

# Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities

**Christopher G. Lamoureux**

Washington University in St. Louis

**William D. Lastrapes**

University of Georgia

*We examine the behavior of measured variances from the options market and the underlying stock market. Under the joint hypotheses that markets are informationally efficient and that option prices are explained by a particular asset pricing model, forecasts from time-series models of the stock-return process should not have predictive content given the market forecast as embodied in option prices. Both in-sample and out-of-sample tests suggest that this hypothesis can be rejected. Using simulations, we show that biases inherent in the procedure we use to imply variances cannot explain this result. Thus, we provide evidence inconsistent with the orthogonality restrictions of option pricing models that assume that variance risk is unpriced. These results also have implications for optimal variance forecast rules.*

According to the Black and Scholes (1973) model of option valuation, equilibrium option prices are determined by the absence of arbitrage profits. This con-

---

We are grateful to Don Andrews, Kerry Back, Phil Dybvig, Ravi Jagannathan, Andrew Lo (the editor), Chester Spatt (the executive editor), and an anonymous referee for numerous suggestions. But we wish to absorb all culpability. Address correspondence to Christopher G. Lamoureux, John M. Olin School of Business, Washington University in St. Louis, Campus Box 1133, One Brookings Drive, St. Louis, MO 63130-4899.

*The Review of Financial Studies* 1993 Volume 6, number 2, pp. 293-326  
© 1993 The Review of Financial Studies 0893-9454/93/\$1.50

dition depends on the assumption that the variance of the underlying stock returns is constant over time or deterministically changing through time. The model is a powerful economic tool because market behavior can be understood without explicitly specifying and estimating preferences of agents.

A problem with the empirical implementation of the Black–Scholes model is that the variance assumption is inconsistent with the data. Since stochastic volatility is manifest in time-series models of stock returns as well as in the empirical variances implied from the Black–Scholes model itself, models of option pricing have been developed in which the variance of the underlying asset returns varies randomly through time. Hull and White (1987) derive a closed-form solution for European call option prices under the assumption that volatility risk is unpriced. They show that, given certain conditions on the stochastic process governing underlying returns, the option price equals the expected value of the Black–Scholes price over the distribution of average variance.

While the stochastic volatility generalization has been shown by Hull and White (1987) and others to improve the explanatory power of the Black–Scholes model, the full implications of the stochastic volatility option pricing models have not been adequately tested. In particular, a clear test of whether the strong assumption of market indifference to volatility risk is consistent with the data is missing in the empirical finance literature. Although this assumption may be unattractive from a theoretical perspective, it does afford simplification of both valuation and variance extraction, and the model lends itself to unambiguous empirical testing. We examine this issue by testing an important implication of the *class* of models represented by Hull and White (1987) in which volatility risk is unpriced. If option markets are informationally efficient, then information available at the time market prices are set cannot be used to predict actual return variance better than the variance forecast embedded in the option price, which represents the subjective expectation of the market. That is, the forecast error of the subjective expectation should be orthogonal to all available information.

To test this orthogonality restriction, we interpret the variance implied from equating the observed market option price to the Hull and White model price as the market's assessment of return variance. However, implying a variance from the closed-form expression of Hull and White may distort, or bias, the market forecast, even assuming that the joint null hypothesis is true since the Black–Scholes formula is nonlinear (Jensen's inequality) and the variance and stock processes may be instantaneously correlated. We carefully calibrate simulations of the stock-price and variance processes for each of the

stocks in the sample to examine the extent to which such bias exists. The simulations demonstrate that this distortion is trivial for our sample of at-the-money options. Furthermore, we couch our statistical inference in the context of the simulations to account for the possible sources of distortion. This procedure enables us to refer to our empirical exercise as a formal test of an asset pricing model. If we did not control for the distortion or couch our inference in terms of this distortion, our exercise would simply be an examination of the extent to which filtering the options and stock-price data with the Black-Scholes model has information.

