JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 45, No. 2, Apr. 2010, pp. 503–533

doi:10.1017/S0022109010000116

Forecasting Volatility Using Long Memory and Comovements: An Application to Option Valuation under SFAS 123R

George J. Jiang and Yisong S. Tian*

Abstract

Horizon-matched historical volatility is commonly used to forecast future volatility for option valuation under the Statement of Financial Accounting Standards (SFAS) 123R. In this paper, we empirically investigate the performance of using historical volatility to forecast long-term stock return volatility in comparison with a number of alternative forecasting methods. In analyzing forecasting errors and their impact on reported income due to option expensing, we find that historical volatility is a poor forecast for long-term volatility and that shrinkage adjustment toward comparable-firm volatility only slightly improves its performance. Forecasting performance can be improved substantially by incorporating both long memory and comovements with common market factors. We also experiment with a simple mixed-horizon realized volatility model and find its long-term forecasting performance to be more accurate than historical forecasts but less accurate than long-memory forecasts.

I. Introduction

As a risk measure, volatility is an important concept in finance and plays a critical role in a wide range of applications such as asset pricing, portfolio management, risk management, and option valuation. Not surprisingly, a large body of research has been devoted to the modeling, estimation, and forecasting of stock return volatility. Substantial progress has been made in our understanding of the time-series properties of stock return volatility (e.g., clustering and persistence),

*Jiang, gjiang@eller.arizona.edu, Eller College of Management, University of Arizona, PO Box 210108, Tucson, AZ 85721; and Tian, ytian@ssb.yorku.ca, Schulich School of Business, York University, 4700 Keele St., Toronto, ON, M3J 1P3, Canada. We thank Melanie Cao, Dan Dhaliwal, Mark Kamstra, Chris Lamoureux, Guanzhong Pan, Tong Yao; conference participants at the annual meetings of the 2006 Northern Finance Association, the 2007 Financial Management Association, and the 2008 European Financial Management Association; and seminar participants at York University for helpful comments and suggestions. We are particularly grateful to the guidance and comments provided by Torben Andersen (the referee) and Hendrik Bessembinder (the editor). The financial support of the Schulich School of Business and Social Sciences and Humanity Research Council of Canada is gratefully acknowledged.

and a number of advanced models have been developed to characterize volatility time series. Nevertheless, the forecasting performance of these volatility models are mostly assessed and investigated over relatively short horizons (e.g., over 1-day to 1-month horizons).[1] In contrast, little attention has been paid to evaluating the forecasting performance of competing volatility models over longer horizons (e.g., over 1- to 5-year horizons).[2] In this paper, we fill this gap in the literature by conducting a large-scale study of long-horizon forecasting performance of competing volatility models. We identify two important features in volatility time series, *long memory* and *comovements* across the market, that are particularly helpful in improving long-term forecasting performance.

Our study is motivated by recent changes in the accounting treatment of stock options. U.S. companies are now required by the Financial Accounting Standards Board (FASB) to recognize, as opposed to disclosing it in footnotes to financial statements, the full cost of stock options granted to their top executives and other employees. Since the FASB (see Statement of Financial Accounting Standards (SFAS) 123R) recognizes horizon-matched historical volatility as an acceptable volatility forecast for option expensing purposes, many firms have adopted this simple forecasting method (e.g., Balsam, Mozes, and Newman (2003), Bartov, Mohanram, and Nissim (2004), Aboody, Barth, and Kasznik (2006), Johnston (2006), and Hodder, Mayew, McAnally, and Weaver (2006)).

In this paper, we evaluate the performance of using historical volatility to forecast long-horizon stock return volatility in comparison with alternative volatility forecasting methods. We argue that long-horizon volatility forecasts should incorporate both long memory (LM) and comovements (VAR) with common market factors. This is consistent with findings in prior research that stock return volatilities exhibit strong persistence over time and comovements with market-wide common factors (e.g., Ding, Granger, and Engle (1993), Bollerslev and Mikkelsen (1996), (1999), Baillie, Bollerslev, and Mikkelsen (1996), Andersen, Bollerslev, Diebold, and Ebens (2001), Andersen, Bollerslev, Diebold, and Labys (2001), (2003), and Pong, Shackleton, Taylor, and Xu (2004)). Modeling the LM feature helps to separate short-term or transitory fluctuations from long-term or permanent trends in the volatility time series and to formulate a volatility forecast with a greater emphasis on the latter than the former. Comovements with common market factors help to identify systematic components in stock return volatility and consequently improve forecasting performance. It is straightforward to incorporate both features into a long-memory vector autoregressive (LM-VAR) framework (e.g., Andersen et al. (2003)).

To provide empirical evidence on long-horizon forecasting performance, we conduct a horse race between horizon-matched historical volatility and competing volatility forecasts using a large sample of U.S. firms. Although we are primarily

[1]See survey papers by Poon and Granger (2003) and Andersen, Bollerslev, Christoffersen, and Diebold (2006) for a comprehensive review of the related literature.

[2]To the best of our knowledge, Alford and Boatsman (1995) is the only previous study that evaluates the long-horizon forecasting performance of competing volatility models (historical vs. shrinkage forecasts) using a large sample of U.S. firms. Some studies (e.g., Bollerslev and Mikkelsen (1996), (1999)) have also used long-memory models to examine the valuation of stock index options.

interested in the comparison between historical and LM-VAR forecasts, several other volatility forecasts are also considered, including two historical volatility-based forecasts (i.e., comparable-firm and shrinkage forecasts) and those based on commonly used "short-memory" models (e.g., AR(1) and ARMA(1,1)). For each volatility forecast, we examine its out-of-sample forecasting performance by comparing the volatility forecast with the future realized volatility. Forecasting errors are then evaluated for different firm-size groups and forecasting horizons. Consistent with the empirical option pricing literature, we use standard deviation as our volatility metric. In light of recent research on volatility forecast evaluation (e.g., Patton (2007)), we also consider alternative volatility metrics such as variance and logarithm of variance in a robustness analysis and find similar relative performance among competing volatility forecasts.

The main findings from our empirical analysis are as follows. First, historical volatility is a poor forecast for long-term stock return volatility.[3] When horizon-matched historical volatility is used to predict future volatility over a 5-year horizon, the median absolute percentage error is 22.8%, 20.7%, and 22.0% for large, medium-sized, and small firms, respectively. This means that historical forecasts miss the target by more than 20% in over half of the cases. Since option value is quite sensitive to changes in volatility, especially for at-the-money options, such magnitude of volatility forecasting errors can directly translate into sizeable valuation errors in reported option expense (as most options are granted at the money).

Second, the LM-VAR forecast is consistently more accurate than all other volatility forecasts across firm-size groups and forecasting horizons. The difference is statistically significant in nearly all cases. For example, the LM-VAR forecast is on average 14% to 30% more accurate than the historical forecast in predicting long-term volatility (based on median absolute percentage errors). As expected, these differences in volatility forecasting errors translate directly into differences in option values and reported income. Evaluating annual stock options granted to chief executive officers (CEOs) in our sample, we find that option valuation errors are on average reduced by 47% to 73% if the LM-VAR forecast is used instead of the historical forecast. Similarly, the error in reported income due to option expensing is reduced by approximately half if the LM-VAR forecast is used instead of the historical forecast. This level of improvements highlights the importance of volatility forecasting and lends strong support for using the LM-VAR model to forecast long-horizon stock return volatility.

In addition, volatility forecasts based on commonly used "short-memory" models (e.g., AR(1) and ARMA(1,1)) perform as poorly as historical volatility in predicting long-term volatility. Some of these forecasts exhibit marginal improvement, while others perform even worse than historical volatility. Intuitively, these volatility forecasts are constructed to place more weight on the short-term rather than long-term dynamics in volatility. Although they may be reasonable methods for forecasting short-term volatility, they are poor candidates for forecasting long-term volatility.

[3]It is well documented that historical volatility is a poor forecast for realized volatility over short horizons. How it does over long forecasting horizons, relative to volatility forecasts based on more sophisticated models, has not been fully addressed in the literature.

Finally, shrinkage forecasts, constructed by adjusting historical volatility toward comparable-firm volatility, perform slightly better than historical volatility in predicting long-term volatility. The improvement in forecasting performance is statistically significant for large and small firms but not for medium-sized firms. As expected, shrinkage forecasts perform substantially worse than the LM-VAR forecasts across all firm-size groups and forecasting horizons.

Of course, the improved forecasting performance is achieved at the cost of increased complexity in empirical implementation. Although the bivariate LM-VAR model is parsimonious, it is nonetheless more difficult to estimate than other volatility models. It is thus not at all certain that corporate users and accounting regulatory bodies (such as the FASB) will embrace the LM-VAR forecasting method. To bridge this gap, we investigate whether it is possible to use Corsi's (2004) mixed-horizon realized volatility model to achieve a comparable forecasting performance. The advantage of this approach is that it incorporates the autoregressive property of historical volatility over mixed horizons and is much simpler to implement. Our empirical analysis shows that it does a reasonable job forecasting long-term volatility, performing better than historical forecast but not as well as the LM forecast. To further improve its forecasting performance, more research is needed to determine the optimal combination of lagged realized volatilities and how the model can be generalized in a multivariate setting.

The rest of the paper proceeds as follows. The next section describes the LM-VAR model and its implementation. Section III describes the sample and summary statistics of stock return volatilities. Section IV empirically investigates the performance of alternative volatility forecasts. Section V examines the economic significance of forecasting errors and implications for option expensing. Section VI evaluates the performance of comparable-firm and shrinkage forecasts. An analysis of Corsi's (2004) mixed-horizon approach and other robustness checks is presented in Section VII. The final section concludes.

II. The LM-VAR Volatility Model

Let $\sigma_{i,\tau}$ be realized volatility for firm i's stock returns in period τ, which is estimated from higher frequency stock returns during each period. We are interested in predicting the realized volatility in period t given a time series of the firm's past realized volatilities (for $\tau = t-1, t-2, \ldots$). In general, this type of volatility forecasts can be written as follows:[4]

$$E_{t-1}(\sigma_{i,t}) = f(\sigma_{i,t-1}, \sigma_{i,t-2}, \ldots). \tag{1}$$

[4]In the empirical investigation, we construct monthly realized volatility from daily returns $\sigma_{i,t}^2 = \sum_{k=1}^{N} r_{i,t,k}^2$, where $r_{i,t,k}$ is firm i's stock return on the kth day of month t and N is the number of trading days in the month. We then focus on the dynamics of monthly volatilities. This approach aggregates out some information in daily return variations. Since our forecasting horizon is 1 to 5 years, the short-term dynamics of daily returns are likely to have only marginal effect on long-horizon volatility forecasts.

The historical forecast used in Alford and Boatsman (1995) is a special case of this more general volatility model when $E_{t-1}(\sigma_{i,t}) = \sigma_{i,t-1}$. It essentially assumes that volatility is a random walk and past volatility is an unbiased forecast for future volatility.

Prior research has established the importance of persistence in volatility time series and demonstrated that volatility persistence can be conveniently captured by a long-memory or fractionally integrated (FI) process. Although transitory shocks to volatility (such as jumps) may play a role in forecasting volatility over short horizons, they are generally less informative about the long-term trend in volatility movements. This is because jumps are rare events and their impact tends to be averaged out or much reduced over longer horizons. By capturing long-term trend in a natural way, long-memory models may provide a more accurate volatility forecast over longer horizons.

Following prior research, we consider the ARFIMA(p, d, q) specification of equation (1):

$$\Phi(B)(1-B)^d \sigma_{i,t} \quad = \quad \Theta(B)\epsilon_{i,t}, \tag{2}$$

where B denotes the lag operator, $\Phi(B)$ the autoregressive (AR) polynomial $1 - \sum_{k=1}^{p} \phi_k B^k$, $\Theta(B)$ is the moving average (MA) polynomial $1 + \sum_{k=1}^{q} \theta_k B^k$, $(1-B)^d = \sum_{k=0}^{\infty}[(\Gamma(d+1))/(\Gamma(k+1)\Gamma(d-k+1))](-1)^k B^k$, and $\epsilon_{i,t}$ is white noise. The autoregressive fractionally integrated moving average (ARFIMA) process is stationary with long memory if $0 < d < 0.5$ and nonstationary if $d > 0.5$. The specification in equation (2) will be referred to as the LM model subsequently. The choice of the ARFIMA specification is also supported by prior research on its demonstrated success in modeling volatilities of stock market indices and currency exchange rates (e.g., Bollerslev and Mikkelsen (1996), (1999), Andersen et al. (2003)). Although other long-memory models have been developed in the literature such as fractionally integrated generalized autoregressive conditional heteroskedasticity (FIGARCH) or fractionally integrated exponential generalized autoregressive conditional heteroskedasticity (FIEGARCH), these alternative models are more appropriate for modeling daily stock returns as opposed to monthly volatilities.

Another important dimension in forecasting long-term volatility is the cross-sectional correlation or comovement between stock return volatilities. As is well known, stock return volatilities are generally positively correlated across the market. In particular, we focus on comovements between a firm's stock return volatility and some common market factors. These common factors can be the volatility of, for example, a market index, a sector or industry index, or similar stocks. Such comovements motivate the modeling of the joint dynamics of volatilities of the individual stock and common factors using a vector-autoregressive (VAR) approach. In particular, the univariate volatility model in equation (2) is generalized into a multivariate LM-VAR model:

$$\Phi(B)(1-B)^d X_{i,t} \quad = \quad \Theta(B)\epsilon_{i,t}, \tag{3}$$

where $X_{i,t}$ is a vector consisting of volatilities for firm i and the common factors in period t.

