Social Sciences and Humanities and Research Council of Canada for financial support. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Glen Donaldson, Faculty of Commerce and Business Administration, University of British Columbia, 2053 Main Mall, Vancouver, British Columbia, Canada, V6T 1Z2, 604-822-8344, 604-822-8521 (fax), glen@donaldson.commerce.ubc.ca, or Mark Kamstra, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, Georgia 30309, 404-498-7094, 404-498-8810 (fax), mark.kamstra@atl.frb.org.

# 1 Introduction and Overview

Market expectations of future return volatility play a crucial role in finance; so too does our understanding of the process by which information is incorporated in security prices through the trading process. This paper seeks to learn something about both of these issues by investigating empirically the role of trading volume (a) in predicting the relative informativeness of volatility forecasts produced by ARCH models versus the volatility forecasts derived from option prices, and (b) in improving volatility forecasts produced by ARCH and option models and combinations of models.

Previous studies have reported that trading volume cannot forecast volatility directly. In this paper we uncover a new result: that volume does indeed have predictive power for forecasting volatility, with volume playing the role of a switching variable between states in which option-implied volatility is more or less informative than ARCH for volatility forecasting. Indeed, we find that the accuracy of volatility forecasts can be significantly improved by accounting for the volume effect and by combining information from ARCH models and option prices accordingly. This finding is made possible because of the novel way we incorporate trading volume into our functional forms and because, while previous papers have added either trading volume or option-implied volatility (but not both) to ARCH models, our study is the first we know of to consider all three factors together.

Traditional ARCH models – including GARCH, EGARCH, and so forth – forecast future return volatility given only information on lagged return innovations.[1] Previous papers that have studied the relationship between return-based volatility forecasts and trading volume have found that, while volume and returns are correlated contemporaneously, *lagged* volume has no power to forecast *future* volatility once the effects of lagged return innovations have been fully accounted for.[2] In other words, results from previous research suggests that in an

[1]Bollerslev, Cho and Kroner [1992] review the traditional ARCH literature. Examples and references to more recent work in areas such as seminonparametric ARCH modeling can be found in Donaldson and Kamstra [1997] and Engle and Ng [1993].

[2]See Brooks [1998], Heimstra and Jones [1994], Lamoureux and Lastrapes [1994] and Richardson and Smith [1994] for an analysis of *lagged* volume effects and returns. The relationship between *contemporaneous* volume and returns is outside the scope of our paper and thus the contemporaneous volume literature is not

ARCH model that already accounts for the impact of lagged return innovations on future volatility, lagged volume will have no marginal power to forecast future volatility. We find a different result on our paper. In particular, when we interact lagged volume with option-implied volatility in an augmented ARCH model, we uncover a significant role for lagged trading volume in forecasting future return volatility. We find this important volume effect, where others have not, because we include implied volatility in our ARCH investigation whereas previous volume/return studies have not.

Previous papers that have added option-implied volatility to ARCH have done so primarily to investigate the efficiency of option markets, not to improve ARCH forecasts *per se*. If option markets are efficient then option prices will contain all available information concerning the expected future volatility of underlying prices – including any information used by ARCH models – and thus volatility forecasts implied by option prices should encompass volatility forecasts from ARCH models. However, most studies to date have found that option-implied volatility cannot encompass ARCH in one-day-ahead volatility forecasting and have thus concluded that either the option market is not efficient or that the option-pricing models employed are misspecified, or are at least problematic for short-term forecasting.[3] We find that option-implied volatility can encompass a simple ARCH model at one day and one month horizons, but when the effects of lagged trading volume and option-implied volatility are incorporated in an augmented ARCH model, an implied volatility forecast is encompassed by this broader time series augmented information set. Considering relative trading volume as a proxy for changes in the set of information available to investors,[4] our findings reveal an important switching role for trading volume between a volatility forecast that reflects relatively stale information (the historical ARCH estimate)

cited here. However, references can be found the in the aforementioned articles and in Karpoff [1987].

[3]See, for example, Day and Lewis [1992] and Lamoureux and Lastrapes [1993]. Note that Christensen and Prabhala [1998] and Fleming [1998] have found that option-implied volatility can outperform ARCH at longer horizons (e.g., one-month-ahead forecasts in Christensen and Prabhala [1998]) once certain biases are accounted for (as in Fleming [1998]). Day and Lewis [1993] argue that applied to the crude oil futures market the options-implied volatility appears to subsume the information contained in other volatility forecasts, at the near-term horizon, though their near-term is 32.5 trading days on average, not one day.

[4]We thank the referee for pointing out this interpretation of relative trading volume.

and the option-implied forward-looking estimate. With trading volume low relative to the recent past ARCH is weighted more heavily in an augmented forecasting model than options for forecasting future stock market volatility and conversely, with volume high relative to the previous recent past, option-implied volatility is weighted more heavily than ARCH for forecasting future volatility.

The data and basic model specifications we employ in our investigation are presented in Section 2 below. In Section 3 we estimate an ARCH model of volatility and combine the resulting forecast with both an option-implied volatility forecast and lagged trading volume in order to observe the importance of each input. Section 4 presents a full Maximum Likelihood modeling of volatility using volume, implied volatility from options prices, and the history of returns, which we term augmented ARCH. In this section we detail our main estimation results. Section 5 has us exploring out-of-sample forecasting results, and Section 6 concludes.

# 2 Basic Models and Data

Data employed in this study range over the period from 1988 to 2003, analyzed at the daily and monthly frequency, with the stock returns coming from the S&P 100 index and option-implied volatilities coming from the Chicago Board of Exchange's VIX index.[5] Model selection is performed on in-sample data from 1988:1-1995:9, and data from 1995:10 to 2003:8 is our holdout sample for out-of-sample forecast evaluation,[6] roughly a half and half split of the data available to us.

Define stock returns, $R_t$, as the arithmetic return based on the daily closing value of the S&P 100 Stock Price Index, multiplied by 100. $E(R_t|I_{t-1})$ is then the conditional forecast of this return such that $R_t = E(R_t|I_{t-1}) + \epsilon_t$, where $I_{t-1}$ is the date $t-1$ conditioning information set on which date $t$ forecasts are based and the additive forecast error, $\epsilon_t$, has

[5]The CBOE recently redefined their VIX index. We use the original series based on S&P 100 index prices.

[6]In the holdout sample we measure realized volatility for volatility forecast evaluation as the squared residual from a model of the mean and the variance as constant, where the constants are measured with data up to but not including the forecast period. There are virtually no differences to using the squared residual from any other model we will discuss below.

zero mean and conditional variance $E(\epsilon_t^2|I_{t-1}) = \sigma_t^2$. A variety of specification tests on the daily S&P 100 data post-1987 (which is the time period we study due to the fact that index option markets were very thin prior to 1988) reveals that expected returns are appropriately modeled with a constant, such that $R_t = \mu + \epsilon_t$. As a baseline in our own investigations below we therefore employ the constant expected returns specification.[7] The simplest model we consider, the "Naive" model, has a constant mean and constant variance, forecasting variance as the average variance from the in-sample period.

A well-documented feature of stock market data is that the return innovations, $\epsilon_t$, appear to be drawn from a time-dependent heteroskedastic distribution. An important goal of the conditional volatility literature is to capture this feature of the data with the appropriate model for the conditional variance process so as to produce a forecasted variance, $\hat{\sigma}_t^2$, along with a return forecast error, $\hat{\epsilon}_t$, such that the standardized residuals, $\hat{\epsilon}_t/\hat{\sigma}_t$, are homoskedastic and independent.

## 2.1 ARCH

In the ARCH family of models the conditioning information set traditionally used to make volatility forecasts contains only the history of $\epsilon$. Lagged $\epsilon^2$ are included to capture the feature of volatility clustering; i.e., future volatility is related to lagged squared return innovations. Levels of lagged $\epsilon$ are also sometimes employed to capture the perception that volatility may be related in an asymmetric way to lagged return innovations, with sharp drops in stock prices causing more future volatility than upturns cause.[8] One specification that nests the popular GARCH model of Bollerslev [1986] and allows asymmetric responses to lagged return shocks is the asymmetric Sign-GARCH model of Glosten, Jagannathan and Runkle [1993] (GJR model), shown in Equations (1)-(3) below.

[7] In the interest of robustness we investigated a variety of alternative specifications for the expected return, including specifications in which expected returns are modeled as a simple AR(1) process, or as a more complicated seasonal process with dummy variables for January and Monday, plus AR terms as necessary to completely whiten the in-sample data. Our volatility results did not change appreciably. We therefore report the simple constant specification in the analysis below.

[8] See Engle and Ng [1993] for further discussion of asymmetric ARCH effects.

$$R_t = \mu + \epsilon_t \quad ; \quad \epsilon_t \sim (0, h_t^2) \tag{1}$$

$$h_t^2 = \alpha + \beta h_{t-1}^2 + \gamma \epsilon_{t-1}^2 + \delta D_{t-1} \epsilon_{t-1}^2 \tag{2}$$

$$D_{t-1} = \begin{cases} 1 & if\ \epsilon_{t-1} < 0 \\ 0 & otherwise \end{cases} \tag{3}$$

A model with $p$ lags of $h_t$, $q$ lags of $\epsilon_t^2$ and $r$ lags of $D_t \epsilon_t^2$ is labeled GJR(p,q,r). Such a model excluding the asymmetric volatility term $D_t \epsilon_t^2$ is labeled GARCH(p,q). For simplicity, models in the ARCH family will be referred to as simply "ARCH" when there is no ambiguity.

For this study, the ARCH model used is the lowest order model which removed evidence of residual autocorrelation, ARCH and sign-ARCH effects[9] and with our data and time period, the a very simple GARCH(1,1) was sufficient and is also ranked best by the Schwarz information criteria.[10] The GARCH(1,1) model uses one lag each of $h^2$ and $\epsilon_t^2$ in Equation (2) below. See Bollerslev [1986] for details on the GARCH model. The models we ranked and from which choose the GARCH(1,1) from are all of the form in Equations (1)-(3), with as many as two lags of $h^2$, $\epsilon^2$ and the asymmetric volatility term. See Appendix 1 for summary statistics and parameter estimates of the various models looked at.

Parameter estimates for the GARCH(1,1) model, estimated with Maximum Likelihood assuming normality,[11] are reported in Section 4 below, along with comparisons to other

[9] The sign-ARCH test probes for asymmetric volatility increases from negative shocks to the returns process. See Engle and Ng [1993] for details on the sign-ARCH tests.