We test the orthogonality restriction for at-the-money call options on individual stocks by comparing the forecast performance of the implied variance from the model with time-series representations of stock-return volatility. We use a simple time-series model of serial dependence in volatility, the generalized autoregressive conditional heteroskedasticity (GARCH) process of Engle (1982) and Bollerslev (1986), to capture available information that can explain the evolution of return variance. Two types of tests of the orthogonality restriction are performed: in-sample and out-of-sample. The in-sample, regression-based tests incorporate the implied variance into the GARCH equation of the return process to measure the marginal predictive power of past information on variance. We then compare the out-of-sample forecast performance of the implied variance with time-series models using standard measures of average forecast error and encompassing regressions. The out-of-sample encompassing analysis helps us to explore an important issue: regardless of the outcome of tests of the orthogonality restrictions, does filtering market data through the model provide information about the future evolution of return variance that is not evident in the past time series of stock returns? This issue is relevant to constructing optimal variance forecast rules.

Our research design differs from most previous tests of stochastic volatility option pricing models. Melino and Turnbull (1990), for example, test such a model for foreign exchange options. They estimate a stochastic volatility process for the underlying asset, then price options on this asset using the parameters from the process and the option pricing model. This price is found to predict the actual option price better than a constant variance price. However, they also find that current information can explain some of the model's forecast error. By testing the implications of the option pricing model in measures of variance rather than option price, we exploit different information than Melino and Turnbull to test the orthogonality restrictions. More importantly, our forecast comparisons are performed ultimately out-of-sample. Out-of-sample analysis is more natural than in-sample analysis for scrutinizing models that depend on the infor-

mation set of agents at any point in time, since the econometrician's conditioning set is a proper subset of the information available to the agents in the economy.

Our work in this article complements concurrent, independent research by Day and Lewis (1992), which also compares implied volatilities from option pricing models with GARCH models. Essentially four major differences occur in the articles. (1) We broaden the data sample by using daily data on individual stocks, whereas Day and Lewis use weekly data on stock indices. (2) As noted, we provide simulation evidence to quantify the extent to which implying variances under Black–Scholes distorts the actual variance forecast *under the null hypothesis* for each stock in our sample. Thus, unlike Day and Lewis we interpret our analysis as a formal test of a specific asset pricing model. (3) Given the daily frequency of data, we are better able than Day and Lewis to adjust for the inconsistencies between the forecast horizons of the time-series models and the maturity of the options in the sample. (4) Perhaps most importantly, we take considerable care to purge problems related to measurement error by effectively using intraday data to construct the daily series. Day and Lewis, for example, use closing prices in both option and stock markets, which do not even close at the same time. To attenuate this problem, they imply the price of the underlying asset from the option price, under the assumption that the model which they are testing is true. Because our data are carefully mapped into our research design, we substantially reduce these errors-in-variables problems.

Although our empirical strategy is discussed in the context of the model developed by Hull and White (1987), we wish to emphasize that we do not attempt to test all of the implications of their model. These authors were interested in explaining the biases of the Black–Scholes restrictions when volatility is stochastic as a function of moneyness and maturity of the option. This aim requires exploiting a broader range of options than we use to infer moments beyond the mean of the subjective distribution of variance. Our strategy is to focus solely on the orthogonality restrictions implied by this *class* of models, hence our reliance on the market's expectation of return variance. In other words, we are not testing the null hypothesis that Black–Scholes is true against the stochastic variance alternative. We are treating the stochastic volatility option pricing model as a special case of more general models of asset pricing that include the possibility of pricing volatility risk.

Both the in-sample tests and the out-of-sample encompassing tests suggest that, while the implied variance helps predict future volatility, the orthogonality restriction of the joint null hypothesis is rejected. One possible reason for the rejection of the null is that volatility risk

is priced. Therefore, further attempts to learn from the data should explicitly model a risk premium on the variance process, as in Heston (1993), for example. Assuming constant relative-risk aversion and using a Fourier inversion formula [Stein and Stein (1991)], Heston has obtained a closed-form expression for an option on an asset with stochastic volatility that allows for the volatility process to be priced.

In the following section, we lay out the analytical framework of the study. We describe the Hull and White (1987) model, state explicitly the orthogonality restrictions that it implies jointly with the assumption of market efficiency, and outline the general test strategy. We also characterize stochastic volatility in the data by using the GARCH model, and we report simulation results to quantify certain biases in our implied variances. In Section 2, we describe the options data used in this study, the specific in-sample and out-of-sample tests of the orthogonality restriction undertaken, and the results from these tests. We offer conclusions in Section 3.