Note that the proposed LM-VAR approach is quite general and includes many standard volatility forecasts as special cases. In a univariate setting, the random walk, AR(p), and ARMA(p, q) models are all special cases of the LM-VAR model. For example, the historical forecast is a special case of the univariate LM model when $p=1$ and $d=q=0$ in equation (2). By embedding these commonly used volatility forecasts within our LM-VAR approach, it is straightforward to set up a horse race and evaluate the forecasting performance of competing volatility models.

The cost of using a more sophisticated model is of course the markedly increased complexity in its estimation. This is true even for the univariate LM model and more so for the multivariate LM-VAR model. After surveying the relevant literature, we choose to implement the Geweke and Porter-Hudak (GPH) (1983) log-periodogram regression estimator as refined by Robinson (1995b). Andersen et al. (2003) use this GPH estimator to study realized volatilities on currency exchange rates and find it works reasonably well. In our context, a key advantage of this estimator is its computational simplicity. Robustness analysis shows that our empirical findings are not materially affected if we use alternative estimators (e.g., Robinson (1995a), Taqqu and Teverovsky (1997), Hurvich, Deo, and Brodsky (1998), Deo and Hurvich (2001), and Deo, Hurvich, and Lu (2006)) or apply shrinkage adjustment toward d values estimated for some common factors (such as volatilities of market indexes or comparable stocks).

To provide further details of the estimation procedure, we outline our implementation of the GPH estimator. Consider first the univariate LM model defined in equation (2) for the volatility time series σ_t for $t=1, 2, \ldots, T$ (with the firm index i suppressed for simplicity). The GPH estimator is a two-step procedure, with the fractional integration parameter d estimated first, followed by the autoregressive moving average (ARMA) parameters p and q. Parameter d is estimated using the following log-periodogram regression (implemented as an ordinary least squares (OLS) regression):

$$(4)\quad \log[I(\omega_j)] = \alpha + \beta\left\{-\ln\left[4\sin^2\left(\frac{\omega_j}{2}\right)\right]\right\} + e_j, \quad \text{for } j = 1, 2, \ldots, m,$$

where

$$I(\omega_j) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T}\sigma_t e^{i\omega_j t}\right|^2, \quad \omega_j = \frac{2\pi j}{T},$$

and m is an integer chosen to be close to $\sqrt{T}$. The slope coefficient (β) is the GPH estimate of d. With the estimated d value, the raw volatility time series is then fractionally differenced by applying the filter $(1 - B)^d$. An ARMA model is then fitted to the filtered volatility time series to obtain estimates for p and q. For the multivariate LM-VAR model, we apply the multivariate extension of the GPH estimator developed by Robinson (1995b) and obtain a common estimate of d value. The remaining steps are identical to the univariate case.

III. Data and Descriptive Statistics

To select the largest possible sample of U.S. firms for our empirical tests, we begin with all firms covered by the daily stock return database from the Center for Research in Security Prices (CRSP) over the time period from January 1, 1962 to December 31, 2004. To be included in our sample, we require firms to have return data available for the entire 15-year period from January 1, 1990 to December 31, 2004. This inclusion criterion results in a final sample of 2,066 firms including 448 large firms, 550 medium-sized firms, and 1,068 small firms, based on the New York Stock Exchange (NYSE) firm-size breakpoints in January 1990. For each firm in the sample, we construct a monthly time series of stock return volatility. Following previous research (e.g., Merton (1980), Poterba and Summers (1986), French, Schwert, and Stambaugh (1987), Schwert (1989), and Andersen et al. (2001), (2003)), we calculate the monthly volatility as the realized volatility during each calendar month using daily stock returns. These monthly volatility series are the subject of our subsequent empirical analysis.

Table 1 reports descriptive statistics of the monthly volatility time series for firms in our sample. These statistics are also reported for the Standard & Poor's 500 (S&P 500) Index, which is our main proxy for common market factors. All reported volatilities and their statistics are annualized. We calculate the mean, median, standard deviation, skewness, kurtosis, and first five order autocorrelations of the monthly volatility series for each firm in our sample. As these statistics vary widely across firms, we report the mean, median, and various percentiles of each statistic for each firm-size group to illustrate variations across firms in the group. Results for large, medium-sized, and small firms are presented in Panels A–C, respectively.

As expected, stock return volatility generally decreases with firm size. Take the firm-level average monthly volatility for example. The median value is 27.4%, 32.0%, and 53.0% for large, medium-sized, and small firms, respectively. In comparison, the corresponding average volatility on the S&P 500 index is much lower at only 13.1% due to the diversification effect. It is even lower than the 5th percentile of every firm-size group. In addition, monthly volatilities are right skewed with fat tails, as suggested by the positive skewness and excess kurtosis. This is consistent with the notion that volatilities usually vary within a normal range but can have occasional spikes.

More relevant to volatility forecasting are the dynamic properties of the volatility time series. The volatility time series of the S&P 500 index exhibits stronger autocorrelation than most individual firms do. As reported in the last column in Table 1, the first-order autocorrelation of the S&P 500 index volatility is 57.1%. In comparison, the average first-order autocorrelation is 47.4%, 45.6%, and 48.7% for large, medium-sized, and small firms, respectively. These high levels of autocorrelation suggest volatility persistence for firms in our sample and particularly for the S&P 500 index. More importantly, the level of autocorrelation decays slowly as the number of lags increases (see the first five order autocorrelations in Table 1). This pattern of slow decay is further illustrated in Figure 1, which plots autocorrelations for up to 24 lags (or 2 years). The signature hyperbolic decay of a long-memory process is evident in the plots.

TABLE 1

Summary Statistics of Monthly Stock Return Volatility

In Table 1, for each firm in the sample, we first calculate summary statistics of monthly stock return volatilities (annualized) including mean, median, standard deviation, skewness, kurtosis, and autocorrelations. The cross-sectional distribution of the firm-level statistics are then summarized and reported in the table for each firm-size group. Here $\rho(i)$s denote the first five order autocorrelations for $i = 1, \ldots, 5$. The corresponding summary statistics of the S&P 500 index are reported in the last column.

	5th Pctl	25th Pctl	Mean	Median	75th Pctl	95th Pctl	S&P 500
Panel A. Large Firms (N = 448)							
Mean	0.178	0.236	0.284	0.274	0.330	0.414	0.131
Median	0.162	0.210	0.257	0.250	0.297	0.381	0.116
Std. dev.	0.079	0.104	0.131	0.122	0.150	0.219	0.070
Skewness	1.113	1.531	2.208	1.943	2.514	4.137	3.745
Kurtosis	4.798	6.903	14.290	9.884	15.540	36.430	35.430
$\rho(1)$	0.306	0.399	0.474	0.478	0.543	0.644	0.571
$\rho(2)$	0.240	0.343	0.417	0.412	0.492	0.601	0.487
$\rho(3)$	0.223	0.324	0.395	0.384	0.465	0.578	0.436
$\rho(4)$	0.150	0.241	0.321	0.309	0.396	0.515	0.313
$\rho(5)$	0.142	0.231	0.310	0.301	0.384	0.502	0.339
Panel B. Medium-Sized Firms (N = 550)							
Mean	0.163	0.255	0.348	0.320	0.422	0.611	
Median	0.154	0.230	0.313	0.290	0.379	0.537	
Std. dev.	0.060	0.118	0.170	0.155	0.204	0.317	
Skewness	0.750	1.200	1.791	1.595	2.121	3.355	
Kurtosis	3.690	5.160	10.110	7.315	10.860	24.460	
$\rho(1)$	0.264	0.374	0.456	0.446	0.539	0.665	
$\rho(2)$	0.185	0.310	0.395	0.389	0.474	0.611	
$\rho(3)$	0.159	0.274	0.361	0.353	0.443	0.578	
$\rho(4)$	0.093	0.206	0.304	0.295	0.387	0.551	
$\rho(5)$	0.065	0.186	0.286	0.281	0.373	0.523	
Panel C. Small Firms (N = 1,068)							
Mean	0.230	0.371	0.552	0.530	0.695	0.946	
Median	0.205	0.328	0.485	0.466	0.616	0.829	
Std. dev.	0.105	0.186	0.319	0.278	0.398	0.651	
Skewness	0.681	1.138	1.790	1.514	2.076	3.711	
Kurtosis	3.427	4.773	10.250	6.619	10.340	26.930	
$\rho(1)$	0.237	0.385	0.487	0.489	0.599	0.739	
$\rho(2)$	0.161	0.302	0.419	0.414	0.526	0.688	
$\rho(3)$	0.137	0.273	0.386	0.382	0.505	0.661	
$\rho(4)$	0.083	0.219	0.341	0.335	0.452	0.625	
$\rho(5)$	0.062	0.199	0.322	0.312	0.434	0.607	

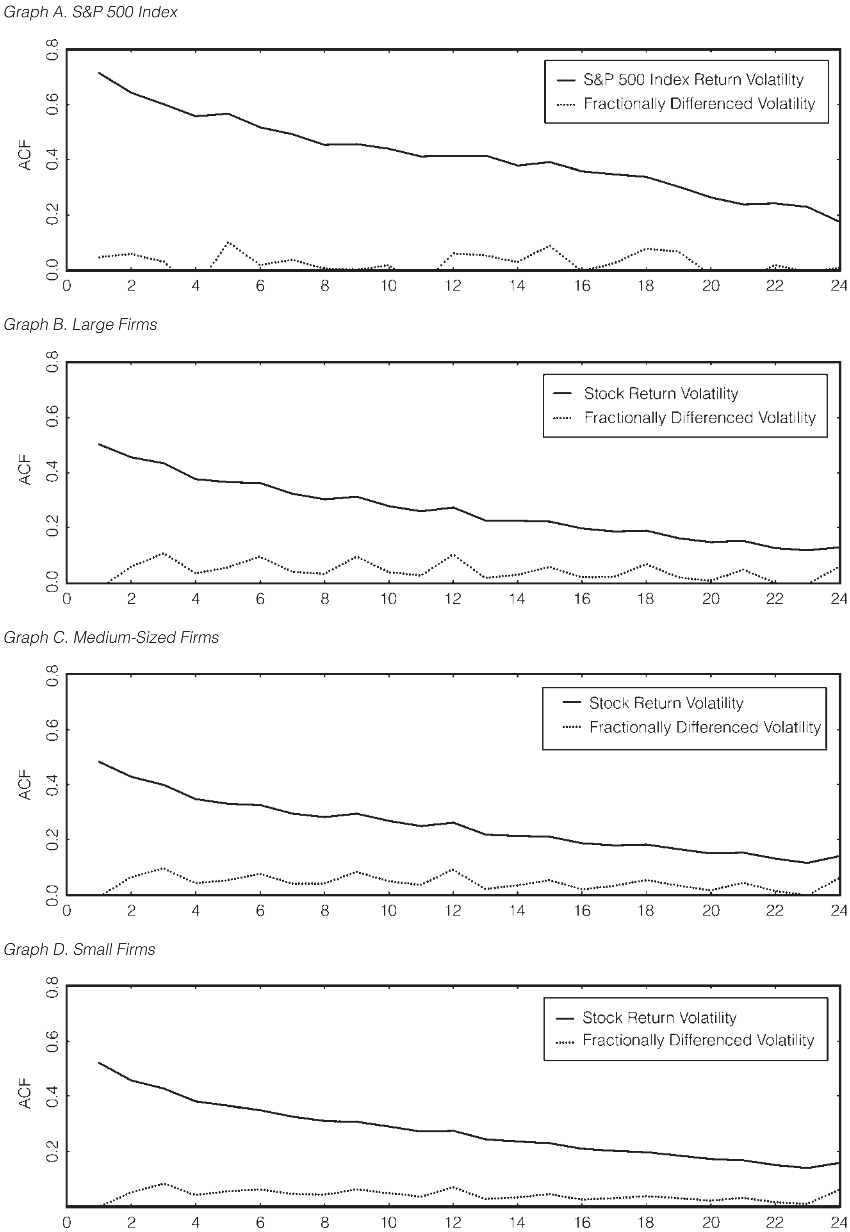
We also perform diagnostic analysis of the volatility time series, which provides not only a better understanding of the dynamic properties of the volatility time series but also useful guidance for model specification. We fit the commonly used AR(1) model to the volatility time series of each individual stock and compute the diagnostic Ljung-Box statistics using 20 lags. The mean, median, and other percentiles of the Ljung-Box statistics across firms in each firm-size group are reported in Table 2. As the *p*-values for the Ljung-Box statistics indicate, the AR(1) is severely misspecified for virtually all firms in our sample.

Table 2 also reports the estimated degree of fractional integration (d). The average estimated value of d is quite similar across the three firm-size groups, at 0.362, 0.352, and 0.375 for large, medium-sized, and small firms in our sample, respectively. In comparison, the corresponding estimate for the S&P 500 volatility series is 0.425. As indicated by the 5th and 95th percentiles reported in Table 2, more than 90% of the estimated d values are between 0.198 and 0.485 for each of the three firm-size groups, suggesting a stationary volatility process with long memory for these firms. Figure 2 plots the estimated values of d separately for

FIGURE 1

Autocorrelations of Monthly Stock and S&P 500 Index Return Volatilities

Figure 1 plots autocorrelations of monthly stock and S&P 500 index return volatilities as well as those of fractionally differenced monthly stock and S&P 500 index return volatilities. ACF stands for autocorrelation function.



large, medium-sized, and small firms in Panels A–C, respectively. It is clear that the overwhelming majority of estimated d values are between 0.2 and 0.5. There is no estimated value of d near or below 0. Although a small fraction of estimated

d values are slightly above 0.5, the difference is rarely statistically significant. These results provide strong evidence that the monthly volatility series is stationary with long memory for the S&P 500 index and for nearly all firms in our sample.