[10] The Schwarz Criterion is widely used in the literature (e.g., Nelson [1991] uses Schwarz to select EGARCH models). The Schwarz Criterion has been shown to provide consistent estimation of the order of linear ARMA models by Hannan [1980]. As noted by Nelson [1991], the asymptotic properties of this criterion are unknown in the context of selecting ARCH models, hence we might also rely on the principle of parsimony among models that do not fail common specification tests to pick the favored model. This would lead us again to the GARCH(1,1) model.

[11] In the context of returns, which are arguably highly non-normal even after modeling their first two moments, this estimation should be viewed as quasi-maximum likelihood. We present standard errors based on Bollerslev and Wooldridge [1992], which are robust to non-normality.

models. Results from various in-sample and out-of-sample diagnostic tests and performance evaluations are also presented below in Sections 4 and 5.

## 2.2 Option-Implied Volatility

As our measure of the options-implied volatility we use the Chicago Board of Exchange (CBOE) VIX index. The VIX is a popular measure of stock market volatility, based on the classic Black-Scholes [1973] option-pricing formula using options prices on the S&P 100. This volatility index is produced using eight S&P 100 index puts and calls, and allows for market microstructure frictions such as bid-ask spreads, the American features of the option contracts, and discrete cash dividends, measuring the implied volatility on a synthetic at-the-money one month maturity option contract. The volatility estimate reported in the VIX is an annualized standard deviation and so must be rescaled for our application.[12] See Whaley [2000] for a very detailed look at the construction of the VIX index.

There is a maturity mismatch when using the VIX for volatility forecasting over horizons less than a month, given the VIX is specifically designed to forecast volatility over a 22 business day period (one calendar month, approximately). Although one can interpret the VIX volatility forecast as the average volatility over the coming month and hence a measure of the volatility over the coming day, at least two outcomes are likely. One, the VIX can be expected to be less efficient at forecasting volatility one day rather than one month ahead. Two, forecast errors from the VIX will be correlated over the month-long period that the VIX forecasts overlap. As a result we will look at the performance of our volatility forecasting techniques over both a daily and monthly period, and when comparing the VIX to other forecasts of volatility on daily data we will employ econometric techniques designed to be robust to a 22 day moving average forecast error process.

[12] While the derivatives and options literature typically makes use of the term "volatility" to mean standard deviation we will continue to use volatility to mean variance.

## 2.3 Volume

The final variable we consider is trading volume at the NYSE. Since our objective is to forecast volatility, we are interested in *lagged* volume; i.e., $Volume_{t-1}$, or some function of $Volume_{t-1}$, as an element of the $I_{t-1}$ forecasting information set. Since one purpose of our investigation is to determine whether ARCH and options behave differently on high versus low volume days, we first consider the high/low volume indicator variable $V_{t-1}$, where:

$$V_{t-1} = \begin{cases} 1 & if\ Volume_{t-1} \geq \frac{1}{(n-1)} \sum_{i=2}^{n} Volume_{t-i} \\ 0 & otherwise \end{cases} \tag{4}$$

We set $n = 5$ so that the variable $V_{t-1}$ is "one" if lagged volume is above its one-week lagged moving average (in the case of daily data), and is "zero" otherwise.[13] As shall be discussed below, this simple discrete indicator variable works as well as, and in many cases better than, continuous alternatives. We shall form this volume variable relative to the past week of volume for the daily forecasting exercises, and relative to the past 5 months of data for the monthly forecasting exercises below.

# 3 Combining Forecasts

## 3.1 The Combining Model and Results

Perhaps the most obvious way to isolate and compare ARCH, option and volume effects is to estimate a simple linear combination of the ARCH forecasts and option-implied forecasts, along with our high/low volume switching variable. We therefore first use Equations (1)-(3) and returns information from period $t-1$ to calculate an ARCH volatility forecast for period $t$, and denote this conditional volatility forecast as $\hat{h}_t^2$, where the "hat" denotes "forecast conditional on time $t-1$ information". We next use the VIX implied volatility estimate (formed in period $t-1$), noted $\hat{S}_t^2$. We then calculate our volume variable, $V_{t-1}$ from

[13] We also investigated a variety of other lag lengths, including a one-month lagged moving average, instead of one week, for daily data. Results from the estimations and tests discussed below were not affected qualitatively.

Equation (4). Finally, we combine these three variables in a joint mean-variance Maximum Likelihood regression to obtain our combined volatility forecast, $\hat{\sigma}_t^2$.

$$\sigma_t^2 = \alpha_0 + \alpha_1 V_{t-1} + \phi_{ARCH,0}\hat{h}_{t-1}^2 + \phi_{ARCH,1}V_{t-1}\sigma_{t-1}^2 + \phi_{Option,0}\hat{S}_{t-1}^2 + \phi_{Option,1}V_{t-1}\hat{S}_{t-1}^2 \quad (5)$$

$$V_{t-1} = \begin{cases} 1 & if\ Volume_{t-1} \geq \frac{1}{(n-1)}\sum_{i=2}^{n} Volume_{t-i} \\ 0 & otherwise \end{cases} \quad (6)$$

Parameter estimates from the combining regression[14] Equation (5) are reported in Table 1 below (with 2-sided t-tests in parentheses and Bollerslev-Wooldridge robust standard errors in brackets). The results in Table 1 are for the daily data we have, January 1988 to September 1995, for a total of 1959 observations.

---

**Table 1**
**Parameter Estimates from the Combining Regression**
**{Robust t-tests in braces}**
**(Bollerslev-Wooldridge robust standard errors in parenthesis)**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\hat{\sigma}_t^2 =$ | 0.009 | $-$ | 0.034 $V_{t-1}$ | $+$ | 0.364 $\hat{h}_t^2$ | $-$ | 0.74$V_{t-1}\hat{h}_t^2$ | $+$ | 0.309$\hat{S}_t^2$ | $+$ | 0.406$V_{t-1}\hat{S}_t^2$ |
| | {.100} | | {0.228} | | {1.54} | | {-2.21**} | | {2.38**} | | {1.89*} |
| | (0.016) | | (0.147) | | (0.237) | | (0.335) | | (0.13) | | (0.215) |

* Significant at the 10% level, two-sided test.
** Significant at the 5% level, two-sided test.
*** Significant at the 1% level, two-sided test.

$\hat{\sigma}_t^2$ = stock market return volatility on period $t$
$V_{t-1}$ = indicator variable - trading volume higher than average over past week
$\hat{h}_t^2$ = forecasted volatility from an ARCH model, conditional on $t-1$ information
$\hat{S}_t^2$ = forecasted volatility from option prices, conditional on $t-1$ information

[14] Forecast combining has a long tradition in the forecasting literature. For a general literature review see Clemens [1989]. For a more recent example see Donaldson and Kamstra [1996].

Log Likelihood = -2269.69, BIC= 4600.026. The jointly estimated mean parameter from Equation (1) is $\mu = 0.043$, with standard error of 0.016. There is little or no evidence of autocorrelation, ARCH or Sign-ARCH. The $R^2$ is 0 as the mean is modeled with only a constant term. A more parsimonious version of this model is presented in Section 4, and more detailed diagnostics are presented there.

---

From the theoretical studies on trading volume cited in the introduction to this paper and discussed more fully below (e.g., Admati and Pfleiderer [1988], He and Wang [1995], etc.), we would expect market prices may be more "informative" during high volume periods and thus would expect option-implied volatility, which is based on market prices, to forecast volatility more accurately, while ARCH models based on a long history of lagged prices would not necessarily prove helpful in forecasting. The parameter estimates we find are consistent with this expectation; when volume is light relative to the past week the weight assigned to the ARCH forecast and the options implied forecast are roughly equal, with a coefficient of 0.0309 on the option-implied term and 0.364 on the ARCH term (although the ARCH coefficient is not quite significant at the 10% level) and when volume is heavy relative to the past week the weight on the option-implied forecast more than doubles to 0.715 (the sum of the coefficients on $\hat{S}_t^2$ and $V_{t-1}\hat{S}_t^2$) and the weight on the ARCH forecast now becomes negative, -0.376 (the sum of the coefficients on $\hat{h}_t^2$ and $V_{t-1}\hat{h}_t^2$).[15]

We conducted several tests to check the robustness of this result. First, we investigated different definitions of volume (e.g., Equation (4) based on a one-month lag instead of a one-week lag), different data series of option-implied volatility and returns (based on the S&P 500) and different definitions of ARCH and the mean equation (e.g., different models, different lag lengths, etc.). The resulting regression coefficients changed somewhat from specification to specification, but the basic finding remained: option-implied volatility dominates ARCH in high volume states and ARCH matches or dominates option-implied volatility in low volume states. Second, we separated our data into various subsamples and re-ran the

[15]Note that there is no restriction that the coefficients on ARCH and options sum to unity in our combining regression and thus the regression intercepts can be nonzero, which is the case above. See Clemens [1989] for a discussion of the pros and cons of restricting the intercept in a combining regression.

regression. Again, coefficients changed somewhat from subperiod to subperiod, but the key result was qualitatively robust. Third, instead of using the variance of the residual from Equation (1) as the dependent variable in the combining regression, we tried defining the dependent variable as option-implied volatility from day $t$. (Thus, we used ARCH, volume and option-implied volatility from day $t-1$ to forecast option-implied volatility from day $t$.) Here again we found our familiar result: option-implied volatility is a better forecaster of future volatility relative to ARCH when volume is high, no matter how the term "volatility" is defined. Regardless of the specification for volatility, this core result was robust. We also considered conducting various option-based tests on our volatility forecasts, including tests for pricing/hedging effectiveness, but did not do so because pricing/hedging effectiveness testing cannot be legitimately undertaken on the volume-ARCH-option combinations we study given the menu of option pricing models currently in existence.[16]

## 3.2 Discussion

As stated in the introduction to our paper, we are not attempting to test market efficiency in our study. It is, however, interesting to note that on days with high volume relative to the past week[17], option-implied volatility dominates ARCH and thus that, using the yardstick suggested by previous authors, the market may indeed be efficient when enough information is flowing into the market (assuming volume is a good proxy for information flow, as is frequently assumed – see for instance Admati and Pfleiderer [1988]). The failure of option-implied volatility to dominate ARCH on low relative volume days might suggest that, if

[16] To test the pricing/hedging effectiveness of combining ARCH and volume effects in an option pricing framework, one would first need to build an option-pricing model that explicitly allows for both volume and ARCH effects, and then test the pricing/hedging effectiveness of that volume-ARCH-option pricing model relative to option pricing models that do not contain volume and/or ARCH effects (for a similar approach to testing pricing/hedging effectiveness of various competing models see Bakshi, Cao and Chen [1997]). In particular, it would not be legitimate to simply plug volatility forecasts from various ARCH-volume models into the Black-Scholes Model (or indeed into any other currently available option pricing model) and observe its resulting pricing/hedging performance. This is because no option-pricing model we know of permits volume effects in the model and thus the option-pricing model, volatility forecasting procedure and test procedure would be inconsistent with each other. While the mechanical exercise is of course feasible, the resulting performance, good or bad, would vary unpredictably with new data since there is no way to reliably establish the properties of such an internally inconsistent methodology.