## 1. Analytical Framework

### 1.1 Theoretical model and test strategy

The framework for the empirical analysis in this article is the model developed by Hull and White (1987), which is an application of Garman (1976). The model represents the class of stochastic volatility options pricing models, including those of Scott (1987), Wiggins (1987), and Johnson and Shanno (1987), that assumes volatility risk does not affect the option price. The Hull–White (HW) model is based upon the following continuous-time process for the underlying stock:

$$dS = \phi S dt + \sqrt{V} S dw, \quad (1)$$

$$dV = \mu V dt + \xi V dz, \quad (2)$$

where  $S$  is the stock price, and the Brownian motions  $dw$  and  $dz$  have an instantaneous correlation of  $\rho$ . Under the assumption that volatility risk is not priced and  $\rho = 0$ , a call option on this stock at time  $t$  will be priced as

$$p_t = \int \text{BS}(\bar{V}_t) b(\bar{V}_t | I_t) d\bar{V}_t = \text{E}[\text{BS}(\bar{V}_t) | I_t], \quad (3)$$

where

$$\bar{V}_t = \frac{1}{T-t} \int_t^T V_i di,$$

$b(\bar{V}_t | V_t)$  is the density of  $\bar{V}_t$  conditional on the current  $V_t$ ,  $T$  is the expiration date of the option,  $I_t$  is the information set at time  $t$ , and

$BS(\cdot)$  is the Black–Scholes pricing formula. Thus, the HW price is the mean Black–Scholes price, evaluated over the conditional distribution of average variance  $\bar{V}_t$ .

To price options according to (3), market participants must form a subjective conditional density on  $\bar{V}_t$ . If these participants are rational, then the subjective density of the market equals the actual density,  $b(\bar{V}_t | I_t)$ . The focus of our study is on the mean of this distribution:

$$E(\bar{V}_t | I_t) = \int \bar{V}_t b(\bar{V}_t | I_t) d\bar{V}_t.$$

The average variance can be decomposed as

$$\bar{V}_t = E(\bar{V}_t | I_t) + \mu_t,$$

where  $\mu_t$  is orthogonal to the conditional expectation and, therefore, to any information available to agents up to time  $t$ . Thus, if option market participants are rational (in the mean sense), then the subjective mean,  $E^s(\bar{V}_t | I_t)$ , equals the actual mean,  $E(\bar{V}_t | I_t)$ , and the subjective forecast error,  $\mu_t^s = \bar{V}_t - E^s(\bar{V}_t | I_t)$ , is orthogonal to all available information. This orthogonality condition is tested in this article.

The market's conditional expectation of variance is not directly observable. But given market prices on options, the HW model implies a variance that can be used to estimate the expectation. Because the variance of the underlying asset is the only unknown argument in the Black–Scholes formula, the implied variance (discussed in more detail later) is the value of that argument that equates the market price to the theoretical Black–Scholes price. It is evident from (3) that the implied variance is not in general a good predictor of the market's evaluation of variance, since the conditional density  $b$  is unaccounted for. However, Cox and Rubinstein (1985, p. 218, figure 5, 6) show that the Black–Scholes formula is essentially a linear function of the standard deviation for at-the-money options, so that  $E[BS(\bar{V}_t) | I_t]$  approximately equals  $BS[E(\bar{V}_t) | I_t]$ .<sup>1</sup> For this reason, we use a sample of at-the-money options and therefore interpret the implied variance as the market's assessment of average stock-return variance over the remaining life of the option,  $E^s(\bar{V}_t | I_t)$ , under the assumption that the HW model is valid.

We test whether the forecast error constructed from  $\bar{V}_t$  and the implied variance is orthogonal to past information, using in-sample and out-of-sample tests. This orthogonality condition is based on a joint hypothesis: (a) option markets are informationally efficient, so

<sup>1</sup> While this linearity has formally been expressed in terms of the standard deviation, our analysis is conducted in terms of the variance. Simulations reported in Section 1.3 show that the linearity applies to the variance as well, for our data. Feinstein (1989) analyzes implied volatilities and confirms the linearity documented by Cox and Rubinstein.

observed option prices contain all relevant, available information; and (b) the HW pricing model is correct, so the implied variances are valid estimates of the subjective variance of the market.