TABLE 2

Diagnostic Statistics of Monthly Stock Return Volatility

In Table 2, for each firm in the sample, we fit the AR(1) model to the monthly stock return volatility series and calculate the Ljung-Box statistic using 20 lags of residuals. The cross-sectional distribution of the Ljung-Box statistic, denoted as Q(20): AR(1), is reported in the table for each firm-size group, together with the corresponding p-value, fractional integration coefficient (d), and correlation with the S&P 500 index volatilities (ρ(SPX)). The corresponding diagnostic statistics of the S&P 500 index are reported in the last column.

	5th Pctl	25th Pctl	Mean	Median	75th Pctl	95th Pctl	S&P 500
Panel A. Large Firms (N = 448)							
Q(20): AR(1)	45.38	75.43	110.7	101.3	134.7	202.2	246.6
p-value	0.000	0.000	0.000	0.000	0.000	0.001	0.000
d	0.252	0.311	0.362	0.359	0.407	0.472	0.425
ρ(SPX)	0.259	0.446	0.525	0.540	0.622	0.741	1.000
Panel B. Medium-Sized Firms (N = 550)							
Q(20): AR(1)	29.28	51.31	93.62	78.66	119.4	216.8	
p-value	0.000	0.000	0.001	0.000	0.000	0.007	
d	0.224	0.290	0.352	0.340	0.407	0.475	
ρ(SPX)	0.101	0.251	0.347	0.372	0.480	0.599	
Panel C. Small Firms (N = 1,068)							
Q(20): AR(1)	26.65	46.15	84.69	67.55	100.0	198.0	
p-value	0.000	0.000	0.002	0.000	0.001	0.013	
d	0.198	0.301	0.375	0.370	0.439	0.485	
ρ(SPX)	0.051	0.113	0.179	0.179	0.300	0.480	

Figure 1 also provides further evidence on the long-memory property of the volatility time series. In addition to the autocorrelations of the raw volatility time series, it also plots the autocorrelations of the fractionally differenced volatility series, which are obtained by applying the filter $(1-B)^d$ to the raw monthly volatilities. Fractional differentiation seems to remove most of the predictive components of the volatility series. This finding suggests that long-memory models are good candidates for modeling volatility time series.

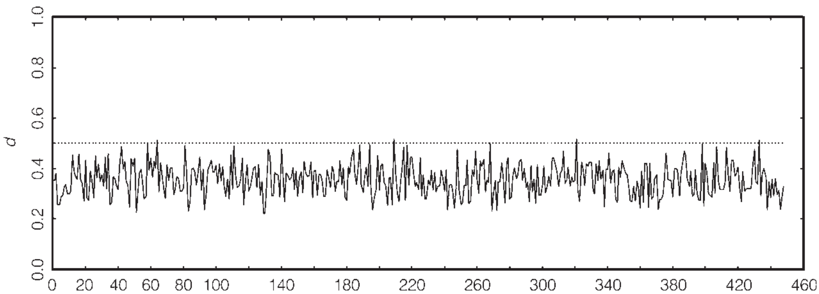
Finally, the last row of each panel in Table 2 also reports the correlation between monthly volatilities of an individual stock and the S&P 500 index. The average correlation is 52.5%, 34.7%, and 17.9% for large, medium-sized, and small firms in our sample, respectively. As expected, the correlation weakens as firm size declines, since the S&P 500 index is primarily a market index for large firms. Correlations also vary widely across firms within each firm-size group. For example, the 5th percentile for large firms is 25.9% compared to the 95th percentile of 74.1%. For small firms, the corresponding numbers are 5.1% and 48.0%, respectively. Nonetheless, it is important to recognize the correlation in volatilities between an individual stock and the S&P 500 index. The presence of such comovement suggests that the S&P 500 volatility is likely helpful in forecasting stock return volatility for these firms (more so for large firms) in a multivariate (VAR) framework. It provides direct support for our subsequent construction of a benchmark volatility forecast based on a bivariate LM-VAR model of the joint volatility time series of an individual stock and the S&P 500 index.

FIGURE 2

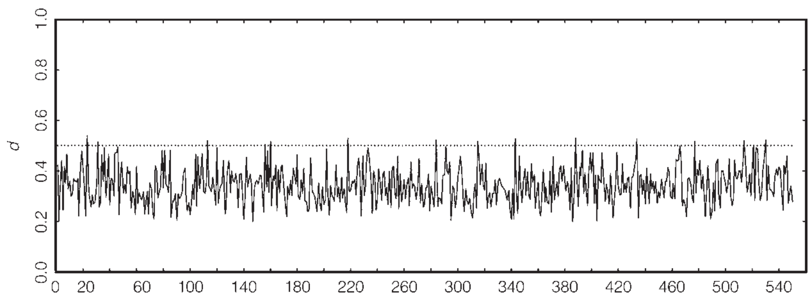
Degree of Fractional Integration of Monthly Stock Return Volatilities

Figure 2 plots the degree of fractional integration (d) of monthly stock return volatilities for large, medium-sized, and small cap stocks.

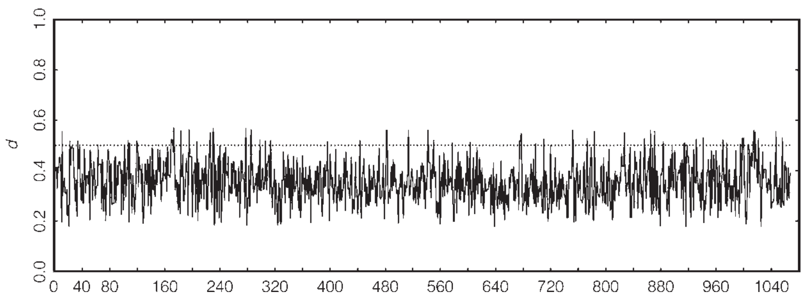
Graph A. Large Firms



Graph B. Medium-Sized Firms



Graph C. Small Firms



IV. Forecasting Performance

In this section, we empirically examine the performance of a number of volatility forecasts including historical forecasts,[5] volatility forecasts based on

[5]In addition to the benchmark historical forecasts based on the past 1-, 3-, and 5-year realized volatilities, we also consider a "long" historical forecast based on the realized volatility of all stock returns prior to the forecasting date. Untabulated results indicate that the long historical forecast performs mostly worse than all three benchmark historical forecasts. The only exception is the large-firm sample at the 5-year forecasting horizon, where the long historical forecast is more accurate than all

commonly used "short-memory" models (e.g., AR(1) and ARMA(1,1)), and those based on the LM-VAR model. For the LM-VAR forecasts, we consider two additional variations that are based on either the LM model or the VAR model. These variations are similar to the LM-VAR model except that one feature (LM or VAR) is dropped in order to determine the marginal contribution of the other feature to forecasting performance.

To construct a volatility forecast over a given future horizon, the underlying volatility model must be estimated first using data available at the time of the forecast. Historical forecasts are straightforward to construct. For the remaining five volatility forecasts, model estimation is necessary. On each forecasting date, the volatility model is reestimated using the full historical time series of the monthly volatilities leading up to that date. As the forecasting date moves forward, we reestimate the model by including additional volatility observations available since the previous forecasting date. The AR(1) and ARMA(1,1) models are estimated using the standard approach. For the univariate and bivariate LM-VAR models, we follow the estimation procedure outlined previously except that we set the AR lag to 1, based on the diagnosis from the Akaike and Schwarz criteria, and we set the MA lag to 0. In other words, we simply follow Andersen et al. (2003) and adopt the ARFIMA$(1, d, 0)$ specification for the LM-VAR models once the memory parameter d is estimated. After each volatility model (e.g., the LM-VAR model) is estimated, volatility forecasts for future months are generated from the estimated model. For the LM-VAR model, we truncate the estimated d value at 0.495 to ensure model stability.[6] The required volatility forecast for a longer horizon (say, a 5-year horizon) is then constructed by aggregating the monthly volatility forecasts over all months in the forecasting horizon. In other words, it is calculated as the square root of the sum of the monthly variance forecasts over the forecasting horizon.

To determine the appropriate forecasting horizon for our empirical tests, we follow the FASB guidelines for determining the fair value of stock options. Firms are required to use the Black-Scholes-Merton model (Black and Scholes (1973), Merton (1973)) or its binomial variations, modified to take into account the option's expected life or early exercise, to determine the fair value of stock options. Although nearly all options are granted with a 10-year maturity, most are exercised between 4 and 6 years after the grant date (see Huddart and Lang (1996), (2003), Hemmer, Matsunaga, and Shevlin (1998), Carpenter (1998), Heath, Huddart, and Lang (1999), and Carpenter and Remmers (2001)). We thus primarily focus on a 5-year forecasting horizon. Other forecasting horizons are also included for comparison purposes.

To evaluate the performance of each volatility model, we compare the volatility forecast with the realized volatility over the forecasting horizon. Model estimation and volatility forecasting are performed at the end of June and December in each calendar year during the period from 1995 to 2003. The

three benchmark historical forecasts but still less accurate than the LM-VAR forecast. We thus do not include the long historical forecast in our tables.

[6]As the results in Table 2 and Figure 2 indicate, this is rarely needed (affecting approximately 3.3% of firms in the sample).

results are summarized in Tables 3–5 for large, medium-sized, and small firms, respectively. As in Alford and Boatsman (1995), the forecasting error is calculated as the difference between the realized volatility and the volatility forecast. The performance of a volatility forecast is then evaluated separately for each firm-size group and forecasting horizon, using various commonly used performance measures (or loss functions) including mean absolute errors, median absolute errors, mean absolute percentage errors, and median absolute percentage errors. Following West (2006), we use a *t*-test to formally evaluate the null hypothesis that the absolute error of a volatility forecast is equal to the corresponding error of the benchmark forecast. Test statistics are calculated for the pooled sample of all firms and jointly for the forecasting errors of all firms. As the results are quite consistent, we only report the statistics from the joint tests in Tables 3–5 (last column). The horizon-matched historical forecast is used as the benchmark forecast, and all other forecasts are evaluated relative to this benchmark.

TABLE 3

The Performance of Alternative Volatility Forecasts (large firms)

Table 3 reports the summary statistics of the forecasting errors of alternative volatility forecasts. HIS_1, HIS_3, and HIS_5 denote, respectively, the historical forecasts based on the 1-, 3-, and 5-year historical volatilities. Other forecasts are based on volatility models estimated using monthly volatility time series over a rolling historical sample period. LM denotes the long-memory model of the firm's monthly volatility series, while VAR denotes the bivariate vector autoregressive model of the joint monthly volatility series of the firm and the S&P 500 index. LM-VAR is the long-memory version of the VAR model. The last column reports the *t*-statistic, showing that the absolute error of the volatility forecast is equal to the horizon-matched historical forecast. The total sample period is from 1962 to 2004, and the forecasting period is from June 1995 to December 2003. N/A indicates not applicable.

	Absolute Error				Absolute Percentage Error				
Forecast	Mean	Median	75th Pctl	95th Pctl	Mean	Median	75th Pctl	95th Pctl	*t*-Test
Panel A. 1-Year Forecasting Horizon									
HIS_1	0.054	0.042	0.075	0.144	0.176	0.152	0.253	0.430	N/A
HIS_3	0.066	0.054	0.092	0.167	0.213	0.188	0.300	0.519	14.718
HIS_5	0.073	0.059	0.102	0.186	0.234	0.209	0.329	0.572	18.375
AR(1)	0.061	0.049	0.085	0.153	0.195	0.171	0.280	0.475	9.509
ARMA(1,1)	0.055	0.044	0.078	0.143	0.181	0.158	0.263	0.440	2.316
LM	0.049	0.038	0.068	0.126	0.159	0.136	0.229	0.394	–8.627
VAR	0.051	0.041	0.072	0.133	0.168	0.144	0.243	0.409	–4.192
LM-VAR	0.045	0.035	0.063	0.116	0.145	0.124	0.209	0.362	–14.865
Panel B. 3-Year Forecasting Horizon									
HIS_1	0.068	0.057	0.095	0.171	0.218	0.199	0.306	0.498	–10.344
HIS_3	0.076	0.067	0.106	0.181	0.243	0.224	0.339	0.552	N/A
HIS_5	0.077	0.066	0.107	0.184	0.244	0.228	0.337	0.562	1.014
AR(1)	0.069	0.059	0.096	0.167	0.220	0.202	0.306	0.498	–9.377
ARMA(1,1)	0.070	0.059	0.097	0.170	0.223	0.205	0.313	0.508	–7.741
LM	0.057	0.049	0.079	0.138	0.182	0.168	0.259	0.405	–21.838
VAR	0.064	0.054	0.088	0.157	0.202	0.188	0.287	0.451	–16.670
LM-VAR	0.053	0.046	0.074	0.126	0.170	0.156	0.242	0.386	–23.498
Panel C. 5-Year Forecasting Horizon									
HIS_1	0.079	0.067	0.110	0.187	0.251	0.235	0.353	0.565	1.904
HIS_3	0.081	0.071	0.112	0.187	0.260	0.242	0.356	0.589	5.281
HIS_5	0.077	0.066	0.107	0.182	0.248	0.228	0.339	0.584	N/A
AR(1)	0.074	0.064	0.103	0.175	0.237	0.222	0.333	0.538	–4.446
ARMA(1,1)	0.079	0.069	0.110	0.184	0.252	0.238	0.350	0.561	2.147
LM	0.071	0.062	0.098	0.162	0.225	0.215	0.314	0.481	–9.162
VAR	0.068	0.060	0.095	0.158	0.218	0.208	0.305	0.468	–11.784
LM-VAR	0.061	0.054	0.084	0.144	0.197	0.185	0.272	0.428	–17.935

TABLE 4

The Performance of Alternative Volatility Forecasts (medium-sized firms)

Table 4 reports the summary statistics of the forecasting errors of alternative volatility forecasts. HIS_1, HIS_3, and HIS_5 denote, respectively, the historical forecasts based on the 1-, 3-, and 5-year historical volatilities. Other forecasts are based on volatility models estimated using monthly volatility time series over a rolling historical sample period. LM denotes the long-memory model of the firm's monthly volatility series, while VAR denotes the bivariate vector autoregressive model of the joint monthly volatility series of the firm and the S&P 500 index. LM-VAR is the long-memory version of the VAR model. The last column reports the *t*-statistic, showing that the absolute error of the volatility forecast is equal to the horizon-matched historical forecast. The total sample period is from 1962 to 2004, and the forecasting period is from June 1995 to December 2003. N/A indicates not applicable.