[17] These are days we speculate exhibit important changes in the information set available to investors.

the market is indeed somewhat inefficient, it may only be so when there is comparatively little information flowing. Alternatively, our results could be interpreted to reveal that the Black-Scholes model is misspecified in some way that is most clearly seen on low volume days and that the market is always efficient.

There appears to be at least two possible (not necessarily mutually exclusive) explanations for our finding that option-implied volatility provides a better volatility forecast relative to ARCH following high volume days: (a) the informativeness of the ARCH volatility forecast declines in high volume states, and/or (b) the informativeness of option-implied volatility increases in high volume states.

We begin by investigating whether there is any change in average volatility following high versus low volume days since this could potentially lead ARCH to under-forecast future volatility following heavy volume days and thus help to explain why ARCH does worse relative to options following high volume days. To examine this possibility we compared average squared errors from Equation (1) and average volatility forecasts from our ARCH and option-implied models on day $t$ when volume was high versus low on day $t-1$. We found[18] that when volume was high on day $t-1$ the average day $t$ squared error is 0.724 and is 0.721 when volume was low on day $t-1$, and that when volume was high (relative to the past week) on day $t-1$ the average day $t$ deviation between the ARCH and option-implied forecasts shrinks compared to the deviation following relatively low volume periods. In other words, squared pricing errors are almost identical following high and low volume days and the ARCH model matches on average the option-implied volatility better on days following high volume. We see the close match of average squared pricing errors on high and low volume days in the insignificant intercepts in the combining regression in Table 1. This suggests that the ARCH versus options effect is not coming from average volatility levels and thus that any explanation of our results is more likely to rest on intertemporal correlations between forecasted and realized volatility.

To investigate correlation effects we computed the simple correlation between realized

[18]This analysis was conduced on our in-sample data, 1988 to September 1995.

volatility at time $t$, as measured by time $t$'s squared return innovation, and the ARCH (option-implied) volatility forecast at time $t-1$.[19] This correlation equals 24.8% (26.0%) when volume is low on day $t-1$ and is 15.8% (23.1%) when volume on day $t-1$ is high. On the full sample, January 1988 to August 2003, we see a similar pattern, with the correlation between realized volatility at time $t$ and the ARCH (option-implied) volatility forecast at time $t-1$ equaling 33.4% (34.5%) when volume is low on day $t-1$ and 26.8% (35.4%) when volume on day $t-1$ is high. In other words, ARCH volatility works best following a low volume market, and option-implied volatility is roughly as good following high or low volume periods (relative to recent volume levels). It therefore seems likely that the ARCH versus options effect is at least partially driven by ARCH doing worse following high volume days than low volume days and option-implied volatility doing as well on either high or low volume days. The conditional analysis and formal tests that follow will suggest further that the benefit derived from the time series information embedded in ARCH and volume come mainly from improving forecasts following low volume states, with the weight on time series information flipping to 0 or negative values following high volume periods.

# 4 Augmented ARCH

Forecast combining was appropriate above because our purpose was to reveal starkly the basic ARCH-volume-option relationship. Forecast combining is also a useful tool for situations in which the econometrician possesses the forecasts produced by various models but not the information sets used to produce the forecasts.[20] However, we do possess the information set on which at least the ARCH forecasts are based and thus, to produce optimal volatility forecasts, we should ideally add option and volume information to the ARCH model directly and estimate an augmented ARCH mega-model. We therefore investigate augmented ARCH models in this section.

[19] Again, this analysis was conduced on our in-sample data, 1988 to September 1995.

[20] See Clemens [1989].

## 4.1 The Model

The augmented-ARCH model we employ is given below, in which $R_t$ is the daily arithmetic stock return (multiplied by 100) and $S_t^2$ is option-implied return variance.

$$R_t = \mu + \epsilon_t \quad ; \quad \epsilon_t \sim (0, \sigma_t^2) \tag{7}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 V_{t-1} + \beta_0 \sigma_{t-1}^2 + \beta_1 V_{t-1} \sigma_{t-1}^2 + \gamma \epsilon_{t-1}^2 + \phi_{Option,0} S_{t-1}^2 + \phi_{Option,1} V_{t-1} S_{t-1}^2 \tag{8}$$

$$V_{t-1} = \begin{cases} 1 & if\ Volume_{t-1} \geq \frac{1}{(n-1)} \sum_{i=2}^{n} Volume_{t-i} \\ 0 & otherwise \end{cases} \tag{9}$$

Equations (7)-(9) are estimated jointly under Maximum Likelihood with $n = 5$, as are a number of interesting restricted and extended versions of (7)-(9). Results are reported below.

To better understand the effects of adding lagged volume and implied volatilities to ARCH, we investigated many possible combinations and permutations within our model, including: each variable alone, each possible combination, variables interacting with each other, and so forth. We also expanded our model to investigate a variety of different functional forms for the conditional volatility, including variables – and groups of variables – added and interacted nonlinearly. Table 2 presents daily data in-sample estimation results for a small collection of models which revealed the most interesting information concerning the effects of volume and implied volatility, for daily data January 1988 to September 1995. Summary statistics on daily and monthly data for these most interesting models follow in Table 3. In all cases our core result – that options provide better forecasts relative to ARCH on high volume days than on low volume days – remains qualitatively robust.

## 4.2 Parameter Estimates

Panel A of Table 2 below reports parameter estimates (with Bollerslev-Wooldridge robust standard errors in parentheses) for the most interesting specifications contained within Equation (5) and Equations (7)-(9). Panel B reports common diagnostics for each model in question. These diagnostics include: the model log likelihood, the p-value from a test for residual autocorrelation (Wald AR),[21] the p-value from a traditional Ljung-Box [1978] test for symmetric ARCH at 24 lags, and the p-values from an Engle-Ng [1993] Sign Bias Test, Negative Sign Bias Test, Positive Sign Bias Test and Joint Sign Bias Test – all at 5 lags – for the presence of asymmetric ARCH effects. P-values below $Z\%$ therefore reject, with $Z\%$ confidence, the null hypothesis that there is no AR/ARCH effect in favor of the alternative hypothesis that the model in question has uncaptured AR/ARCH. The model $R^2$ is 0, as the mean equation includes only an intercept term, and so this statistic is not reported.

We begin our analysis of Table 2 by considering the results from the basic GARCH(1,1) specification, as reported in Column 3 (labeled "ARCH"). Note from Panel A Column 3 that all the parameter estimates from the ARCH model are of the expected sign and magnitude and, from Panel B, that the model passes all standard specification tests at conventional significance levels (e.g., there are no p-values below 0.050 in Panel B, Column 2). Note in particular from Panel A that the parameter on lagged conditional volatility, $\beta_0$, is close to unity, which reveals the highly persistent nature of stock return volatility.

Column 1 ("Naive") of Table 2 reports results from a constant mean and constant variance model, forecasting next period variance as the average variance from the in-sample period. This model displays gross evidence of misspecification with very strong residual ARCH and Sign-ARCH effects, though no evidence of autocorrelation.

Column 2 ("Options Only") of Table 2 reports results from option-implied volatility alone; i.e., results from the Equations (7)-(9) estimated with all parameters in Equations (7)-(8) set to zero except for the intercepts and $\phi_{Option,0}$. The interesting results from Column 2

[21] This test is a $\chi^2$ Wald test on 5 lags of the residual in the mean equation.

come from Panel B. Note in particular that the log likelihood from option-implied volatility alone exceeds the log likelihood from ARCH in Column 3. Also note from Panel B that option-implied volatility passes all of the various ARCH tests at the 5% significance level.

**Table 2**
**Estimation Results and Diagnostics for**
**Volatility Models based on Equations (5), (7)-(9) and Daily Data**

| | **Panel A: Parameter Estimates** (Robust Standard Error) | | | | |
|---|---|---|---|---|---|
| Column → | 1 | 2 | 3 | 4 | 5 |
| Parameter Name ↓ | Naive | Options Only | ARCH Model | Combining Model | Full MLE |
| $\mu$ | 0.047** | 0.042** | 0.047*** | 0.043*** | 0.045*** |
| | ( 0.019) | ( 0.016) | ( 0.016) | ( 0.016) | ( 0.016) |
| $\alpha_0$ | 0.723*** | 0.010 | 0.002 | 0.019 | 0.032*** |
| | ( 0.049) | ( 0.050) | ( 0.001) | ( 0.053) | ( 0.010) |
| $\beta_0$ | | | 0.980*** | | 0.842*** |
| | | | ( 0.008) | | ( 0.114) |
| $\beta_1$ | | | | | -1.31*** |
| | | | | | ( 0.336) |
| $\gamma$ | | | 0.016** | | -.016 |
| | | | ( 0.007) | | ( 0.014) |
| $\phi_{Option,0}$ | | 0.519*** | | 0.307** | 0.049 |
| | | ( 0.059) | | ( 0.130) | ( 0.055) |
| $\phi_{Option,1}$ | | | | 0.411* | 0.709*** |
| | | | | ( 0.212) | ( 0.184) |
| $\phi_{ARCH,0}$ | | | | 0.350 | |
| | | | | ( 0.222) | |
| $\phi_{ARCH,1}$ | | | | -.706** | |
| | | | | ( 0.330) | |
| | **Panel B: Diagnostics** | | | | |
| Log Likelihood | -2461.41 | -2274.27 | -2311.62 | -2269.75 | -2269.71 |
| BIC | 4945.565 | 4578.857 | 4661.135 | 4592.553 | 4600.070 |
| Wald AR Test | 0.119 | 0.133 | 0.124 | 0.095 | 0.071 |
| LM ARCH Test | 0.000 | 0.965 | 0.980 | 0.968 | 0.965 |
| Sign Test | 0.529 | 0.609 | 0.521 | 0.639 | 0.639 |
| Neg. Sign Test | 0.000 | 0.851 | 0.446 | 0.792 | 0.801 |
| Pos. Sign Test | 0.000 | 0.063 | 0.059 | 0.042 | 0.100 |
| Joint Sign Test | 0.000 | 0.612 | 0.224 | 0.485 | 0.555 |

* Significant at the 10% level, two-sided test.
** Significant at the 5% level, two-sided test.
*** Significant at the 1% level, two-sided test.