The test design is analogous to tests in the international finance literature that examine the hypothesis that forward exchange rates are optimal predictors of future spot exchange rates [e.g., Hansen and Hodrick (1980)]. If the foreign exchange market is informationally efficient, then the difference between the realized future spot rate and the market's subjective expectation of that rate is orthogonal to obtainable information. As with our approach, an economic model must be used to link the market's expectation to observable variables (i.e., market prices). Given this link, the forecast error can be measured and regressed on past information available to traders. The primary difference between the exchange rate tests and our own is that we exploit option price data to focus on the variance, not the mean, of the underlying asset.

### 1.2 Characterization of stochastic volatility

Tests of the orthogonality restrictions will lack power unless information available to the market can be used to predict return volatility. We show in this subsection that the returns in our sample are accurately characterized by the GARCH process, a parametric model of persistence in conditional variance. As such, the model links return volatility to the past behavior of the return process itself, which is included in  $I_t$ . Under the null hypothesis, GARCH momentum will be fully used by options traders in forming expectations of variance; thus, the GARCH model will have no marginal predictive power over the implied variance.

Consider the following GARCH (1, 1) model for stock returns:

$$r_t = \bar{r} + \epsilon_t, \quad (4)$$

$$\epsilon_t \mid \epsilon_{t-1}, \epsilon_{t-2}, \dots \sim N(0, h_t), \quad (5)$$

$$h_t = C + \alpha \epsilon_{t-1}^2 + \beta h_{t-1} + \gamma \zeta_{t-1}, \quad (6)$$

where  $r_t$  is the return over day  $t$ ;  $\bar{r}$ ,  $C$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$  are parameters, and  $\zeta_{t-1}$  is a vector of exogenous variables. In this section, we constrain  $\gamma$  to be zero, so (6) is the conventional GARCH specification. It is easily verified that if  $\alpha + \beta = 1$  there is no mean reversion in the variance. In this case, the conditional variance is integrated, and unconditional variance is undefined. Persistence of shocks to variance increases rapidly as this sum approaches unity from below. The return process is stationary if  $\alpha + \beta < 1$ .

The GARCH model is a discrete-time approximation to the diffusion process in Equations (1) and (2) and is therefore consistent with the

HW model. We can think of the actual stock price and its variance as being generated in continuous time by (1) and (2), but with the data observed discretely (daily). It is evident that (4) is the discrete-time analog to (1), where  $\bar{r}$  approximates the drift  $\phi$ . To see how (6) maps into (2) (assuming  $\gamma = 0$ ), subtract  $b_{t-1}$  from both sides of (6) to obtain

$$b_t - b_{t-1} = [C/b_{t-1} - (1 - \alpha - \beta)]b_{t-1} + \alpha b_{t-1}(\theta_{t-1}^2 - 1), \quad (6)$$

where we have used the identity  $\epsilon_t^2 = b_t \theta_t^2$ , and  $\theta$  is an i.i.d. standardized normal random variable. Nelson (1990) has shown that as the time interval goes to zero this expression approaches the diffusion in (2), where  $C/b_{t-1} - (1 - \alpha - \beta)$  approaches  $\mu dt$  and  $\alpha$  approaches  $\xi\sqrt{dt}/2$ . This approximation is not unique, but it is consistent with our use of the GARCH model in the subsequent analysis.

It is not intuitive that the continuous-time limit of the GARCH process has two sources of randomness [ $dw$  and  $dz$  in (2)], whereas the discrete-time process appears to have a single source of randomness [ $\epsilon$  is the only stochastic term in (4) to (6)]. The intuition for this is as follows. By definition, the conditional variance of *next period's* residual is not stochastic in the GARCH framework. However, the (conditional) forecast of the variance over the next two periods depends on the realization of  $\epsilon_{t+1}$  and hence is random. Now, as the interval between time periods shrinks to its limit, the ability to distinguish between time  $t + 1$  and time  $t$  is lost, which yields the second source of randomness in the limiting case.

The GARCH model is estimated for a sample of daily returns for 10 individual stocks over the period April 19, 1982, to March 30, 1984 (496 trading days), except for company 10, the sample for which begins June 30, 1982. The sample is chosen to conform to our tests of the orthogonality restrictions, as discussed in the following section. The return data come from the CRSP tapes; daily returns are thus calculated as the rate of change of the last transaction price of the day.