	Absolute Error				Absolute Percentage Error				
Forecast	Mean	Median	75th Pctl	95th Pctl	Mean	Median	75th Pctl	95th Pctl	*t*-Test
Panel A. 1-Year Forecasting Horizon									
HIS_1	0.066	0.047	0.088	0.193	0.184	0.159	0.268	0.443	N/A
HIS_3	0.076	0.055	0.102	0.221	0.211	0.178	0.308	0.521	10.024
HIS_5	0.082	0.058	0.110	0.237	0.226	0.193	0.323	0.564	14.358
AR(1)	0.069	0.049	0.093	0.204	0.192	0.164	0.281	0.470	3.014
ARMA(1,1)	0.067	0.047	0.090	0.193	0.186	0.161	0.269	0.451	0.429
LM	0.059	0.042	0.079	0.171	0.164	0.141	0.240	0.402	–9.708
VAR	0.062	0.044	0.083	0.180	0.173	0.148	0.254	0.427	–5.091
LM-VAR	0.055	0.040	0.074	0.162	0.156	0.132	0.228	0.385	–12.220
Panel B. 3-Year Forecasting Horizon									
HIS_1	0.084	0.061	0.116	0.236	0.228	0.199	0.328	0.553	–3.080
HIS_3	0.087	0.064	0.121	0.244	0.235	0.205	0.335	0.575	N/A
HIS_5	0.088	0.065	0.120	0.248	0.236	0.207	0.335	0.567	0.840
AR(1)	0.081	0.059	0.112	0.232	0.219	0.191	0.314	0.531	–5.889
ARMA(1,1)	0.083	0.060	0.116	0.233	0.226	0.196	0.328	0.543	–4.182
LM	0.071	0.052	0.097	0.203	0.194	0.170	0.281	0.467	–14.944
VAR	0.076	0.056	0.106	0.216	0.208	0.182	0.302	0.496	–10.969
LM-VAR	0.065	0.048	0.090	0.182	0.178	0.155	0.258	0.431	–19.244
Panel C. 5-Year Forecasting Horizon									
HIS_1	0.093	0.070	0.130	0.262	0.252	0.221	0.364	0.608	3.063
HIS_3	0.093	0.069	0.128	0.264	0.247	0.218	0.349	0.605	2.623
HIS_5	0.090	0.067	0.123	0.259	0.240	0.207	0.337	0.598	N/A
AR(1)	0.087	0.065	0.121	0.249	0.233	0.206	0.329	0.573	–3.223
ARMA(1,1)	0.092	0.070	0.127	0.258	0.247	0.218	0.355	0.596	1.403
LM	0.083	0.063	0.115	0.231	0.222	0.198	0.319	0.528	–7.926
VAR	0.083	0.063	0.114	0.230	0.222	0.197	0.315	0.526	–8.167
LM-VAR	0.074	0.057	0.101	0.207	0.198	0.178	0.280	0.466	–15.308

It is also important to emphasize that an inappropriate choice of loss function may lead to incorrect inferences on the relative performance of alternative volatility forecasts, since realized volatility is not observable. Recent research by Patton (2007) provides theoretical guidance on the choice of appropriate loss functions. Most relevant to our study is the finding that certain loss functions may not provide a correct ranking of volatility forecasts if volatility is measured with error. However, this problem is mitigated in our case, since realized volatility is calculated using returns sampled at sufficiently high frequencies. In particular, Patton (2007) shows (in his Table II) that most loss functions (including ours) provide the correct ranking of volatility forecasts if realized volatility is calculated using returns sampled at a frequency of 78 times per period. Since we use daily returns to calculate realized volatility and our forecasting horizon ranges from 1 to 5 years, our realized volatilities are calculated using returns sampled at a frequency of at least 250 times per forecasting period. This is much higher than the minimum requirement determined by Patton (2007). We thus do not expect our choice of loss function to bias the ranking of volatility forecasts. In our subsequent

TABLE 5

The Performance of Alternative Volatility Forecasts (small firms)

Table 5 reports the summary statistics of the forecasting errors of alternative volatility forecasts. HIS_1, HIS_3, and HIS_5 denote, respectively, the historical forecasts based on the 1-, 3-, and 5-year historical volatilities. Other forecasts are based on volatility models estimated using monthly volatility time series over a rolling historical sample period. LM denotes the long-memory model of the firm's monthly volatility series, while VAR denotes the bivariate vector autoregressive model of the joint monthly volatility series of the firm and the S&P 500 index. LM-VAR is the long-memory version of the VAR model. The last column reports the *t*-statistic, showing that the absolute error of the volatility forecast is equal to the horizon-matched historical forecast. The total sample period is from 1962 to 2004, and the forecasting period is from June 1995 to December 2003. N/A indicates not applicable.

	Absolute Error				Absolute Percentage Error				
Forecast	Mean	Median	75th Pctl	95th Pctl	Mean	Median	75th Pctl	95th Pctl	*t*-Test
Panel A. 1-Year Forecasting Horizon									
HIS_1	0.110	0.073	0.145	0.343	0.198	0.164	0.287	0.499	N/A
HIS_3	0.126	0.086	0.169	0.372	0.232	0.189	0.333	0.600	11.885
HIS_5	0.137	0.098	0.185	0.396	0.258	0.209	0.369	0.678	18.612
AR(1)	0.116	0.081	0.157	0.343	0.215	0.179	0.311	0.551	5.246
ARMA(1,1)	0.111	0.075	0.147	0.343	0.201	0.168	0.288	0.509	0.830
LM	0.099	0.067	0.132	0.307	0.181	0.152	0.264	0.458	–9.366
VAR	0.106	0.071	0.140	0.327	0.193	0.162	0.281	0.491	–3.470
LM-VAR	0.089	0.061	0.118	0.271	0.164	0.137	0.238	0.414	–17.508
Panel B. 3-Year Forecasting Horizon									
HIS_1	0.133	0.094	0.187	0.384	0.244	0.207	0.356	0.609	–5.521
HIS_3	0.139	0.100	0.193	0.405	0.258	0.212	0.371	0.663	N/A
HIS_5	0.144	0.103	0.199	0.419	0.270	0.221	0.386	0.708	3.401
AR(1)	0.132	0.095	0.183	0.377	0.243	0.202	0.351	0.618	–6.936
ARMA(1,1)	0.132	0.095	0.184	0.382	0.243	0.204	0.353	0.615	–6.216
LM	0.110	0.080	0.153	0.315	0.202	0.171	0.294	0.511	–23.500
VAR	0.122	0.088	0.171	0.351	0.224	0.190	0.326	0.559	–15.094
LM-VAR	0.099	0.072	0.138	0.288	0.185	0.155	0.268	0.468	–27.899
Panel C. 5-Year Forecasting Horizon									
HIS_1	0.147	0.107	0.205	0.419	0.268	0.226	0.389	0.672	0.389
HIS_3	0.146	0.105	0.203	0.421	0.269	0.222	0.384	0.696	–0.497
HIS_5	0.144	0.103	0.202	0.419	0.267	0.220	0.388	0.687	N/A
AR(1)	0.140	0.101	0.196	0.397	0.258	0.211	0.374	0.659	–5.945
ARMA(1,1)	0.144	0.104	0.201	0.414	0.264	0.220	0.382	0.671	–1.531
LM	0.127	0.092	0.178	0.363	0.233	0.197	0.339	0.577	–16.669
VAR	0.134	0.097	0.186	0.382	0.245	0.207	0.357	0.611	–10.333
LM-VAR	0.115	0.084	0.160	0.326	0.210	0.178	0.305	0.526	–24.642

robustness analysis, we also consider alternative volatility metrics (variance and logarithm of variance) and find consistent ranking of our volatility forecasts.

Analysis of forecasting errors reported in Tables 3–5 leads to several interesting findings. First, historical volatility is a poor forecast for long-term volatility. When the 5-year historical volatility is used to predict realized volatility over the next 5 years, the median absolute percentage error is 22.8%, 20.7%, and 22.0% for large, medium-sized, and small firms, respectively. This means that historical forecasts miss the target by more than 20% in over half of the cases. Interestingly, our results also offer some support for the FASB recommendation that horizon-matched historical volatility should be included as a key factor in volatility estimation.[7] For the 1-year forecasting horizon (Panel A of Tables 3–5), the horizon-matched 1-year historical volatility is more accurate than either the 3- or 5-year historical volatility. The difference is statistically significant at the 1% level in all cases. At the 5-year forecasting horizon, horizon matching also produces the

[7]See Section A32 of the FASB's final statement on share-based payment, SFAS 123R.

most accurate forecast, with the difference mostly statistically significant (three cases at 1%, one at 10%, and two insignificant). The 3-year horizon is the exception, as the horizon-matched 3-year historical volatility is always less accurate than the 1-year historical volatility (significant at the 1% level).

Second, volatility forecasts based on "short-memory" models such as AR(1) and ARMA(1,1) perform as poorly as the historical forecast in predicting long-term volatility. Volatility forecasts based on the AR(1) model are generally more accurate than historical forecasts at longer forecasting horizons (e.g., the 5-year horizon) but less accurate at shorter forecasting horizons (e.g., the 1-year horizon), although the difference in forecasting accuracy appears to be quite minor both economically and statistically. The performance of the ARMA(1,1) model is similar except that it seems to perform more poorly than the AR(1) model at longer horizons (e.g., the 5-year horizon). There appears to be a typical trade-off between in-sample fitting and out-of-sample forecasting performance. While the ARMA(1,1) fits the volatility process better in sample as a result of a more flexible model specification, the out-of-sample forecasting performance does not necessarily improve.

More importantly, the LM-VAR model consistently provides the most accurate forecast for long-term volatility across all firm-size groups and forecasting horizons. Measured by median absolute (percentage) error, the LM-VAR forecast is 14.9% to 31.3% (14.0% to 30.4%) more accurate than the historical volatility across firm-size groups and forecasting horizons. Consider the results for large firms at the 5-year forecasting horizon (Panel C of Table 3), for example. The 5-year historical volatility (HIS_5) has a median absolute (percentage) error of 0.066 (0.228). In comparison, the LM-VAR forecast has a median absolute (percentage) error of 0.054 (0.185) or a reduction of 18.2% (18.9%). A formal statistical test (last column in the table) further confirms that the LM-VAR forecast is the most accurate among all volatility forecasts, with the difference in forecasting errors statistically significant in all cases. Both LM and VAR features are important in improving forecasting performance, as both the LM and VAR forecasts are substantially more accurate than the historical forecast (significant at the 1% level in all cases). Neither feature alone is sufficient in achieving the best, or close to the best, forecasting performance because the error of the LM-VAR forecast is significantly smaller than the corresponding error of either the LM or VAR forecast (all at the 1% level except one at the 10% level). This finding highlights the importance of both LM and VAR features in predicting long-term volatility.

V. The Impact of Forecasting Errors on Option Expensing

Option expensing is based on the grant-date value of stock option grants. As option value is quite sensitive to changes in volatility, volatility forecasting errors are expected to have a substantial impact on option value and reported income. To quantify their economic significance, we next empirically investigate the impact of volatility forecasting errors on option valuation. We determine option value first using the future realized volatility and then the volatility forecast. The difference in option value is our measure of option valuation error due to volatility forecasting error. To proceed along this line, we need stock option data for firms in our sample. We thus cross-reference firms in our sample with the Standard and Poor's ExecuComp database. We only keep firms that are covered by ExecuComp and grant stock options to their top executives. Of the 448 large firms in the original sample, 390 firms still remain after cross-referencing with ExecuComp. As the firm size declines, the number of firms remaining drops off significantly. While 316 of the original 550 medium-sized firms remain after cross-referencing with ExecuComp, only 176 of the original 1,068 small firms are left. The sharp drop-off, especially for small firms, is expected as ExecuComp only includes firms covered by the S&P 1500 index. Nonetheless, we still have a sizeable sample of 804 firms.