Column 4 ("Combining Model") of Table 2 reports results from the combining regression model of Equation (5) constraining the volatility intercept dummy variable $V_{t-1}$ to have a coefficient of 0. This model also removes most evidence of Sign-ARCH effects and looks very similar to the combining model including the intercept volume dummy variable. The log likelihood is, of course, improved relative to either the ARCH or option-implied model alone, although the BIC criteria favors the simple option-implied volatility model over the more highly parameterized combining model.

Column 5 ("Full MLE") of Table 2 reports results from adding option-implied volatility and our high/low volume indicator variable to the standard ARCH model, a representative but parsimonious model of the set of models that could be constructed from interacting the volume variable with the various ARCH variables and the option-implied volatility variable. Unreported results reveal that, either alone or when added to a standard ARCH model, lagged volume has no power to *predict* volatility. From this one might be tempted to conclude, as previous researchers have concluded, that lagged volume has no power to forecast future volatility once the effects of lagged return innovations have been accounted for. However, such a conclusion would be premature. It would be more accurate to argue that, while volume cannot by itself forecast volatility, it does play an important regime-switching role, interacting with other variables in the model as we have already seen. Here we see that, on high volume days, the weight on option-implied volatility increases and the weight on the lagged conditional variance decreases (i.e., $\beta_1 < 0$ and $\phi_{Option,1} > 0$). In other words, the "Full MLE" model from Table 2 confirms our findings from the simple combining exercise we reported in Table 1. The results from this model are, however, even more dramatic than those from the simple combining exercise. The weight on the option-implied variable is effectively 0 in the low volume state, equal to 0.049 and statistically insignificant, and the weight on the lagged conditional variance flips to nearly -0.5 in the high volume (relative to the last weeks' average volume). Compared to the Combining Model reported in Table 2, however, the log likelihood and BIC give little reason to prefer the Full MLE model over the Combining Model on daily data. As will be presented, out-of-sample forecast performance confirms this for daily data but shows advantage to the Full MLE model for monthly forecast horizons.

Results for the daily data model estimations for the full sample, 1988 to 2003, as well as for monthly data, 1988-1995 and the full sample, can be found in Appendix 2. These results are qualitatively identical to the results discussed above. An outstanding difference is that residual Sign-ARCH effects for the daily data full sample are not expunged by any of the models, though standard ARCH and autocorrelation effects are not statistically significant for these models and data. Also, the full sample daily results are not as statistically significant and coefficient estimates of interest are not as large in magnitude as for the subsample 1988-1995. In contrast, the monthly results are quite strong both in and out of sample, and change very little as the sample is extended to 2003 from 1995. Given that the option-implied volatility estimate is calibrated to make a one month volatility forecast, not a one day forecast, it is perhaps not surprising that the one-month horizon results are more stable. Related issues will be explored with out-of-sample forecasting experiments, reported further below.

Finally, we investigated a number of nonlinear extensions to Equations (7)-(9), including models based on the Artificial Neural Network specifications. Interestingly, results from various (unreported) tests revealed that adding lagged volume and lagged option volatility to ARCH models removes the need to explicitly model nonlinear effects in lagged return innovations. In other words, volume, and especially options, seem to account for the nonlinear effects otherwise omitted from standard ARCH models. This suggests that ARCH modelers may benefit at least as much from expanding the traditional ARCH information set to include variables such as option-implied volatility and volume as from building ever more complex nonlinear models based only on lagged return innovations.

### 4.3 Summary Statistics

Tables 3.1 and 3.2 presents some summary statistics on all of Table 2's models, at the daily and monthly data frequencies.[22] The first row of statistics in Tables 3.1 and 3.2, marked "Raw Data", reports statistics for the S&P 100 returns (recall that returns are arithmetic and multiplied by 100). The row marked "ARCH" is for the basic GARCH(1,1) model in Equations (1)-(3). "Option" signifies variance defined as lagged option-implied volatility[23]; i.e., Equations (7)-(9) with all parameters zero except $\mu$, $\alpha_0$ and $\phi_{Option,0}$. The "Combining" model is a modified version of the model presented in Section 3 of this paper, in which the volume dummy variable $V_{t-1}$ is omitted, as in Column 4 of Table 2. The last row of Table 3 reports results from the "Full MLE", which is the maximum likelihood combination of ARCH, volume and options obtained by estimating Equations (7)-(9) including interaction terms between volume and the option-implied volatility, and volume and the lagged conditional variance, as in Column 5 of Table 2 above.[24]

**TABLE 3.1**
**Summary Statistics**
**Daily Data, January 1, 1988 to September 29, 1995**

| Method | Variance Forecasts | | | | | Standardized Returns | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std | Skew | Kurtosis | RMSE | Std | Skew | Kurtosis |
| Raw Data | 0.723 | 2.161 | 13.660 | 275.757 | 2.991 | 1.000 | -0.499 | 6.963 |
| ARCH | 0.723 | 0.456 | 1.855 | 4.130 | 2.725 | 1.001 | -0.573 | 5.397 |
| Option | 0.715 | 0.510 | 2.396 | 8.092 | 2.503 | 1.000 | -0.495 | 4.278 |
| Combining | 0.712 | 0.512 | 2.615 | 11.006 | 2.464 | 1.000 | -0.490 | 4.084 |
| Full MLE | 0.707 | 0.507 | 2.774 | 11.709 | 2.470 | 1.000 | -0.470 | 4.113 |

[22] The column reporting kurtosis presents the excess kurtosis, that is, the kurtosis in excess of 3. Were the data normally distributed the excess kurtosis would be 0.

[23] Suitably rescaled for the forecast horizon.

[24] In-sample, the Naive model yields a constant variance forecast and standardized returns equal to the raw data, and so are not reported.

**TABLE 3.2**
**Summary Statistics**
**Monthly Data, January 1988 to September 1995**

| Method | Variance Forecasts | | | | | Standardized Returns | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std | Skew | Kurtosis | RMSE | Std | Skew | Kurtosis |
| Raw Data | 12.353 | 20.892 | 2.637 | 6.764 | 1.691 | 1.000 | -0.051 | 0.944 |
| ARCH | 12.276 | 5.479 | 0.687 | -0.340 | 1.583 | 1.007 | -0.293 | 0.566 |
| Option | 13.468 | 11.181 | 3.012 | 12.220 | 1.514 | 1.005 | -0.245 | 0.364 |
| Combining | 13.117 | 10.041 | 2.459 | 7.507 | 1.448 | 1.005 | -0.157 | 0.151 |
| Full MLE | 12.931 | 10.810 | 2.388 | 7.165 | 1.260 | 1.005 | 0.027 | -0.382 |

Looking first at the variance (squared error) forecasts, all the forecasts, daily and monthly, share some characteristics. Each has a mean close to the average squared error – the raw data – (the raw option-implied volatility is biased but our estimate corrects this bias for the most part by estimating an intercept and slope term, $\alpha_0$ and $\phi_{Option,0}$). All models produce forecasts that are, naturally, less volatile than the actual squared error – the raw data. Similarly, the forecasts are generally less skewed and kurtotic than the raw data. The root mean squared error (RMSE column) of each forecast method is a goodness-of-fit metric calculated by standardizing the return innovation with its forecasted standard deviation, squaring this, subtracting one, squaring the result, summing over the sample, and taking the square root. We would expect a more highly parameterized model would have a lower RMSE, at least in-sample, and for the most part this is true. For the Raw Data raw, the RMSE is calculated by taking the return, subtracting its' sample mean value, dividing by the sample standard deviation, squaring this, subtracting one, squaring the result, summing over the sample, and taking the square root.

Of particular interest in Tables 3.1 and 3.2 are results from the columns on Standardized Returns; i.e., $\hat{\epsilon}_t/\hat{\sigma}_t$. A desirable model is one that produces standardized returns with less (excess) kurtosis than the raw data. By this criterion the models with ARCH, volume and options (i.e., the Combining and the Full MLE models) perform best (i.e., deliver the lowest kurtosis), while basic ARCH does worst. Tables 3.1 and 3.2 therefore confirm our conclusion from Table 2 that, by adding volume and implied volatility to a basic ARCH model, we are able to capture important effects that a basic ARCH model alone cannot capture, at least in-sample.

# 5 Forecasts

As important as a model's ability to fit the data in-sample is its ability to forecast future volatility.[25] We therefore investigate our models' forecasting abilities. We find that the interaction of volume with options-implied and ARCH forecasts produces volatility forecasts that encompass both the options-implied and ARCH forecasts themselves, with the cleanest results being for the monthly data, the forecast horizon that the options-implied forecasts are designed for.

To explore out-of-sample forecasting, we first use data from January 4, 1988 through September 29th, 1995, to estimate model parameters. Then we use these parameters with the data from January 1988 through September 1995 to produce for October 2, 1995 (the next business day that markets were open), a one-step-ahead out-of-sample forecast of both the level and volatility of the expected return innovation (i.e., both $\hat{\epsilon}_t$ and $\hat{h}_t$). Next, we update our information set by one period and thus use data from January 4, 1988 through to October 2, 1995 produce for October 3, 1995 a one-step-ahead out-of-sample forecast of both the level and volatility of the expected return innovation. Then we again update, re-estimate, and re-forecast to produce a one-step-ahead out-of-sample forecast for October 4, 1995, then October 5, 1995, and so forth. This process continues until we have obtained one-step-ahead out-of-sample forecasts for October 2, 1995, through August 8, 2003. Tests based on these one-step-ahead out-of-sample forecasts from the various models are reported below.

[25] To measure realized volatility, we employ the squared residual from the Naive model. There are virtually no differences to using the residual from any of the other models.

**TABLE 4.1**
**Summary Statistics**
**Daily Data, October 2, 1995 to August 8, 2003**

| Method | Variance Forecasts | | | | | Standardized Returns | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std | Skew | Kurtosis | RMSE | Std | Skew | Kurtosis |
| Raw Data | 1.737 | 3.609 | 5.986 | 53.673 | 3.916 | 1.000 | 0.003 | 2.324 |
| Naive | 0.917 | 0.177 | 0.315 | -1.184 | 4.006 | 1.359 | -0.069 | 2.505 |
| ARCH | 1.638 | 1.087 | 1.609 | 3.250 | 2.177 | 1.061 | -0.340 | 1.710 |
| Option | 1.372 | 0.814 | 1.638 | 3.214 | 2.098 | 1.099 | -0.194 | 0.977 |
| Combining | 1.397 | 0.856 | 1.664 | 3.283 | 2.118 | 1.097 | -0.189 | 1.063 |
| Full MLE | 1.363 | 0.815 | 1.655 | 3.318 | 2.148 | 1.107 | -0.195 | 1.029 |

**TABLE 4.2**
**Summary Statistics**
**Monthly Data, October 1995 to July 2003**

| Method | Variance Forecasts | | | | | Standardized Returns | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std | Skew | Kurtosis | RMSE | Std | Skew | Kurtosis |
| Raw Data | 25.947 | 32.242 | 2.211 | 5.093 | 2.150 | 1.000 | -0.049 | -0.453 |
| Naive | 15.063 | 2.873 | 0.275 | -1.471 | 2.268 | 1.315 | -0.010 | -0.412 |
| ARCH | 20.892 | 12.747 | 0.952 | 1.046 | 2.000 | 1.217 | 0.013 | -0.206 |
| Option | 23.262 | 13.901 | 1.650 | 3.303 | 1.466 | 1.107 | -0.247 | -0.619 |
| Combining | 22.541 | 14.862 | 1.752 | 4.513 | 1.993 | 1.185 | -0.411 | -0.100 |
| Full MLE | 23.642 | 17.174 | 1.903 | 5.162 | 1.525 | 1.134 | -0.172 | -0.704 |

Tables 4.1 and 4.2 present statistics analogous to those presented above in Tables 3.1 and 3.2, but in addition we now include forecasts from the Naive model. Most of the models manage to produce out-of-sample normalized returns that are less kurtotic than the raw data for the daily data forecast horizon with the exception of the Naive model. For the monthly forecast horizon there is no excess kurtosis, even for the raw data. All the models produce one-step-ahead forecasts which are biased downward, so that the normalized returns are too volatile. For both daily and monthly horizons, the option-implied forecasts yield the smallest mean squared (forecast) error. We would hope to see the RMSE of the more complex forecasting techniques to be much lower than that of the Naive model on the out-of-sample data, and this is true, with the RMSE of the Naive model nearly twice that of all the models on daily data, and as much as 50% larger than the best model on monthly data.

To more carefully evaluate the out-of-sample forecasting performance of each model relative to the other models, we conduct forecast encompassing tests similar to Chong and Hendry (1986), Fair and Shiller [1990] and Day and Lewis [1992,1993]. To examine forecast encompassing we estimate a single bivariate regression with both forecasts as regressors and test for the significance of the parameter estimates in the regression in Equation (10),

$$\hat{\epsilon}_t^2 = \alpha + \beta_j \hat{h}_{t,j} + \beta_i \hat{h}_{t,i} + \nu_t \tag{10}$$

in which ($\hat{\epsilon}_t^2$) is the Naive Model's forecast error,[26] $\hat{h}_{t,j}$ is Model $j$'s forecast, $\hat{h}_{t,i}$ is Model $i$'s forecast, and $\nu$ is a random error. Multicollinearity can lead to both $\beta$ coefficients being insignificant when Equation (10) is estimated, while sufficiently non-overlapping information sets can lead to both estimated $\beta$ coefficients being significant.

The estimation of Equation (10) parameters and standard errors reported below in Tables 5.1 and 5.2 was based on Hansen's (1982) Generalized Method of Moments (GMM) and Newey and West (1987) heteroskedasticity and autocorrelation consistent (HAC) covariances, employing 1 lag of regressors as instruments and 22 lags for the construction of the HAC standard errors in the case of daily data (to correct for the overlapping nature of the option-implied forecast errors). The $\mathrm{R}^2$ reported below are based on a simple OLS regression.

The first block of results in the tables below are single regressor models for which the out-of-sample squared residuals are regressed one-at-a-time on each of the models we consider. The remaining blocks of results detail pairwise regressions described by Equation (10).

The daily and monthly out-of-sample regressions contained in the first panel of Tables 5.1 and 5.2 (the univariate results), provide very strong evidence that the Naive model is inadequate compared to all the other models, with a very large jump in out-of-sample $R^2$ moving from the single regressor Naive forecast regression to any of the other forecast regressions (2.2% $R^2$ in Table 5.1 for the Naive Model, first row of results, to 12.1% $R^2$ for the Combining model, with even more dramatic improvements for monthly forecasting, Table 5.2). There is a similar, though less extreme advantage of the remaining models over

[26] Results are not sensitive to this choice.

the simple ARCH model in terms of $R^2$s or total out-of-sample explanatory power. All the forecasts are individually statistically significant in the case of the daily data forecast horizon,[27] but only the Combining and Full MLE forecasts are individually significant at the monthly horizon.

[27] There is also evidence that most of the forecasts are biased, with slope coefficients significantly different than one and intercepts different from 0, (by inspection, not reported) echoing the results of research such as Day and Lewis [1993]. As the relative efficiency of different forecast methods is our focus, this is issue will not be pursued here.

**Table 5.1**
**Daily Data**
**Out-of-Sample Regression Results**

| Regressors ↓ | Parameter Estimates (Robust Std Error) Naive | ARCH | Options | Combining Model | Full MLE | $R^2$ |
|---|---|---|---|---|---|---|
| Naive | 3.15*** ( .432 ) | | | | | .022 |
| ARCH | | .870*** ( .120 ) | | | | .068 |
| Options | | | 1.45*** ( .214 ) | | | .119 |
| Combining | | | | 1.47*** ( .207 ) | | .121 |
| Full MLE | | | | | 1.51*** ( .219 ) | .116 |
| Naive and ARCH | .488 ( .485 ) | .828*** ( .132 ) | | | | .068 |
| Naive and Options | -.44 ( .625 ) | | 1.56*** ( .238 ) | | | .120 |
| Naive and Combining | -.86 ( .691 ) | | | 1.66*** ( .251 ) | | .122 |
| Naive and Full MLE | -.52 ( .648 ) | | | | 1.64*** ( .258 ) | .117 |
| ARCH and Options | | -.14 ( .215 ) | 1.65*** ( .388 ) | | | .120 |
| ARCH and Combining | | -.11 ( .238 ) | | 1.50*** ( .420 ) | | .125 |
| ARCH and Full MLE | | -.10 ( .202 ) | | | 1.65*** ( .391 ) | .116 |
| Options and Combining | | | .772 ( .621 ) | .637 ( .620 ) | | .121 |
| Options and Full MLE | | | 2.84** ( 1.26 ) | | -1.5 ( 1.25 ) | .120 |
| Combining and Full MLE | | | | 1.45** ( .730 ) | .023 ( .701 ) | .121 |

* Significant at the 10% level, two-sided test.
** Significant at the 5% level, two-sided test.
*** Significant at the 1% level, two-sided test.

Regression coefficients, standard errors and P-values are based on GMM estimation and HAC covariances. The $R^2$ is based on a simple OLS regression.

**Table 5.2**
**Monthly Data**
**Out-of-Sample Regression Results**

| Regressors ↓ | Parameter Estimates (Robust Std Error) Naive | ARCH | Options | Combining Model | Full MLE | $R^2$ |
|---|---|---|---|---|---|---|
| Naive | 1.00 | | | | | .022 |
| | ( .952 ) | | | | | |
| ARCH | | .387 | | | | .053 |
| | | ( .287 ) | | | | |
| Options | | | .539 | | | .071 |
| | | | ( .333 ) | | | |
| Combining | | | | .673** | | .097 |
| | | | | ( .338 ) | | |
| Full MLE | | | | | .661** | .137 |
| | | | | | ( .290 ) | |
| Naive | .051 | .347 | | | | .055 |
| and ARCH | ( 1.48 ) | ( .425 ) | | | | |
| Naive | -.58 | | .714 | | | .071 |
| and Options | ( 1.55 ) | | ( .461 ) | | | |
| Naive | -.67 | | | .781* | | .097 |
| and Combining | ( 1.49 ) | | | ( .438 ) | | |
| Naive | -.27 | | | | .722** | .137 |
| and Full MLE | ( 1.23 ) | | | | ( .340 ) | |
| ARCH | | .074 | .526 | | | .081 |
| and Options | | ( .399 ) | ( .477 ) | | | |
| ARCH | | .004 | | .599 | | .101 |
| and Combining | | ( .451 ) | | ( .502 ) | | |
| ARCH | | .029 | | | .654* | .145 |
| and Full MLE | | ( .323 ) | | | ( .356 ) | |
| Options | | | -.09 | .531 | | .099 |
| and Combining | | | ( .478 ) | ( .581 ) | | |
| Options | | | -.72 | | 1.20** | .158 |
| and Full MLE | | | ( .527 ) | | ( .530 ) | |
| Combining | | | | -.15 | .747* | .139 |
| and Full MLE | | | | ( .395 ) | ( .401 ) | |

* Significant at the 10% level, two-sided test.
** Significant at the 5% level, two-sided test.
*** Significant at the 1% level, two-sided test.

Regression coefficients, standard errors and P-values are based on GMM estimation and HAC covariances. The $R^2$ is based on a simple OLS regression.

When we look at the first set of bivariate regressions, the second panel of Tables 5.1 and 5.2, we see that the addition of the Naive model forecast (the second panel in each of Tables 5.1 and 5.2) to any of the other methods' forecasts does little to increase the explanatory power of the regression, at both the daily and monthly forecast horizons. Also, the $\beta$ coefficient on the Naive forecast is in each case greatly reduced in magnitude, from over 3 to below 1 in absolute magnitude, is negative when included with all but the ARCH forecast, and is statistically insignificant in all cases. In the bivariate regressions of the second panel of Tables 5.1 and 5.2 all but the Naive forecast have statistically significant regression coefficients in the case of the daily data forecast horizon, but only the Combining and Full MLE forecasts are significant at the monthly horizon. Altogether these results suggest that the Naive model is easily encompassed by the remaining models, which is consistent with volatility being predictable.

The second set of bivariate regressions, in the third panel of Tables 5.1 and 5.2, permit us to explore the incremental value of the time series ARCH forecast out-of-sample. For both the daily and monthly horizons, adding ARCH to any of the remaining forecasts, Options, Combining or Full MLE, increases the $R^2$ very little relative to not including ARCH. Most telling, the coefficient estimate on the ARCH forecast is not only statistically insignificant (which it might be simply because of multicollinearity) but also very near 0 and greatly reduced in magnitude from the regression with ARCH only. These results suggest strongly that simple univariate time series information has little to add relative to the information embedded in options prices, even when we look at a daily forecast horizon with its concurrent mismatch of forecast horizon relative to the option-implied volatility forecast calibrated for a monthly horizon. Altogether these results point to the ARCH model being encompassed by the Options, Combining and Full MLE models.