For the 804 firms remaining in the sample, we evaluate the economic significance of forecasting errors by calculating the value of all stock option grants received by the CEO during the year the volatility forecast is constructed. We choose the CEO, as he or she is the highest ranked executive in the firm and usually receives the most option grants. Only new option grants are included in our analysis because ExecuComp provides details of the option terms (such as strike price, maturity date, and number of options in each grant) for new options only. We also make several simplifying assumptions: The Black-Scholes-Merton model is a reasonable model for estimating option value, all options have a 5-year maturity, the risk-free rate is 5%, and the expected dividend yield over the next 5 years is identical to the actual dividend yield over the previous 5 years. These assumptions are not unreasonable, since our goal is to examine the impact of volatility forecasting errors across different volatility models and these assumptions are unlikely to change the relative performance between volatility models in any biased manner.

A potential inconsistency does arise when volatility forecasts from time-series-based models (such as the LM-VAR model) are used as an input to the Black-Scholes-Merton model. While stock returns are assumed to be normally distributed with constant volatility in the latter model, they are not in the former. For long-term options, this is unlikely to be a problem, as stock returns are roughly normally distributed over long horizons. For firms in our sample, this is indeed the case. We perform several normality tests (e.g., Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling tests) on monthly, quarterly, semiannual, and annual returns for these firms. Untabulated test results show that while normality is strongly rejected for most firms if monthly returns are used, it is rarely rejected if annual returns are used. Take the results from the Kolmogorov-Smirnov test, for example. Normality is rejected at the 1% significance level for 87.7% of firms when monthly returns are tested, but only for 7.8% of firms when annual returns are tested. Results from the other three normality tests are similar, supporting normality in annual stock returns. It is thus reasonable to use the Black-Scholes-Merton model to value options with 1-year or longer maturities.

Table 6 summarizes results on the impact of volatility forecasting errors on option value. The results for large, medium-sized, and small firms are shown in Panels A–C, respectively. For each volatility forecast, we first calculate the value of the CEO's option grants, in millions of dollars, using the volatility forecast. We repeat the calculation using the realized volatility over the expected life of the option (i.e., the 5-year period after the grant date). The absolute error and absolute percentage error are then calculated for each volatility forecast. As expected, a more accurate volatility forecast does translate into a smaller error in option value. In particular, the errors from using the LM-VAR forecast are generally less than half the size of the errors from using the 5-year historical volatility across all firm-size groups and alternative measures of forecasting errors. This level of improvement is clearly economically significant.

TABLE 6

The Impact of Forecasting Errors on Option Value

Table 6 summarizes the impact of forecasting errors on option value for alternative volatility forecasts. Both absolute errors and absolute percentage errors are reported for each firm-size group. HIS_1, HIS_3, and HIS_5 denote, respectively, the historical forecasts based on the 1-, 3-, and 5-year historical volatilities. Other forecasts are based on volatility models estimated using monthly volatility time series over a rolling historical sample period. LM denotes the long-memory model of the firm's monthly volatility series, while VAR denotes the bivariate vector autoregressive model of the joint monthly volatility series of the firm and the S&P 500 index. LM-VAR is the long-memory version of the VAR model. The total sample period is from 1962 to 2004, and the forecasting period is from June 1995 to December 2003.

Forecast	HIS_1	HIS_3	HIS_5	AR(1)	ARMA(1,1)	LM	VAR	LM-VAR
Panel A. Large Firms								
Absolute Error in Option Value ($million)								
Mean	0.6967	0.8736	1.0838	0.8143	0.7144	0.5934	0.5531	0.4790
Median	0.1539	0.2415	0.3331	0.1899	0.1581	0.1240	0.1095	0.0905
75th pctl	0.4338	0.6662	0.8732	0.5717	0.4965	0.4066	0.3678	0.3305
95th pctl	2.8117	3.8390	4.4872	3.5728	2.6546	2.4221	2.2034	1.9536
Absolute Percentage Error in Option Value								
Mean	0.1311	0.1795	0.2165	0.1563	0.1355	0.1142	0.1052	0.0912
Median	0.0790	0.1297	0.1667	0.1051	0.0810	0.0648	0.0560	0.0519
75th pctl	0.1673	0.2419	0.2867	0.2012	0.1782	0.1432	0.1282	0.1119
95th pctl	0.4842	0.5305	0.5256	0.5151	0.5003	0.4573	0.3576	0.3092
Fraction positive	0.3810	0.0644	0.0280	0.1233	0.2353	0.3137	0.3642	0.4174
Panel B. Medium-Sized Firms								
Absolute Error in Option Value ($million)								
Mean	0.2393	0.2147	0.2522	0.2142	0.2231	0.2063	0.2060	0.1684
Median	0.0555	0.0576	0.0771	0.0520	0.0578	0.0493	0.0446	0.0410
75th pctl	0.1503	0.2124	0.2786	0.1988	0.1762	0.1322	0.1408	0.1132
95th pctl	0.7336	0.8307	0.9798	0.8554	0.7659	0.7472	0.7150	0.6984
Absolute Percentage Error in Option Value								
Mean	0.1147	0.1372	0.1641	0.1209	0.1108	0.0991	0.0998	0.0852
Median	0.0773	0.0985	0.1217	0.0814	0.0766	0.0646	0.0623	0.0562
75th pctl	0.1543	0.1818	0.2257	0.1488	0.1408	0.1226	0.1217	0.1098
95th pctl	0.3059	0.4228	0.4998	0.4049	0.3254	0.3307	0.3159	0.2631
Fraction positive	0.4613	0.1624	0.0664	0.2251	0.3210	0.3469	0.3026	0.4244
Panel C. Small Firms								
Absolute Error in Option Value ($million)								
Mean	0.1830	0.2361	0.2592	0.2079	0.1902	0.1662	0.1789	0.1377
Median	0.0579	0.0586	0.0763	0.0467	0.0489	0.0472	0.0473	0.0388
75th pctl	0.1396	0.1621	0.1859	0.1373	0.1400	0.1223	0.1321	0.1008
95th pctl	0.5191	0.8462	1.0404	0.7222	0.6117	0.5319	0.6142	0.4281
Absolute Percentage Error in Option Value								
Mean	0.1049	0.1186	0.1283	0.1047	0.1047	0.0902	0.0973	0.0749
Median	0.0681	0.0808	0.0883	0.0719	0.0745	0.0598	0.0636	0.0478
75th pctl	0.1307	0.1634	0.1735	0.1440	0.1342	0.1149	0.1221	0.0935
95th pctl	0.3651	0.3943	0.4618	0.3494	0.3572	0.3292	0.3650	0.2544
Fraction positive	0.4034	0.1705	0.1818	0.2898	0.3409	0.3580	0.3352	0.3921

Consider first the impact of forecasting errors on option value for large firms (Panel A of Table 6). The mean absolute (percentage) error in option value is \$1.084 million (21.6%) for the 5-year historical forecast (HIS_5). The corresponding figure for the LM-VAR forecast is \$0.479 million (9.1%), which is 55.8%

(57.9%) smaller than the error for the historical forecast. For medium-sized firms (Panel B), the mean absolute (percentage) error in option value for the LM-VAR forecast is 33.2% (48.1%) smaller than for the historical forecast. For small firms (Panel C), the corresponding reduction in forecasting errors is 46.9% (41.6%). The reduction in option valuation error is thus substantial in all cases if the LM-VAR forecast is used instead of the historical forecast. Similar or even stronger improvement is observed when we examine the median, 75th, and 95th percentiles of the absolute (percentage) errors.

To further illustrate the economic significance of volatility forecasting errors, we assess the impact of forecasting errors on the firm's reported income due to option expensing. In order to do so, we calculate the total value of all options granted by each firm (to all employees including the CEO) during the year, first using the volatility forecast and then the realized volatility. We then determine the error in the estimated option value by calculating the percentage difference in the two option values. Finally, we translate both the option value and the absolute error as a fraction of the firm's net income in order to capture the impact of option expensing and volatility forecasting errors on reported income. Note that stock options are usually expensed over the vesting period rather than just in the fiscal year they are granted. During the initial phase of implementing option expensing, the annual option expense may vary substantially from year to year. Once the steady state is reached, the annual option expense is expected to be more stable and roughly equal to the value of all newly granted options in the same year. In our calculations here, we assume that all firms have already reached the steady state and estimate the annual option expense by simply calculating the total value of newly granted options. As we are interested in the relative impact of forecasting errors on option expensing across volatility forecasts, this simplification is unlikely to induce any bias in our inferences.

Table 7 presents the results on the impact of volatility forecasting errors on reported income. Take the large-firm sample (Panel A), for instance. As reported in the top half of the panel, the mean (median) ratio of total option value to net income varies from 23.6% to 27.6% (from 5.2% to 6.4%) depending on the choice of volatility forecasts. These numbers indicate that mandatory option expensing would result in substantial downward revisions to reported income for a large fraction of these firms. The choice of volatility forecasting methods also matters given the range of variations across volatility forecasts. As for the impact of forecasting errors on reported income (the bottom half of the panel), there are also wide variations across different volatility forecasts. Using the 5-year historical volatility forecast, the mean (median) absolute error in total option value is 3.7% (0.8%) of net income. In comparison, the corresponding figure is 1.8% (0.3%) if the LM-VAR forecast is used. As a result, the LM-VAR forecast cuts the impact of forecasting errors on reported income by about half relative to that of the historical forecast. For option valuation and expensing, the choice of volatility forecasts is thus quite important.

For medium-sized and small firms in our sample (Panels B and C of Table 7, respectively), the impact of forecasting errors on option expensing is even stronger. For medium-sized firms, the mean (median) ratio of total option value to net income varies from 46.5% to 49.4% (from 6.0% to 7.3%), depending on the choice

TABLE 7

The Impact of Forecasting Errors on Option Expensing

Table 7 summarizes the impact of forecasting errors on option expensing for alternative volatility forecasts. Both absolute errors and absolute percentage errors are reported for each firm-size group. HIS_1, HIS_3, and HIS_5 denote, respectively, the historical forecasts based on the 1-, 3-, and 5-year historical volatilities. Other forecasts are based on volatility models estimated using monthly volatility time series over a rolling historical sample period. LM denotes the long-memory model of the firm's monthly volatility series, while VAR denotes the bivariate vector autoregressive model of the joint monthly volatility series of the firm and the S&P 500 index. LM-VAR is the long-memory version of the VAR model. The total sample period is from 1962 to 2004, and the forecasting period is from June 1995 to December 2003.

Forecast	HIS_1	HIS_3	HIS_5	AR(1)	ARMA(1,1)	LM	VAR	LM-VAR
Panel A. Large Firms								
Option Expense as Percentage of Net Income (%)								
Mean	27.0	24.8	23.6	25.2	26.4	27.1	27.3	27.6
Median	6.3	5.6	5.2	5.8	6.2	6.3	6.4	6.4
75th pctl	14.4	12.6	12.2	13.2	13.6	13.7	14.2	14.6
95th pctl	51.1	47.1	46.3	47.4	50.8	52.1	52.3	51.9
Absolute Dollar Error as Percentage of Net Income (%)								
Mean	3.0	3.0	3.7	2.6	2.6	2.2	2.1	1.8
Median	0.4	0.7	0.8	0.6	0.4	0.4	0.3	0.3
75th pctl	1.2	1.5	1.8	1.4	1.2	0.9	0.9	0.8
95th pctl	4.9	6.1	8.2	4.6	4.7	4.3	4.0	3.3
Panel B. Medium-Sized Firms								
Option Expense as Percentage of Net Income (%)								
Mean	49.0	47.3	46.5	47.7	48.3	49.0	48.7	49.4
Median	7.3	6.3	6.0	6.6	6.9	7.0	6.8	7.2
75th pctl	21.5	19.3	19.3	20.8	21.4	21.5	21.3	21.6
95th pctl	152.0	144.0	139.4	144.3	149.4	147.6	147.2	148.1
Absolute Dollar Error as Percentage of Net Income (%)								
Mean	2.9	2.9	3.4	2.9	2.9	2.3	2.4	1.9
Median	0.6	0.6	0.7	0.6	0.6	0.5	0.5	0.4
75th pctl	1.9	1.7	2.0	1.6	1.6	1.4	1.5	1.3
95th pctl	9.1	9.2	10.9	10.3	9.2	7.7	7.7	6.4
Panel C. Small Firms								
Option Expense as Percentage of Net Income (%)								
Mean	122.5	118.0	118.0	120.2	120.0	121.0	120.3	123.5
Median	15.9	15.1	15.4	15.6	15.8	15.5	15.6	16.0
75th pctl	49.1	48.3	50.6	49.6	48.3	49.0	48.8	50.3
95th pctl	195.0	191.4	190.1	192.4	193.7	193.3	193.1	196.6
Absolute Dollar Error as Percentage of Net Income (%)								
Mean	5.1	8.2	8.2	6.4	6.8	5.9	6.3	3.6
Median	1.1	1.1	1.3	1.0	1.0	0.9	1.0	0.7
75th pctl	3.3	3.0	3.6	3.1	2.8	2.8	3.0	2.1
95th pctl	9.8	11.5	10.2	9.9	10.2	9.9	10.0	6.9

of volatility forecasts. The corresponding figures for small firms are from 118.0% to 123.5% (from 15.1% to 16.0%). Not surprisingly, the impact of forecasting errors on reported income is also larger for these firms. The mean (median) absolute error in total option value from using the 5-year historical volatility is 3.4% (0.7%) of net income for medium-sized firms and 8.2% (1.3%) for small firms. For the LM-VAR volatility method, the corresponding figure is 1.9% (0.4%) for medium-sized firms and 3.6% (0.7%) for small firms. Again, the LM-VAR forecast reduces the impact of forecasting errors on reported income by about half relative to that of the historical forecast.