The third set of bivariate regressions, in the fourth panel Table 5.1, reveal at the daily horizon that there is too much multicollinearity between the Options forecast of volatility and the Combining and Full MLE models (models that incorporate the option-implied volatility with times series information) to discriminate between them. The leveraged coefficients we see when Options and Full MLE are included together (Options having a coefficient of 2.84 and Full MLE of -1.5) are not particularly meaningful, in our opinion, as virtually no increase in the $R^2$ is associated with this pairing relative to either forecast used individually. Similarly in the fifth panel of Table 5.1, the daily data horizon case, the paring of the Combining model and the Full MLE barely budges the $R^2$ relative to either forecast individually. Overall, this suggests that there may be no value to the volume interaction variable and ARCH time series forecast when using daily data, even though in-sample the volume interaction terms appear significant. It may also be true that we do not have a long enough sample of daily out-of-sample data to uncover a significant impact from volume and ARCH.

The monthly data horizon is more definitive for determining the incremental value of the new time series information we introduce in this paper, the volume data, relative to information embedded in the option-implied volatility forecast. Consider the third set of bivariate regressions, the fourth panel of Table 5.2. Now we find that adding either the Combining forecast or the Full MLE forecast to the Options forecast both drives down the magnitude of the Options forecast coefficient estimate and drives out the coefficient estimates' statistical significance. As well, the $R^2$ of the models using volume information, Combining and Full MLE, is a great deal more than the remaining models with use only univariate time series information or option-implied volatility. The Full MLE model, dynamically estimating the univariate time series coefficients, the volume interaction coefficients and the option-implied volatility coefficient, demonstrates a large advantage out-of-sample relative to all the rest of the models, being the only model which has a statistically significant coefficient regardless of what other forecast is added to it, and the only other model that does not have its coefficient driven down to 0 or negative values by any other forecast.

# 6 Summary and Conclusions

Previous studies have reported that trading volume cannot forecast volatility directly. In this paper we uncover a new result: that volume does indeed have predictive power for forecasting volatility, with volume playing the role of a switching variable between states in which option-implied volatility is more or less informative than ARCH for volatility forecasting, perhaps reflecting important changes in information attendant with increases in trading volume on the NYSE. We find that the accuracy of volatility forecasts can be significantly improved by accounting for the volume effect and by combining information from ARCH models and option prices accordingly. This finding is made possible because of the novel way we incorporate trading volume into our functional forms and because, while previous papers have added either trading volume or option-implied volatility (but not both) to ARCH models, our study is the first we know of to consider all three factors together.

Results produced by our investigation reveal that if trading volume was "lower than normal" during period $t-1$ then the best forecast of time $t$ volatility is found by combining the ARCH forecast with the option-implied volatility forecast, with similar weight being given to ARCH than to options. Conversely, if trading volume was "higher than normal" during period $t-1$, then the best forecast of time $t$ volatility is obtained by placing more weight on options and less on ARCH. This result is robust to a variety of perturbations of the in-sample period and model specification and seems to be largely driven by improvements in the quality of option-implied volatility forecasts relative to ARCH forecasts during high volume periods.

Results from the combining exercise in Section 3 reveal that option prices appear to contain better information about future volatility on high volume days than on low volume days relative to ARCH forecasts. This suggests that market prices contain more information relative to historical sources in high volume periods than in low volume periods (indeed, our work suggests a new way to test the relative informativeness of market prices in various volume regimes). Our results also suggest either that option markets are more efficient in high volume periods – that prices encompass historical information only in high volume periods – or else that option pricing models are less misspecified in high volume periods.

Results from the various ARCH tests in Section 4 further reveal that adding option and volume information to ARCH models greatly improves the ARCH models' forecasting performance, though most markedly at a monthly rather than a daily forecast horizon. This suggests that ARCH modelers may profit from expanding the traditional ARCH information set to include volume, options, and other types of information in addition to the history of lagged return innovations.

## Appendix 1

### Model Selection: GARCH Models

Parameter estimates and summary statistics for models of the form:

$$R_t = \mu + \epsilon_t \quad ; \quad \epsilon_t \sim (0, h_t^2) \tag{11}$$

$$h_t^2 = \alpha + \sum_{i=1}^{2} \left(\beta_i h_{t-i}^2\right) + \sum_{i=1}^{2} \left(\gamma_i \epsilon_{t-i}^2\right) + \sum_{i=1}^{2} \left(\delta_i D_{t-i} \epsilon_{t-i}^2\right) \tag{12}$$

$$D_{t-i} = \begin{cases} 1 & if\ \epsilon_{t-i} < 0 \\ 0 & otherwise \end{cases} \tag{13}$$

are found in Panel A for Tables A1.1 to A1.4 below. A model with $p$ lags of $h_t$, $q$ lags of $\epsilon_t^2$ and $r$ lags of $D_t\epsilon_t^2$ is labeled GJR(p,q,r). Such a model excluding the asymmetric volatility term $D_t\epsilon_t^2$ is labeled GARCH(p,q). These models were estimated on the period from January 4, 1988 to the end of September 1995. Presented in Panel B of the tables are the model Log Likelihood, Schwarz information criteria (BIC), and tests for residual autocorrelation (Wald AR), ARCH and Sign-ARCH. The test for autocorrelation is a $\chi^2$ Wald test on 5 (3) lags of the daily (monthly) data residuals in the mean equation. The test for ARCH is a standard LM test, having us regress the current squared standardized residual on 24 (3) lags of the daily (monthly) data squared standardized residual, collecting the regression $R^2$ and forming $z = nR^2$, where $n$ is the number of observations, and $z$ is distributed $\chi^2$ with 24 (3) degrees of freedom. The Sign-ARCH tests were proposed by Engle and Ng [1993] and test for stronger (asymmetric) volatility increases from negative shocks to the returns process. Our Sign-ARCH tests on daily data used 5 lags, 3 lags for monthly data.

The $R^2$ for all the models below are 0 as the mean regression equation in each model contains only a constant, and so this statistic is not included. Of the models that remove significant evidence (at the 5% level) of time series dependence, including residual autocorrelation, ARCH and Sign-ARCH, the GARCH(1,1) specification has the best BIC as well as being the most simple model. The GJR(0,1,2) and GJR(0,1,1) models do very nearly as well on all criteria we investigate failing only one specification test at the 5% level, and slightly better by the BIC, but we will favor the more familiar GARCH(1,1) model.

**Table A1.1**

**Daily Data, In-Sample 1988:1-1995:9**

**Estimation Results and Diagnostics for**

**Volatility Models based on Equations (11)-(13)**

| | Panel A: Parameter Estimates (Robust Standard Error) | | | | | |
|---|---|---|---|---|---|---|
| Parameter | MLE GARCH (1,0) | MLE GARCH (2,0) | MLE GARCH (1,1) | MLE GARCH (2,1) | MLE GARCH (1,2) | MLE GARCH (2,2) |
| $\mu$ | 0.041** | 0.044** | 0.047*** | 0.047*** | 0.047*** | 0.047*** |
| | ( 0.019) | ( 0.019) | ( 0.016) | ( 0.017) | ( 0.017) | ( 0.016) |
| $\alpha$ | 0.628*** | 0.610*** | 0.002 | 0.002 | 0.003 | 0.004 |
| | ( 0.047) | ( 0.054) | ( 0.001) | ( 0.001) | ( 0.003) | ( 0.003) |
| $\beta_1$ | | | 0.980*** | 0.982*** | 0.527 | 0.010 |
| | | | ( 0.008) | ( 0.008) | ( 1.168) | ( 0.101) |
| $\beta_2$ | | | | | 0.444 | 0.949*** |
| | | | | | ( 1.147) | ( 0.100) |
| $\gamma_1$ | 0.141*** | 0.132*** | 0.016** | 0.036 | 0.023 | 0.013 |
| | ( 0.044) | ( 0.044) | ( 0.007) | ( 0.034) | ( 0.019) | ( 0.012) |
| $\gamma_2$ | | 0.033 | | -.021 | | 0.020** |
| | | ( 0.026) | | ( 0.034) | | ( 0.010) |
| | Panel B: Diagnostics | | | | | |
| Log Likelihood | -2440.90 | -2439.30 | -2311.62 | -2310.45 | -2310.94 | -2311.22 |
| BIC | 4912.125 | 4916.499 | 4661.135 | 4666.378 | 4667.369 | 4675.501 |
| LM ARCH | 0.000 | 0.000 | 0.980 | 0.998 | 0.994 | 0.959 |
| Wald AR | 0.057 | 0.015 | 0.124 | 0.133 | 0.130 | 0.096 |
| Sign Test | 0.583 | 0.566 | 0.521 | 0.470 | 0.539 | 0.610 |
| Negative Sign Test | 0.001 | 0.001 | 0.445 | 0.293 | 0.629 | 0.451 |
| Positive Sign Test | 0.000 | 0.000 | 0.058 | 0.040 | 0.068 | 0.223 |
| Joint Sign Test | 0.000 | 0.000 | 0.223 | 0.242 | 0.304 | 0.446 |

* Significant at the 10% level, two-sided test.

** Significant at the 5% level, two-sided test.

*** Significant at the 1% level, two-sided test.

**Table A1.2**

**Daily Data, In-Sample 1988:1-1995:9**

**Estimation Results and Diagnostics for**

**Volatility Models based on Equations (11)-(13)**

| | Panel A: Parameter Estimates (Robust Standard Error) | | | | | |
|---|---|---|---|---|---|---|
| Parameter | MLE GJR (1,0,1) | MLE GJR (2,0,1) | MLE GJR (1,1,1) | MLE GJR (2,1,1) | MLE GJR (1,2,1) | MLE GJR (2,2,1) |
| $\mu$ | 0.039** | 0.042** | 0.043** | 0.044*** | 0.043** | 0.041** |
| | ( 0.019) | ( 0.019) | ( 0.017) | ( 0.017) | ( 0.017) | ( 0.017) |
| $\alpha$ | 0.631*** | 0.613*** | 0.002 | 0.002 | 0.002 | 0.003* |
| | ( 0.048) | ( 0.055) | ( 0.001) | ( 0.001) | ( 0.002) | ( 0.002) |
| $\beta_1$ | | | 0.983*** | 0.984*** | 0.511 | 0.108 |
| | | | ( 0.008) | ( 0.008) | ( 0.999) | ( 0.170) |
| $\beta_2$ | | | | | 0.465 | 0.859*** |
| | | | | | ( 0.985) | ( 0.171) |
| $\gamma_1$ | 0.108* | 0.096 | 0.008 | 0.022 | 0.010 | -.013 |
| | ( 0.059) | ( 0.059) | ( 0.012) | ( 0.030) | ( 0.016) | ( 0.016) |
| $\gamma_2$ | | 0.033 | | -.014 | | 0.024 |
| | | ( 0.027) | | ( 0.033) | | ( 0.019) |
| $\delta_1$ | 0.058 | 0.063 | 0.012 | 0.010 | 0.018 | 0.030 |
| | ( 0.089) | ( 0.088) | ( 0.014) | ( 0.014) | ( 0.026) | ( 0.022) |
| | Panel B: Diagnostics | | | | | |
| Log Likelihood | -2440.45 | -2438.75 | -2309.62 | -2309.14 | -2308.83 | -2307.49 |
| BIC | 4918.802 | 4922.974 | 4664.720 | 4671.332 | 4670.722 | 4675.621 |
| LM ARCH | 0.000 | 0.000 | 0.994 | 0.998 | 0.998 | 0.989 |
| Wald AR | 0.055 | 0.015 | 0.156 | 0.152 | 0.163 | 0.208 |
| Sign Test | 0.605 | 0.619 | 0.412 | 0.386 | 0.428 | 0.564 |
| Negative Sign Test | 0.001 | 0.006 | 0.609 | 0.380 | 0.720 | 0.689 |
| Positive Sign Test | 0.000 | 0.000 | 0.061 | 0.044 | 0.064 | 0.366 |
| Joint Sign Test | 0.000 | 0.000 | 0.299 | 0.283 | 0.353 | 0.686 |

* Significant at the 10% level, two-sided test.

** Significant at the 5% level, two-sided test.

*** Significant at the 1% level, two-sided test.

**Table A1.3**

**Daily Data, In-Sample 1988:1-1995:9**

**Estimation Results and Diagnostics for**

**Volatility Models based on Equations (11)-(13)**

| | Panel A: Parameter Estimates (Robust Standard Error) | | | | | |
|---|---|---|---|---|---|---|
| Parameter | MLE GJR (1,0,2) | MLE GJR (2,0,2) | MLE GJR (1,1,2) | MLE GJR (2,1,2) | MLE GJR (1,2,2) | MLE GJR (2,2,2) |
| $\mu$ | 0.037** | 0.037* | 0.043** | 0.043** | 0.044*** | 0.041** |
| | ( 0.019) | ( 0.019) | ( 0.017) | ( 0.017) | ( 0.000) | ( 0.017) |
| $\alpha$ | 0.599*** | 0.604*** | 0.001 | 0.001 | 0.000 | 0.003 |
| | ( 0.051) | ( 0.055) | ( 0.001) | ( 0.001) | ( 0.000) | ( 0.002) |
| $\beta_1$ | | | 0.985*** | 0.984*** | 1.966*** | 0.212 |
| | | | ( 0.007) | ( 0.007) | ( 0.000) | ( 0.145) |
| $\beta_2$ | | | | | -.965*** | 0.760*** |
| | | | | | ( 0.000) | ( 0.143) |
| $\gamma_1$ | 0.077 | 0.075 | 0.008 | -.038* | 0.000 | -.037** |
| | ( 0.053) | ( 0.052) | ( 0.012) | ( 0.022) | ( 0.000) | ( 0.017) |
| $\gamma_2$ | | -.011 | | 0.046* | | 0.049* |
| | | ( 0.020) | | ( 0.026) | | ( 0.027) |
| $\delta_1$ | 0.079 | 0.081 | 0.087 | 0.125** | 0.034*** | 0.063** |
| | ( 0.083) | ( 0.082) | ( 0.055) | ( 0.054) | ( 0.000) | ( 0.025) |
| $\delta_2$ | 0.141*** | 0.149*** | -.080 | -.116** | -.034*** | -.042 |
| | ( 0.055) | ( 0.054) | ( 0.053) | ( 0.055) | ( 0.000) | ( 0.030) |
| | Panel B: Diagnostics | | | | | |
| Log Likelihood | -2431.77 | -2431.50 | -2305.13 | -2302.01 | -2289.17 | -2305.36 |
| BIC | 4909.016 | 4916.068 | 4663.321 | 4664.670 | 4638.985 | 4678.933 |
| LM ARCH | 0.000 | 0.000 | 0.999 | 0.999 | 0.969 | 0.995 |
| Wald AR | 0.010 | 0.011 | 0.145 | 0.166 | 0.000 | 0.133 |
| Sign Test | 0.542 | 0.577 | 0.423 | 0.487 | 0.328 | 0.620 |
| Negative Sign Test | 0.006 | 0.006 | 0.658 | 0.537 | 0.627 | 0.823 |
| Positive Sign Test | 0.000 | 0.000 | 0.042 | 0.069 | 0.033 | 0.633 |
| Joint Sign Test | 0.000 | 0.000 | 0.309 | 0.488 | 0.284 | 0.869 |

* Significant at the 10% level, two-sided test.

** Significant at the 5% level, two-sided test.

*** Significant at the 1% level, two-sided test.

**Table A1.4**

**Daily Data, In-Sample 1988:1-1995:9**

**Estimation Results and Diagnostics for**

**Volatility Models based on Equations (11)-(13)**

| | Panel A: Parameter Estimates (Robust Standard Error) | | | | | |
|---|---|---|---|---|---|---|
| Parameter | MLE GJR (0,0,1) | MLE GJR (0,1,1) | MLE GJR (0,2,1) | MLE GJR (0,0,2) | MLE GJR (0,1,2) | MLE GJR (0,2,2) |
| $\mu$ | 0.042$^{**}$ | 0.042$^{**}$ | 0.042$^{**}$ | 0.039$^{**}$ | 0.042$^{**}$ | 0.043$^{**}$ |
| | ( 0.019) | ( 0.017) | ( 0.017) | ( 0.019) | ( 0.017) | ( 0.017) |
| $\alpha$ | 0.673$^{***}$ | 0.001 | 0.002 | 0.628$^{***}$ | 0.001 | 0.003$^{*}$ |
| | ( 0.052) | ( 0.001) | ( 0.001) | ( 0.056) | ( 0.001) | ( 0.002) |
| $\beta_1$ | | 0.988$^{***}$ | 0.448 | | 0.990$^{***}$ | -.004 |
| | | ( 0.004) | ( 0.780) | | ( 0.004) | ( 0.015) |
| $\beta_2$ | | | 0.535 | | | 0.977$^{***}$ |
| | | | ( 0.772) | | | ( 0.012) |
| $\delta_1$ | 0.138$^{**}$ | 0.018$^{**}$ | 0.027 | 0.134$^{**}$ | 0.088$^{*}$ | 0.016 |
| | ( 0.063) | ( 0.007) | ( 0.017) | ( 0.058) | ( 0.049) | ( 0.011) |
| $\delta_2$ | | | | 0.146$^{***}$ | -.073 | 0.025$^{***}$ |
| | | | | ( 0.056) | ( 0.050) | ( 0.009) |
| | Panel B: Diagnostics | | | | | |
| Log Likelihood | -2445.54 | -2311.06 | -2310.00 | -2434.96 | -2306.74 | -2308.75 |
| BIC | 4921.409 | 4660.028 | 4665.489 | 4907.815 | 4658.966 | 4670.569 |
| LM ARCH | 0.000 | 0.991 | 0.997 | 0.000 | 0.998 | 0.977 |
| Wald AR | 0.174 | 0.283 | 0.284 | 0.127 | 0.250 | 0.343 |
| Sign Test | 0.532 | 0.355 | 0.359 | 0.496 | 0.345 | 0.350 |
| Negative Sign Test | 0.001 | 0.731 | 0.759 | 0.006 | 0.710 | 0.568 |
| Positive Sign Test | 0.000 | 0.037 | 0.034 | 0.000 | 0.017 | 0.026 |
| Joint Sign Test | 0.000 | 0.243 | 0.286 | 0.000 | 0.214 | 0.175 |

## Appendix 2

## Additional Daily and Monthly Estimation Results

Parameter estimates and summary statistics for models based on Equations (5), and (7)-(9) are found in Panel A for Tables A2.1 to A2.3 below. Presented in Panel B of the tables are the model Log Likelihood, Schwarz information criteria (BIC), and tests for residual autocorrelation (Wald AR), ARCH and Sign-ARCH. The test for autocorrelation is a $\chi^2$ Wald test on 5 (3) lags of the daily (monthly) data residuals in the mean equation. The test for ARCH is a standard LM test, having us regress the current squared standardized residual on 24 (3) lags of the daily (monthly) data squared standardized residual, collecting the regression $R^2$ and forming $z = nR^2$, where $n$ is the number of observations, and $z$ is distributed $\chi^2$ with 24 (3) degrees of freedom. The Sign-ARCH tests were proposed by Engle and Ng [1993] and test for stronger (asymmetric) volatility increases from negative shocks to the returns process. Our Sign-ARCH tests on daily data used 5 lags, 3 lags for monthly data.

The $R^2$ for all the models below are 0 as the mean regression equation in each model contains only a constant, and so this statistic is not included.

**TABLE A2.1**

**Monthly Data, In-Sample 1988:1-1995:9**

**Estimation Results and Diagnostics for**

**Volatility Models based on Equations (5), (7)-(9)**

| | **Panel A: Parameter Estimates** (Robust Standard Error) | | | | |
|---|---|---|---|---|---|
| Column → | 1 | 2 | 3 | 4 | 5 |
| Parameter Name ↓ | Naive | Options Only | ARCH Model | Combining Model | Full MLE |
| $\mu$ | 0.992*** | 0.849*** | 0.921*** | 0.918*** | 1.137*** |
| | ( 0.364) | ( 0.316) | ( 0.328) | ( 0.312) | ( 0.290) |
| $\alpha_0$ | 12.35*** | -.090 | 0.137 | -1.16 | -1.11 |
| | ( 2.155) | ( 2.196) | ( 0.372) | ( 3.005) | ( 1.339) |
| $\beta_0$ | | | 0.947*** | | 0.408* |
| | | | ( 0.047) | | ( 0.222) |
| $\beta_1$ | | | | | -1.18*** |
| | | | | | ( 0.381) |
| $\gamma$ | | | 0.025 | | -.111* |
| | | | ( 0.016) | | ( 0.060) |
| $\phi_{Option,0}$ | | 0.423*** | | 0.123 | 0.260** |
| | | ( 0.136) | | ( 0.163) | ( 0.108) |
| $\phi_{Option,1}$ | | | | 0.356 | 0.635*** |
| | | | | ( 0.354) | ( 0.219) |
| $\phi_{ARCH,0}$ | | | | 0.630 | |
| | | | | ( 0.584) | |
| $\phi_{ARCH,1}$ | | | | -.534 | |
| | | | | ( 0.774) | |
| | **Panel B: Diagnostics** | | | | |
| Log Likelihood | -248.856 | -242.903 | -244.158 | -242.080 | -239.223 |
| BIC | 511.310 | 503.936 | 510.980 | 515.888 | 514.707 |
| Wald AR Test | 0.066 | 0.275 | 0.428 | 0.039 | 1.000 |
| LM ARCH Test | 0.114 | 0.129 | 0.186 | 0.090 | 0.185 |
| Sign Test | 0.257 | 0.888 | 0.796 | 0.800 | 0.619 |
| Neg. Sign Test | 0.145 | 0.754 | 0.951 | 0.778 | 0.587 |
| Pos. Sign Test | 0.190 | 0.511 | 0.478 | 0.544 | 0.401 |
| Joint Sign Test | 0.167 | 0.863 | 0.877 | 0.913 | 0.902 |

* Significant at the 10% level, two-sided test.