VI. Comparable-Firm and Shrinkage Forecasts

Previous research (e.g., Alford and Boatsman (1995)) suggests that a shrinkage adjustment toward comparable-firm volatilities can improve the forecasting performance of historical volatility.[8] The basic idea is to combine a firm's historical volatility with those of comparable firms. As stock returns of comparable firms tend to exhibit a strong correlation, the median or average volatility of comparable firms is likely to be useful in establishing the long-term trend in volatility movement for these firms. In this section, we empirically evaluate the performance of both comparable-firm and shrinkage forecasts. The comparable-firm forecast is constructed as the median historical volatility of the comparable firms, while the shrinkage forecast is an equal-weighted average of the historical and comparable-firm forecasts.[9]

To search for comparable firms, we follow prior research (e.g., Lev (1983), Christie (1982), and Karolyi (1993)) by considering firms in the same industry with similar size and leverage. For each firm in our sample (called the target firm), we search for 10 comparable firms matched on industry, firm size, and leverage. We use the first two digits of the Standard Industrial Classification (SIC) code to identify industries. Firm size is proxied by market capitalization, while leverage is defined as the ratio of total debt to total assets. To qualify as a comparable firm, it must be in the same industry and have both its size and leverage within 50% of the target firm. If fewer than seven firms meet the requirement, the target firm is not included in the out-of-sample test on that particular forecasting date. On average, we only lose 6% of the sample firms due to insufficient number of comparable firms. If more than 10 firms meet the requirement, we select the top 10 firms based on how closely they match up with the target firm.[10] Since firm size and leverage change over time, a different set of comparable firms are likely matched to the same target firm on different dates.

Table 8 reports summary statistics of the monthly volatility series for both target and comparable firms. For comparable firms, we report the median volatility of the comparable firms. These statistics suggest that the volatilities of the target and comparable firms are matched quite well with similar mean, median, and other percentiles for all three size-based groups. More importantly, the correlation (ρ) between the two volatility series indicates substantial comovements between the two volatility series. As expected, the correlation is higher for large firms (with a mean of 0.291) than for small firms (with a mean of 0.220). In addition, the degree of fractional integration (d) reported in Table 8 suggests that the target and comparable firms share similar dynamic properties.

Table 9 summarizes the performance of comparable-firm and shrinkage forecasts in comparison with a number of alternative forecasts. To conserve space, we only tabulate the results for the 5-year forecasting horizon, which is the most

[8]For a more general discussion of the shrinkage methods and a survey of forecast combinations, see Timmermann (2006).

[9]We follow Alford and Boatsman (1995) and construct the comparable-firm forecast as the *median* historical volatility of the comparable firms. We also consider the *average* historical volatility of the comparable firms and unequally weighted average of historical and comparable-firm forecasts (e.g., [2/3, 1/3] or [1/3, 2/3]). Untabulated results indicate that none of these changes materially improves the forecasting performance.

[10]The quality of the match is proxied by a score based on the sum of squared percentage deviation in firm size and squared percentage deviation in leverage. We select the 10 firms with the lowest scores as our choice of comparable firms for the target firm.

TABLE 8

Summary and Diagnostic Statistics of Monthly Volatility (target and comparable firms)

In Table 8, we match each firm (the target firm) in the sample with 10 comparable firms based on industry, size, and leverage. For each monthly period, the comparable-firm volatility is the median volatility of the 10 matched firms. Summary statistics (including mean, median, and various percentiles) are first calculated at the firm level. The cross-sectional distributions of each statistic for the target and comparable firms are then reported. Diagnostic statistics are also reported, including the fractional integration coefficient (d), the coefficient of correlation with the S&P 500 index volatilities (ρ(SPX)), and the correlation of the volatility series between the target and comparable firms (ρ).

	Firm Type	5th Pctl	25th Pctl	Mean	Median	75th Pctl	95th Pctl
Panel A. Large Firms (N = 432)							
Mean	Target	0.150	0.223	0.280	0.268	0.334	0.410
	Comparable	0.158	0.238	0.290	0.276	0.328	0.419
Median	Target	0.121	0.204	0.257	0.246	0.306	0.423
	Comparable	0.129	0.217	0.259	0.250	0.324	0.434
d	Target	0.236	0.318	0.365	0.376	0.411	0.481
	Comparable	0.218	0.294	0.354	0.361	0.398	0.462
ρ(SPX)	Target	0.214	0.370	0.437	0.464	0.547	0.669
	Comparable	0.209	0.363	0.425	0.451	0.533	0.656
ρ		0.052	0.129	0.291	0.282	0.410	0.612
Panel B. Medium-Sized Firms (N = 506)							
Mean	Target	0.171	0.250	0.355	0.341	0.409	0.617
	Comparable	0.160	0.246	0.362	0.348	0.412	0.620
Median	Target	0.136	0.229	0.328	0.313	0.384	0.530
	Comparable	0.131	0.227	0.331	0.326	0.398	0.539
d	Target	0.200	0.305	0.363	0.366	0.398	0.471
	Comparable	0.184	0.277	0.349	0.348	0.405	0.468
ρ(SPX)	Target	0.121	0.221	0.322	0.326	0.420	0.549
	Comparable	0.139	0.220	0.345	0.342	0.430	0.541
ρ		0.047	0.089	0.258	0.266	0.365	0.549
Panel C. Small Firms (N = 1,001)							
Mean	Target	0.222	0.358	0.545	0.527	0.708	0.946
	Comparable	0.202	0.317	0.494	0.455	0.623	0.923
Median	Target	0.202	0.314	0.497	0.478	0.643	0.825
	Comparable	0.196	0.269	0.453	0.426	0.576	0.811
d	Target	0.186	0.312	0.378	0.382	0.440	0.489
	Comparable	0.165	0.294	0.375	0.376	0.431	0.480
ρ(SPX)	Target	0.087	0.147	0.199	0.194	0.284	0.459
	Comparable	0.079	0.143	0.215	0.228	0.322	0.490
ρ		0.043	0.082	0.220	0.219	0.307	0.459

relevant for option expensing. The results are similar for the other two forecasting horizons.[11] We evaluate 10 volatility forecasts, including three historical volatility-based forecasts (i.e., historical, comparable-firm, and shrinkage forecasts) and seven additional forecasts that incorporate either the LM, VAR, or both features. The horizon-matched historical forecast (HIS), the comparable-firm forecast (COM), and the shrinkage forecast (SHR) are constructed as previously described. The VAR_{COM} ($VAR_{S\&P}$) forecast is constructed from a bivariate VAR model that combines the volatility time series of the target firm and the comparable firms (the S&P 500 index). The VAR3 forecast is constructed from a

[11]One minor difference is that the shrinkage forecast is more (less) accurate than the historical forecast over the 5-year (1-year) horizon. This is likely due to the performance of the comparable-firm forecast. Although it is consistently worse than the historical forecast (across all forecasting horizons and firm-size groups), it is much worse at the 1-year horizon than at the 5-year horizon.

trivariate VAR model that combines the volatility time series of the target firm, the comparable firms, and the S&P 500 index. The LM forecast is constructed from the univariate volatility time series of the target firm, incorporating the LM feature. Finally, the three LM-VAR forecasts ($\text{LM-VAR}_{\text{COM}}$, $\text{LM-VAR}_{\text{S\&P}}$, and LM-VAR3) are similar to the corresponding VAR forecasts except that the LM feature is incorporated in each volatility time series.

TABLE 9

The Performance of Shrinkage and Comparable-Firm Forecasts

Table 9 reports summary statistics of the forecasting errors of alternative volatility forecasts at the 5-year forecasting horizon. HIS, COM, and SHR denote the horizon-matched historical, comparable-firm, and shrinkage forecasts, respectively. Other forecasts are based on volatility models estimated using monthly volatility time series over a rolling historical sample period. VAR_{COM} ($\text{VAR}_{\text{S\&P}}$) denotes the bivariate VAR model of target-firm and comparable-firm (S&P 500 index) volatilities. VAR3 denotes the trivariate VAR model of the target-firm, comparable-firm, and S&P 500 index volatilities. LM denotes the univariate long-memory model of the target firm's volatility series. $\text{LM-VAR}_{\text{COM}}$, $\text{LM-VAR}_{\text{S\&P}}$, and LM-VAR3 are the LM version of the corresponding VAR models. The last column reports the *t*-statistic, showing that the absolute error of the volatility forecast is equal to the horizon-matched historical forecast. The total sample period is from 1962 to 2004, and the forecasting period is from June 1995 to December 2003. N/A indicates not applicable.

	Absolute Error				Absolute Percentage Error				
Forecast	Mean	Median	75th Pctl	95th Pctl	Mean	Median	75th Pctl	95th Pctl	*t*-Test
Panel A. Large Firms									
HIS	0.086	0.071	0.114	0.216	0.284	0.261	0.380	0.667	N/A
COM	0.101	0.079	0.127	0.264	0.322	0.295	0.414	0.879	13.588
SHR	0.080	0.064	0.106	0.219	0.282	0.253	0.368	0.630	–2.446
VAR_{COM}	0.074	0.063	0.104	0.184	0.258	0.212	0.343	0.695	–9.653
$\text{VAR}_{\text{S\&P}}$	0.071	0.061	0.100	0.180	0.249	0.206	0.332	0.676	–11.836
VAR3	0.068	0.058	0.095	0.170	0.235	0.195	0.321	0.606	–14.936
LM	0.070	0.060	0.099	0.177	0.243	0.201	0.332	0.630	–13.391
$\text{LM-VAR}_{\text{COM}}$	0.066	0.057	0.093	0.166	0.233	0.194	0.311	0.621	–14.911
$\text{LM-VAR}_{\text{S\&P}}$	0.064	0.055	0.089	0.160	0.224	0.185	0.302	0.590	–16.400
LM-VAR3	0.062	0.053	0.087	0.156	0.215	0.179	0.294	0.555	–17.823
Panel B. Medium-Sized Firms									
HIS	0.090	0.067	0.118	0.263	0.269	0.248	0.371	0.664	N/A
COM	0.119	0.081	0.152	0.293	0.325	0.269	0.417	0.912	11.500
SHR	0.088	0.064	0.117	0.262	0.268	0.217	0.318	0.744	–0.322
VAR_{COM}	0.078	0.057	0.111	0.219	0.236	0.183	0.335	0.650	–10.258
$\text{VAR}_{\text{S\&P}}$	0.077	0.056	0.110	0.217	0.233	0.181	0.332	0.644	–10.817
VAR3	0.075	0.055	0.108	0.213	0.224	0.176	0.319	0.617	–12.647
LM	0.076	0.056	0.109	0.217	0.227	0.178	0.323	0.625	–12.047
$\text{LM-VAR}_{\text{COM}}$	0.074	0.054	0.105	0.209	0.225	0.173	0.317	0.626	–13.390
$\text{LM-VAR}_{\text{S\&P}}$	0.072	0.053	0.102	0.205	0.217	0.171	0.309	0.605	–13.829
LM-VAR3	0.070	0.051	0.100	0.199	0.209	0.165	0.299	0.583	–16.207
Panel C. Small Firms									
HIS	0.166	0.121	0.230	0.519	0.307	0.240	0.417	0.865	N/A
COM	0.196	0.143	0.267	0.552	0.374	0.264	0.453	1.054	9.784
SHR	0.157	0.112	0.216	0.499	0.299	0.227	0.392	0.844	–3.257
VAR_{COM}	0.142	0.104	0.201	0.413	0.268	0.214	0.380	0.722	–10.975
$\text{VAR}_{\text{S\&P}}$	0.143	0.104	0.204	0.416	0.270	0.215	0.384	0.729	–10.947
VAR3	0.137	0.100	0.194	0.401	0.257	0.204	0.371	0.693	–13.866
LM	0.132	0.097	0.187	0.388	0.248	0.197	0.355	0.672	–15.369
$\text{LM-VAR}_{\text{COM}}$	0.126	0.092	0.178	0.366	0.239	0.190	0.338	0.651	–16.633
$\text{LM-VAR}_{\text{S\&P}}$	0.127	0.093	0.179	0.369	0.241	0.192	0.340	0.656	–16.017
LM-VAR3	0.118	0.088	0.167	0.344	0.224	0.177	0.320	0.608	–18.859

Note that the three VAR forecasts (VAR_{COM}, $\text{VAR}_{\text{S\&P}}$, and VAR3) are expected to perform more poorly than the corresponding LM-VAR forecasts. We consider these VAR forecasts because they have some similarities with the shrinkage forecast. In particular, the VAR_{COM} forecast is constructed using the same two volatility time series (of the target and comparable firms) as the shrinkage forecast and imposes a relatively minimal model structure. Intuitively, the VAR_{COM} forecast can be considered as a generalized form of the shrinkage forecast. The advantage of the VAR_{COM} forecast is that the weights between historical and comparable-firm volatilities are determined by how the two volatility series correlate instead of the ad hoc scheme used in the shrinkage forecast. It is thus interesting to see how the VAR_{COM} forecast compares with the shrinkage forecast in predicting long-term volatility.