** Significant at the 5% level, two-sided test.

*** Significant at the 1% level, two-sided test.

## TABLE A2.2

## Monthly Data, Full Sample 1988:1-2003:8

## Estimation Results and Diagnostics for

## Volatility Models based on Equations (5), (7)-(9)

| | **Panel A: Parameter Estimates** (Robust Standard Error) | | | | |
|---|---|---|---|---|---|
| Column → | 1 | 2 | 3 | 4 | 5 |
| Parameter Name ↓ | Naive | Options Only | ARCH Model | Combining Model | Full MLE |
| $\mu$ | 0.884*** | 0.842*** | 1.001*** | 0.848*** | 0.936*** |
| | ( 0.319) | ( 0.264) | ( 0.259) | ( 0.263) | ( 0.251) |
| $\alpha_0$ | 19.13*** | -1.20 | 0.648 | -1.03 | -1.17 |
| | ( 2.018) | ( 1.561) | ( 0.528) | ( 1.736) | ( 0.877) |
| $\beta_0$ | | | 0.831*** | | 0.430** |
| | | | ( 0.075) | | ( 0.189) |
| $\beta_1$ | | | | | -1.12*** |
| | | | | | ( 0.286) |
| $\gamma$ | | | 0.129** | | -.084 |
| | | | ( 0.056) | | ( 0.053) |
| $\phi_{Option,0}$ | | 0.525*** | | 0.105 | 0.273*** |
| | | ( 0.079) | | ( 0.132) | ( 0.096) |
| $\phi_{Option,1}$ | | | | 0.416* | 0.655*** |
| | | | | ( 0.246) | ( 0.175) |
| $\phi_{ARCH,0}$ | | | | 0.700** | |
| | | | | ( 0.330) | |
| $\phi_{ARCH,1}$ | | | | -.659 | |
| | | | | ( 0.465) | |
| | **Panel B: Diagnostics** | | | | |
| Log Likelihood | -544.186 | -529.191 | -533.230 | -528.104 | -524.272 |
| BIC | 1104.081 | 1079.328 | 1092.643 | 1092.863 | 1090.435 |
| Wald AR Test | 0.692 | 0.380 | 0.941 | 0.429 | 0.712 |
| LM ARCH Test | 0.000 | 0.972 | 0.450 | 0.943 | 0.734 |
| Sign Test | 0.001 | 0.342 | 0.259 | 0.342 | 0.111 |
| Neg. Sign Test | 0.000 | 0.779 | 0.798 | 0.700 | 0.376 |
| Pos. Sign Test | 0.619 | 0.567 | 0.907 | 0.426 | 0.424 |
| Joint Sign Test | 0.000 | 0.844 | 0.554 | 0.805 | 0.432 |

* Significant at the 10% level, two-sided test.

** Significant at the 5% level, two-sided test.

*** Significant at the 1% level, two-sided test.

**TABLE A2.3**

**Daily Data, Full Sample 1988:1-2003:8**

**Estimation Results and Diagnostics for**

**Volatility Models based on Equations (5), (7)-(9)**

| | **Panel A: Parameter Estimates** (Robust Standard Error) | | | | |
|---|---|---|---|---|---|
| Column → | 1 | 2 | 3 | 4 | 5 |
| Parameter Name ↓ | Naive | Options Only | ARCH Model | Combining Model | Full MLE |
| $\mu$ | 0.042** | 0.036*** | 0.056*** | 0.038*** | 0.037*** |
| | ( 0.018) | ( 0.014) | ( 0.014) | ( 0.014) | ( 0.014) |
| $\alpha_0$ | 1.232*** | -.109*** | 0.006*** | -.102** | -.087** |
| | ( 0.048) | ( 0.040) | ( 0.002) | ( 0.042) | ( 0.038) |
| $\beta_0$ | | | 0.951*** | | 0.232 |
| | | | ( 0.008) | | ( 0.283) |
| $\beta_1$ | | | | | -.202 |
| | | | | | ( 0.367) |
| $\gamma$ | | | 0.045*** | | 0.010 |
| | | | ( 0.009) | | ( 0.024) |
| $\phi_{Option,0}$ | | 0.704*** | | 0.436*** | 0.526*** |
| | | ( 0.038) | | ( 0.076) | ( 0.187) |
| $\phi_{Option,1}$ | | | | 0.290** | 0.148 |
| | | | | ( 0.135) | ( 0.234) |
| $\phi_{ARCH,0}$ | | | | 0.375*** | |
| | | | | ( 0.101) | |
| $\phi_{ARCH,1}$ | | | | -.404** | |
| | | | | ( 0.176) | |
| | **Panel B: Diagnostics** | | | | |
| Log Likelihood | -5997.58 | -5447.72 | -5526.98 | -5440.02 | -5447.05 |
| BIC | 12019.99 | 10928.56 | 11095.34 | 10937.98 | 10960.32 |
| Wald AR Test | 0.143 | 0.109 | 0.016 | 0.072 | 0.104 |
| LM ARCH Test | 0.000 | 0.495 | 0.927 | 0.739 | 0.551 |
| Sign Test | 0.000 | 0.003 | 0.013 | 0.001 | 0.008 |
| Neg. Sign Test | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Pos. Sign Test | 0.000 | 0.003 | 0.003 | 0.000 | 0.006 |
| Joint Sign Test | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

* Significant at the 10% level, two-sided test.

** Significant at the 5% level, two-sided test.

*** Significant at the 1% level, two-sided test.

# References

Admati, A. and P. Pfleiderer, 1988, "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies* 1, 3-40.

Bakshi, G.; C. Cao and Z. Chen, 1997, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance* 52, 105-134.

Black, F., and M. S. Scholes, 1973, "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, 637-659.

Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* 31, 307-327.

Bollerslev, T., R.Y. Chou, and K.F. Kroner, 1992, "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52, 5-59.

Bollerslev, T. and J.M. Wooldridge, 1992, "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances," *Econometric Reviews* 11, 143-172.

Brooks, 1998, "Predicting Stock Market Volatility: Can Market Volume Help?" *Journal of Forecasting* 17, 59-80.

Christensen, B.J. and N.R. Prabhala, 1998, "The Relation Between Implied and Realized Volatility," *Journal of Financial Economics* 50, 125-150.

Chong, Y.Y. and D.F. Hendry, 1986, "Econometric Evaluation of Linear Macroeconomic Models," *Review of Economic Studies* 53, 671-90.

Clemens, R.T., 1989, "Combining Forecasts: A Review and Annotated Bibliography," *International Journal of Forecasting* 5, 559-583.

Day, T. and C. Lewis, 1992, "Stock Market Volatility and the Information Content of Index Options," *Journal of Econometrics* 52, 267-287.

Day, T. and C. Lewis, 1993, "Forecasting Futires Market Volatility," *The Journal of Derivatives* 1(2), 33-50.

Donaldson, R.G. and M. Kamstra, 1997, "An Artificial Neural Network-GARCH Model for International Stock Return Volatility," *Journal of Empirical Finance* 4, 17-46.

Donaldson, R.G. and M. Kamstra, 1996, "Forecast Combining with Neural Networks," *Journal of Forecasting* 15, 49-61.

Engle, R.F. and V.K. Ng, 1993, "Measuring and Testing the Impact of News on Volatility," *Journal of Finance* 48, 1749-1778.

Fair, R.C. and R.J. Shiller, 1990, "Comparing Information in Forecasts from Econometric Models," *American Economic Review* 80, 375-389.

Fleming, J., 1998, "The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices," *Journal of Empirical Finance* 5, 317-345.

Glosten, L.R., R. Jagannathan and D.E. Runkle, 1993, "The Relationship Between Expected Value and the Volatility of the Nominal Excess Return On Stocks," *Journal of Finance* 48, 1779-1801.

Hannan, E.J., 1980, "The Estimation of the Order of an ARMA Process" *The Annals of Statistics* 8, 1071-1081.

Hansen, L.P. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 1982, *50*, 1029-1054.

He, H. and J. Wang, 1995, "Differential Information and the Dynamic Behavior of Stock Trading Volume," *Review of Financial Studies* 8, 919-972.

Heimstra, C. and J. Jones, 1994, "Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation," *Journal of Finance* 49, 1639-1664.

Karpoff, J., 1987, "The Relation Between Price Changes and Trading Volume: A Survey," *Journal of Financial and Quantitative Analysis* 22, 109-126.

Lamoureux, C. and W. Lastrapes, 1993, "Forecasting Stock-Return Variance: Toward and Understanding of Stochastic Implied Volatility," *Review of Financial Studies* 6, 293-326.

Lamoureux, C. and W. Lastrapes, 1994, "Endogenous Trading Volume and Momentum in Stock Return Volatility," *Journal of Business and Economic Statistics* 12, 253-260.

Ljung, G.M. and G.E.P. Box, 1978, "On a Measure of Lack of Fit in Time Series Models," *Biometrika* 65, 297-303.

Nelson, D.B., 1991, "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica* 59, 347-370.

Newey, Whitney K. and Kenneth D. West. A Simple, Positive, Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 1987, *55*, 703-708.

Richardson, M. and T. Smith, 1994, "A Direct Test of the Mixture of Distributions Hypothesis: Measuring the Daily Flow of Information," *Journal of Financial and Quantitative Analysis* 29, 101-116.

Whaley, R. E., 2000, "The Investor Fear Gauge", *Journal of Portfolio Management* 26, 12-17.