As reported in Table 9, the three historical volatility-based forecasts (i.e., historical, comparable-firm, and shrinkage forecasts) are all poor candidates for forecasting long-term volatility, with the shrinkage (comparable-firm) forecast slightly more (substantially less) accurate than the historical forecast. As expected, all seven LM-VAR-based forecasts are more accurate than the historical forecast, and the difference in forecasting errors is statistically significant at the 1% level in all cases. While the trivariate LM-VAR3 model provides the most accurate forecast, the two bivariate LM-VAR models (i.e., LM-$VAR_{S\&P}$ and LM-VAR_{COM}) are also good candidates for forecasting long-term volatility and come quite close to matching the forecasting performance of the trivariate LM-VAR3 model.

To focus on the shrinkage forecast, we further examine its forecasting performance in comparison with other directly related forecasts. First, the shrinkage forecast is consistently less accurate than the univariate LM forecast across all firm-size groups and forecasting horizons. Take the sample of medium-sized firms with the 5-year forecasting horizon (Panel B in Table 9), for instance. The LM forecast has a median absolute (percentage) error of 0.056 (0.178). In comparison, the corresponding error from the shrinkage forecast is 0.064 (0.217), or 14% (22%) larger. This is an interesting finding, as the univariate LM forecast is constructed using the target firm's volatility time series alone, while the shrinkage forecast is constructed using the volatility time series of both the target and comparable firms. The shrinkage adjustment is thus much less effective than the LM model in extracting relevant information from historical data to project future stock return volatility.

In addition, the shrinkage forecast is also consistently less accurate than the bivariate VAR_{COM} forecast across all firm-size groups and forecasting horizons. Take the sample of medium-sized firms (Panel B in Table 9), for example. The VAR_{COM} forecast has a median absolute (percentage) error of 0.057 (0.183). In comparison, the corresponding error from the shrinkage forecast is 0.064 (0.217), or 12.3% (18.6%) larger. Since both volatility forecasts are constructed using the same two volatility time series, the reported differences in forecasting performance highlight the importance of incorporating comovement between volatility series. The shrinkage adjustment does not fully take into account the correlation between the two volatility series and uses a simple average of the historical and comparable-firm forecasts.

VII. Robustness Analysis

It is important to recognize that the LM-VAR methodology is new to the accounting profession, and some resistance is inevitable due to the added complexity in its implementation. In this section, we present further analysis on the robustness of our findings and the likelihood of making the LM-VAR model more accessible to the accounting community.

A. Alternative Estimators of the Fractional Integration Parameter *d*

As discussed previously, we use the GPH estimator, as refined by Robinson (1995b), to estimate the fractional integration parameter d in our empirical analysis. While it is a simple estimator requiring only OLS regressions, it has well-known biases in estimated values. We thus wish to conduct a series of robustness analyses to analyze the impact of such bias on estimated d values and forecasting errors.

We first consider an improved estimator proposed by Deo and Hurvich (1998) and Hurvich and Deo (1999) that uses an alternative bandwidth for the log-periodogram ordinates. Instead of truncating the log-periodogram ordinates by setting m in equation (4) roughly equal to the square root of the number of observations in the time series (i.e., $m \approx \sqrt{T}$), Deo and Hurvich (1998) and Hurvich and Deo (1999) show that m should be chosen as $CT^{4/5}$ where the constant C is optimized using log-periodogram regression. Implementing the Hurvich-Deo GPH estimator for firms in our sample, we find the estimated d values to be only slightly higher than those reported in Table 2. For large firms, the mean (median) d value is now 0.368 (0.360) compared to the previously estimated value of 0.362 (0.359) (in Panel A of Table 2). A similar level of differences is also observed for medium-sized and small firms. Given such a small effect on estimated d values, we do not expect any material impact on forecasting performance. Indeed, untabulated results confirm that the forecasting performance of the LM model is virtually unchanged if the Robinson GPH estimator is replaced by the Hurvich-Deo GPH estimator.

Next, we evaluate whether shrinkage-adjusted d estimates, toward d of either the market index or comparable firms, may improve forecasting performance. We begin with an analysis of shrinkage adjustment toward the d of the S&P 500 index. We use the bivariate LM-VAR model of the volatilities of the firm and the S&P 500 index to carry out the analysis. We compare four estimates of d values—the univariate estimate from the firm's volatility series, the univariate estimate from the index volatility series, the average of the two univariate estimates from each volatility series, and the joint estimate using both series simultaneously. Note that the results in Tables 3–5 are based on the joint estimate of d value, which reflects the LM properties of each individual volatility series and the common dynamics of the two series. If shrinkage adjustment can indeed improve forecasting performance, we should see substantial differences in the results using the four different estimates. Untabulated results indicate that forecasting performance is not materially affected regardless of which estimate is used. Take the 5-year forecasting horizon, for example. If the univariate d of the firm's volatility is used, the mean absolute error is 0.060, 0.076, and 0.118 for large, medium-sized, and small firms, respectively. The magnitude of these errors is virtually identical to those reported in Tables 3–5 (0.061, 0.074, and 0.115, respectively). If the univariate d of the index volatility is used, the forecasting performance is slightly worse, with mean absolute error of 0.063, 0.077, and 0.120 for large, medium-sized, and small firms, respectively. This is expected, since we are predicting the firm's volatility rather than the market's volatility. Finally, if the average of the two univariate d values is used to forecast volatility, the mean absolute error is 0.062, 0.076, and 0.117 for large, medium-sized, and small firms, respectively. These errors are not materially different from those reported in Tables 3–5. These findings suggest that shrinkage adjustment toward volatility of the market index does not materially improve (or deteriorate) the forecasting performance. This is true for all forecasting horizons. In a parallel analysis of shrinkage adjustment toward the d of comparable firms, we find that it does not materially affect forecasting performance as long as the weight placed on the d of comparable-firm volatility is not too excessive. Forecasting performance is clearly worse when 100% of the weight is placed on the d of comparable-firm volatility, consistent with our findings in Section VI on comparable-firm forecasts.

B. Mixed-Horizon Historical Volatility Forecasts

How do we make the LM-VAR methodology more accessible to the accounting profession? One possible solution is Corsi's (2004) mixed-horizon realized volatility model. It is based on historical volatilities, which are already well understood and accepted by the accounting profession, and yet it has the built-in flexibility to reflect the LM feature of the time series. By combining volatilities over different horizons (e.g., lagged 1-month, 6-month, and 1-year realized volatilities), his so-called heterogeneous autoregressive (HAR) volatility model may capture both short- and long-term components in volatility. Since it is just a linear combination of lagged historical volatilities with different horizons, the HAR model is straightforward to estimate using OLS regressions. For 1-day to 2-week forecasting horizons, Corsi (2004) finds that the HAR model performs comparably with, and sometimes even outperforms, the LM model. In the context of this paper, we need to verify whether the good performance of the HAR model extends over longer forecasting horizons. If it does, the HAR model should provide a much simpler method to characterize long memory and make it easier for accounting practitioners to accept LM-based forecasts.

To investigate the effectiveness of Corsi's (2004) approach in our setting, we implement the HAR model with three mixed horizons. We use 1- and 6-month realized volatilities to capture short-term and medium-term components in volatility. To capture the long-term component, we use either 1-, 3-, or 5-year realized volatility. We thus consider three specifications of the HAR model, HAR(3,1), HAR(3,3), and HAR(3,5), with the two indexes denoting the number of horizons and the longest horizon (in years) of the realized volatilities, respectively. These specifications are summarized in the following regression equation:

$$\sigma_t \quad = \quad \alpha + \beta_1 \sigma_{t-1} + \beta_2 \sigma_{t-1,t-6} + \beta_3 \sigma_{t-1,t-m} + \varepsilon_t,$$

where σ_t is the realized volatility in month t and $\sigma_{t-1,t-m}$ is the realized volatility over an m-month period immediately before month t. As before, all realized volatilities are calculated using daily returns, and the firm subscript is suppressed for convenience.

To empirically test the HAR models, we focus on a subsample of 308 firms that have at least 40 years of daily return data. This subsample includes 194 large firms, 65 medium-sized firms, and 49 small firms. The forecasting period is the 15-year period from 1990 to 2004. On each forecasting date, the HAR model is estimated using the three lagged realized volatility series over the past 20 years. For each of these 308 firms, we fit and compare the forecasting performance of eight volatility models including the horizon-matched historical volatility, AR(1), ARMA(1,1), LM, as well as the three HAR models. In order to focus on the effectiveness of the Corsi (2004) approach in replicating the LM feature, we use the univariate LM forecast as the benchmark and compare the performance of all other volatility models to the LM model. The results are summarized in Table 10.

TABLE 10

Corsi's Mixed-Horizon Historical Volatility Model

In Table 10, the forecasting performance of Corsi's mixed-horizon historical volatility model is compared with alternative models. HIS_1, HIS_3, and HIS_5 denote, respectively, historical forecasts based on the 1-, 3-, and 5-year historical volatilities. Other forecasts are based on volatility models estimated using monthly volatility time series over a rolling historical sample period as the forecasting date moves forward. LM denotes the long-memory model of the firm's monthly volatility series. HAR(3,1), HAR(3,3), and HAR(3,5) are volatility forecasts based on the mixed-horizon historical volatility model with three lagged realized volatilities. While 1- and 6-month horizons are used for the first two lagged realized volatilities, the horizon of the last lagged realized volatility varies from 1 to 5 years. The last column reports the t-statistic, showing that the absolute error of the volatility forecast is equal to the LM forecast. The sample contains 308 common stocks that have at least 480 monthly observations and remain in the CRSP database at the end of 2004. The total sample period is from 1962 to 2004, and the forecasting period is from June 1990 to December 2003. N/A indicates not applicable.

	Absolute Error				Absolute Percentage Error				
Forecast	Mean	Median	75th Pctl	95th Pctl	Mean	Median	75th Pctl	95th Pctl	t-Test
Panel A. 1-Year Forecasting Horizon									
HIS_1	0.055	0.037	0.071	0.158	0.175	0.139	0.240	0.445	5.416
AR(1)	0.061	0.043	0.078	0.179	0.201	0.160	0.276	0.523	7.314
ARMA(1,1)	0.063	0.045	0.083	0.183	0.214	0.173	0.288	0.558	8.914
LM	0.049	0.032	0.064	0.147	0.138	0.123	0.216	0.404	N/A
HAR(3,1)	0.053	0.034	0.070	0.150	0.142	0.130	0.221	0.409	2.061
HAR(3,3)	0.055	0.035	0.071	0.156	0.145	0.133	0.220	0.418	3.409
HAR(3,5)	0.056	0.035	0.074	0.155	0.147	0.134	0.223	0.422	3.448
Panel B. 3-Year Forecasting Horizon									
HIS_3	0.059	0.041	0.075	0.169	0.195	0.152	0.262	0.526	5.776
AR(1)	0.065	0.044	0.080	0.181	0.197	0.159	0.269	0.497	5.729
ARMA(1,1)	0.067	0.049	0.082	0.185	0.200	0.167	0.266	0.489	8.347
LM	0.047	0.033	0.061	0.146	0.166	0.131	0.227	0.441	N/A
HAR(3,1)	0.052	0.036	0.068	0.155	0.172	0.141	0.236	0.463	3.092
HAR(3,3)	0.051	0.034	0.067	0.151	0.170	0.137	0.233	0.465	1.538
HAR(3,5)	0.051	0.035	0.067	0.154	0.171	0.139	0.235	0.464	2.985
Panel C. 5-Year Forecasting Horizon									
HIS_5	0.058	0.039	0.073	0.177	0.196	0.148	0.259	0.551	4.531
AR(1)	0.063	0.043	0.081	0.179	0.191	0.150	0.255	0.513	5.726
ARMA(1,1)	0.064	0.045	0.084	0.179	0.192	0.151	0.251	0.507	6.809
LM	0.050	0.034	0.065	0.152	0.171	0.132	0.236	0.486	N/A
HAR(3,1)	0.056	0.037	0.071	0.166	0.179	0.146	0.245	0.520	3.157
HAR(3,3)	0.053	0.037	0.069	0.163	0.174	0.139	0.246	0.497	4.137
HAR(3,5)	0.053	0.036	0.068	0.162	0.175	0.139	0.245	0.499	2.836

As expected, the LM forecast performs the best among the eight volatility models across all forecasting horizons. As indicated by the t-statistic (last column in Table 10), the difference in forecasting errors is statistically significant in all but one case. More importantly, the HAR models provide more accurate forecasts than either the historical volatility, AR(1), or ARMA(1,1) models across nearly all performance measures and forecasting horizons. Compared to the LM forecast, the HAR forecasts are still less accurate but come close to matching it in some cases. This is especially true for the horizon-matched HAR model, which includes the horizon-matched historical volatility as one of the three historical volatility series (e.g., HAR(3,5) for the 5-year forecasting horizon). In fact, the forecasting error of the HAR(3,3) model is not statistically different from the corresponding error of the LM model, as the t-statistic indicates. The HAR approach is thus quite promising and may be a useful tool in the transition from historical forecast to LM-VAR forecast. Further research is needed to investigate the optimal combination of lagged realized volatilities and how the model can be generalized in a multivariate setting.

C. Volatility Metric

In measuring forecasting errors, we use standard deviation (of stock returns) as our volatility metric. Although this metric is consistent with the standard practice in the option pricing literature, it may not necessarily be the best volatility metric to evaluate volatility forecasts in light of Patton's (2007) recent research on performance evaluation. We thus conduct further robustness analysis and examine whether our inferences on forecasting performance are materially affected if variance or logarithm of variance is used as the volatility metric. Historical forecasts of variance or its logarithm are readily calculated from the corresponding standard deviation. For all other forecasts, the time-series model must be reestimated using the new volatility metric in order to construct the various volatility forecasts. We do this for all 2,066 firms in our sample and replicate the analysis in Tables 3–5 for variance and logarithm of variance. For brevity, we only report the results (in Table 11) for medium-sized firms at the 5-year forecasting horizon. The results for other firm-size groups and/or forecasting horizons are qualitatively similar.

As reported in Table 11, the LM-VAR model remains the most accurate in forecasting long-term volatility. Its forecasting error is consistently the smallest regardless of the performance measure used. This is true whether we use variance or logarithm of variance as the volatility metric, similar to the results in Table 4, where standard deviation is used as the volatility metric. Other inferences remain unchanged as well, and we continue to find evidence that historical volatility is a poor forecast for long-term volatility. AR(1) and ARMA(1,1) do not provide much improvement beyond the historical forecast, and both LM and VAR features are important in improving long-term forecasting performance. The t-statistics reported in the last column of the table further confirm the statistical significance of these results. Our empirical findings are thus robust to the choice of volatility metric.

TABLE 11

Forecasting Performance Using Other Volatility Metrics

Table 11 reports the performance of alternative volatility forecasts when either variance or logarithm of variance is used as the volatility metric. Only medium-sized firms are included, and the forecasting horizon is 5 years. HIS_1, HIS_3, and HIS_5 denote, respectively, the historical forecasts based on the 1-, 3-, and 5-year historical volatilities. Other forecasts are based on volatility models estimated using monthly volatility time series over a rolling historical sample period. LM denotes the long-memory model of the firm's monthly volatility series, while VAR denotes the bivariate vector autoregressive model of the joint monthly volatility series of the firm and the S&P 500 index. LM-VAR is the long-memory version of the VAR model. The last column reports the *t*-statistic, showing that the absolute error of the volatility forecast is equal to the horizon-matched historical forecast. The total sample period is from 1962 to 2004, and the forecasting period is from June 1995 to December 2003. N/A indicates not applicable.

	Absolute Error				Absolute Percentage Error				
Forecast	Mean	Median	75th Pctl	95th Pctl	Mean	Median	75th Pctl	95th Pctl	*t*-Test
Panel A. Variance Forecasts									
HIS_1	0.095	0.053	0.118	0.329	0.480	0.433	0.657	1.164	–0.503
HIS_3	0.098	0.053	0.114	0.337	0.486	0.418	0.642	1.293	0.666
HIS_5	0.096	0.051	0.113	0.331	0.478	0.402	0.624	1.266	N/A
AR(1)	0.093	0.050	0.110	0.328	0.460	0.399	0.615	1.184	–1.545
ARMA(1,1)	0.095	0.052	0.115	0.327	0.477	0.426	0.644	1.190	–0.699
LM	0.086	0.047	0.102	0.298	0.429	0.386	0.588	1.029	–4.547
VAR	0.085	0.046	0.102	0.287	0.427	0.382	0.582	1.031	–4.767
LM-VAR	0.078	0.042	0.092	0.266	0.385	0.346	0.527	0.922	–7.441
Panel B. Log Variance Forecasts									
HIS_1	0.545	0.450	0.770	1.349	0.242	0.191	0.346	0.653	3.226
HIS_3	0.536	0.439	0.744	1.345	0.237	0.187	0.342	0.634	1.915
HIS_5	0.522	0.421	0.718	1.343	0.229	0.176	0.329	0.621	N/A
AR(1)	0.504	0.414	0.700	1.274	0.223	0.176	0.320	0.604	–2.397
ARMA(1,1)	0.534	0.443	0.753	1.346	0.237	0.189	0.340	0.636	1.913
LM	0.482	0.404	0.679	1.193	0.214	0.170	0.307	0.578	–5.069
VAR	0.480	0.404	0.674	1.182	0.214	0.171	0.305	0.575	–5.353
LM-VAR	0.427	0.363	0.595	1.054	0.190	0.153	0.270	0.508	–10.026

VIII. Conclusions

In this paper, we empirically investigate the forecasting performance of historical forecasts in comparison with a number of alternative volatility forecasts. Unlike prior researchers, we focus on long-horizon volatility forecasting and its impact on the valuation and expensing of stock options. Our empirical analysis suggests that horizon-matched historical volatility is a poor forecast for long-term stock return volatility and shrinkage adjustment toward common factors (e.g., volatility of comparable firms or market index) only slightly improves its performance. Forecasting performance can be improved substantially, however, if the volatility forecast incorporates both long memory and comovements with common factors. Our results indicate that the impact of forecasting errors on option expensing is reduced by approximately 50% if the LM-VAR forecast is used instead of the historical forecast. Although volatility forecasting remains a challenging task, our research provides insight on the key factors influencing forecasting performance and how volatility forecasts should be constructed.

Our research also provides a better approach to evaluating managerial incentives to manipulate its estimate of stock return volatility. Prior research typically detects the use of managerial discretion in volatility estimation by comparing the reported volatility and horizon-matched historical volatility. An opportunistic use of managerial discretion is recognized if the reported volatility represents a downward adjustment from historical volatility. However, this is not necessarily a correct assessment, since historical volatility is a rather poor forecast for future volatility. It is also quite possible that management is using private information to update historical experience in order to make a more accurate volatility forecast. The LM-VAR forecast offers a better benchmark for assessing managerial discretion in volatility forecasting.

References

Aboody, D.; M. E. Barth; and R. Kasznik. "Do Firms Understate Stock Option-Based Compensation Expense Disclosed under SFAS 123?" *Review of Accounting Studies*, 11 (2006), 429–461.

Alford, A. W., and J. R. Boatsman. "Predicting Long-Term Stock Return Volatility: Implications for Accounting and Valuation of Equity Derivatives." *Accounting Review*, 70 (1995), 599–618.

Andersen, T. G.; T. Bollerslev; P. F. Christoffersen; and F. X. Diebold. "Volatility and Correlation Forecasting." In *Handbook of Economic Forecasting*, 1, G. Elliot, C. W. J. Granger, and A. Timmermann, eds. Amsterdam: North Holland (2006), 777–878.

Andersen, T. G.; T. Bollerslev; F. X. Diebold; and H. Ebens. "The Distribution of Realized Stock Return Volatility." *Journal of Financial Economics*, 61 (2001), 43–76.

Andersen, T. G.; T. Bollerslev; F. X. Diebold; and P. Labys. "The Distribution of Realized Exchange Rate Volatility." *Journal of the American Statistical Association*, 96 (2001), 42–55.

Andersen, T. G.; T. Bollerslev; F. X. Diebold; and P. Labys. "Modeling and Forecasting Realized Volatility." *Econometrica*, 71 (2003), 579–625.

Baillie, R. T.; T. Bollerslev; and H. O. Mikkelsen. "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, 74 (1996), 3–30.

Balsam, S.; H. A. Mozes; and H. A. Newman. "Managing *Pro Forma* Stock Option Expense under SFAS No. 123." *Accounting Horizons*, 17 (2003), 31–45.

Bartov, E.; P. S. Mohanram; and D. Nissim. "Stock Option Expense, Forward-Looking Information, and Implied Volatilities of Traded Options." Working Paper, Columbia University and New York University (2004).

Black, F., and M. Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, 81 (1973), 637–654.

Bollerslev, T., and H. O. Mikkelsen. "Modeling and Pricing Long Memory in Stock Market Volatility." *Journal of Econometrics*, 73 (1996), 151–184.

Bollerslev, T., and H. O. Mikkelsen. "Long-Term Equity Anticipation Securities and Stock Market Volatility Dynamics." *Journal of Econometrics*, 92 (1999), 75–99.

Carpenter, J. N. "The Exercise and Valuation of Executive Stock Options." *Journal of Financial Economics*, 48 (1998), 127–158.

Carpenter, J. N., and B. Remmers. "Executive Stock Option Exercises and Inside Information." *Journal of Business*, 74 (2001), 513–534.

Christie, A. A. "The Stochastic Behavior of Common Stock Variances: Value, Leverage, and Interest Rate Effects." *Journal of Financial Economics*, 10 (1982), 407–432.

Corsi, F. "A Simple Long Memory Model of Realized Volatility." Working Paper, University of Lugano and Swiss Finance Institute (2004).

Deo, R. S., and C. M. Hurvich. "Linear Trend with Fractionally Integrated Errors." *Journal of Time Series Analysis*, 19 (1998), 379–397.

Deo, R. S., and C. M. Hurvich. "On the Log Periodogram Regression Estimator of the Memory Parameter in Long Memory Stochastic Volatility Models." *Econometric Theory*, 17 (2001), 686–710.

Deo, R.; C. Hurvich; and Y. Lu. "Forecasting Realized Volatility Using a Long-Memory Stochastic Volatility Model: Estimation, Prediction and Seasonal Adjustment." *Journal of Econometrics*, 131 (2006), 29–58.

Ding, Z.; C. W. J. Granger; and R. F. Engle. "A Long Memory Property of Stock Market Returns and a New Model." *Journal of Empirical Finance*, 1 (1993), 83–106.

French, K. R; G. W. Schwert; and R. F. Stambaugh. "Expected Stock Returns and Volatility." *Journal of Financial Economics*, 19 (1987), 3–29.

Geweke, J., and S. Porter-Hudak. "The Estimation and Application of Long Memory Time Series Models." *Journal of Time Series Analysis*, 4 (1983), 221–238.

Heath, C.; S. Huddart; and M. Lang. "Psychological Factors and Stock Option Exercise." *Quarterly Journal of Economics*, 114 (1999), 601–627.

Hemmer, T.; S. Matsunaga; and S. Shevlin. "Optimal Exercise and the Cost of Granting Employee Stock Options with a Reload Provision." *Journal of Accounting Research*, 36 (1998), 231–255.

Hodder, L. D; W. J Mayew; M. L. McAnally; and C. D. Weaver. "Employee Stock Option Fair-Value Estimates: Do Managerial Discretion and Incentives Explain Accuracy?" Working Paper, Texas A&M University (2006).

Huddart, S., and M. Lang. "Employee Stock Option Exercises: An Empirical Analysis." *Journal of Accounting and Economics*, 21 (1996), 5–43.

Huddart, S., and M. Lang. "Information Distribution within Firms: Evidence from Stock Option Exercises." *Journal of Accounting and Economics*, 34 (2003), 3–31.

Hurvich, C. M., and R. S. Deo. "Plug-In Selection of the Number of Frequencies in Regression Estimates of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis*, 20 (1999), 331–341.

Hurvich, C. M.; R. Deo; and J. Brodsky. "The Mean Squared Error of Geweke and Porter-Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis*, 19 (1998), 19–46.

Johnston, D. "Managing Stock Option Expense: The Manipulation of Option-Pricing Model Assumptions." *Contemporary Accounting Research*, 23 (2006), 395–425.

Karolyi, G. A. "A Bayesian Approach to Modeling Stock Return Volatility for Option Valuation." *Journal of Financial and Quantitative Analysis*, 28 (1993), 579–594.

Lev, B. "Some Economic Determinants of Time-Series Properties of Earnings." *Journal of Accounting and Economics*, 5 (1983), 31–48.

Merton, R. C. "The Theory of Rational Option Pricing." *Bell Journal of Economics and Management Science*, 4 (1973), 141–183.

Merton, R. C. "On Estimating the Expected Return on the Market: An Exploratory Investigation." *Journal of Financial Economics*, 8 (1980), 323–361.

Patton, A. J. "Volatility Forecast Comparison Using Imperfect Volatility Proxies." Working Paper, London School of Economics (2007).

Pong, S.; M. B. Shackleton; S. J. Taylor; and X. Xu. "Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models." *Journal of Banking and Finance*, 28 (2004), 2541–2563.

Poon, S.-H., and C. W. J. Granger. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature*, 41 (2003), 478–539.

Poterba, J. M., and L. H. Summers. "The Persistence of Volatility and Stock Market Fluctuations." *American Economic Review*, 76 (1986), 1142–1151.

Robinson, P. M. "Gaussian Semiparametric Estimation of Long Range Dependence." *Annals of Statistics*, 23 (1995a), 1630–1661.

Robinson, P. M. "Log-Periodogram Regression of Time Series with Long Range Dependence." *Annals of Statistics*, 23 (1995b), 1048–1072.

Schwert, G. W. "Why Does Stock Market Volatility Change Over Time?" *Journal of Finance*, 44 (1989), 1115–1153.

Taqqu, M. S., and V. Teverovsky. "Robustness of Whittle-Type Estimators for Time Series with Long-Range Dependence." *Stochastic Models*, 13 (1997), 723–757.

Timmermann, A. "Forecast Combinations." In *Handbook of Economic Forecasting, Vol. 1*, G. Elliott, C. W. J. Granger, and A. Timmermann, eds. Amsterdam: North-Holland (2006), 135–196.

West, K. "Forecast Evaluation." In *Handbook of Economic Forecasting, Vol. 1*, G. Elliott, C. W. J. Granger, and A. Timmermann, eds. Amsterdam: North-Holland (2006), 99–134.