Estimates from the GARCH model are reported in Table 1, for each of the 10 companies, as model 2. Ticker symbols are provided in Table 1, and we use these in reference to particular companies. Maximum likelihood estimation of the GARCH model is carried out by using a variant of the Berndt, Hall, Hall, and Hausman optimization algorithm that constrains parameter estimates in the variance equations to be nonnegative [Biegler and Cuthrell (1985)]. First derivatives are calculated numerically.<sup>2</sup>

<sup>2</sup> We tested all 10 companies for the presence of an AR(1) process in returns, allowing for GARCH residuals. All 10 had positive first-order serial correlation, but this was significant at the 10 percent level for only one company (number 10). The magnitude of this serial correlation is trivial: the largest first-order serial correlation coefficient (company 10) is .13; the median is .06.



The evidence in Table 1 shows that the GARCH model provides a good fit for the 10 stocks in the sample. In all cases but one, the GARCH parameters are statistically different from zero at small significance levels, according to the asymptotic  $t$ -statistics. For DEC (panel B),  $\alpha$  is significant at a 10 percent level. For most of the stocks, the sum  $\alpha + \beta$  exceeds .9.

### 1.3 Characterization of bias

The interpretation of our tests relies on equating the variance implied from the data and the Black–Scholes model with the market's subjective variance of returns over the remaining life of the option. However, there are potentially important ways in which the implied variance can be expected to differ from the subjective variance, even assuming the HW model is valid and markets are efficient. In this subsection, we attempt to quantify this bias in order to determine if it affects our inferences.

The implied variance may deviate from the subjective variance as a result of measurement error in the option price and nonsynchronous stock and option prices. However, because of our careful choice and handling of the data, as discussed in Section 2.1, the bias from this source is likely to be trivial and will not influence our results.

Two other sources of bias, however, are potentially more serious. First, as noted, the virtual linearity of the Black–Scholes formula for at-the-money options means that the implied variance will only be an approximation of the true subjective variance. Without further analysis, however, it is not clear how good the approximation will be. Hull and White (1987), for example, indicate that large values for the standard deviation of the variance process can lead to large biases in implied variances even for at-the-money options. Second, the stock return distribution may be skewed, implying a nonzero  $\rho$  and additional bias. In order to understand the importance, *in our data*, of the linearity approximation and the assumption that  $\rho = 0$ , we perform a Monte Carlo simulation of the continuous-time process for returns in (1) and (2). The simulation allows  $\rho$  to be nonzero. The magnitude of the bias from these two sources is measured by comparing the implied variance from the simulated data with the actual variance inherent in this data.

We calibrate the simulation to be consistent with the market data used in this study. Because the GARCH model approximates the diffusion process, the parameters in (1) and (2) are constructed from the GARCH estimates in Table 1. For example,  $\xi = \sqrt{2\alpha}$ ,  $dt$  is taken to be 1, and  $\alpha$  is estimated directly from the GARCH model. Sample values for  $\rho$  are obtained by estimating the sample correlation between  $r_t$  and  $b_t$ , the fitted value from the GARCH equation (6).

**Table 1**  
**Specifications of conditional variance**

Model	$\mathcal{L}$	$\bar{r}$ ( <i>t</i> -stat.)	$C$ ( <i>t</i> -stat.)	$\alpha$ ( <i>t</i> -stat.)	$\beta$ ( <i>t</i> -stat.)	$\gamma$ ( <i>t</i> -stat.)
A: Computer Sciences Corp. (CSC) (20536310)						
1	-2281.8	0.648 (0.56)	590.52 (18.98)			
2	-2270.3	0.385 (0.33)	238.15 (1.99)	0.110 (2.04)	0.484 (2.10)	
3	-2275.0	0.000 (0.00)	262.62 (3.53)			0.576 (3.79)
4	-2264.9	-0.034 (-0.03)	58.85 (1.02)	0.058 (1.76)	0.612 (2.82)	0.237 (1.60)
B: Digital Equipment Corp. (DEC) (25384910)						
1	-2292.0	0.484 (0.42)	614.71 (38.71)			
2	-2280.2	0.011 (0.01)	413.79 (2.44)	0.094 (1.77)	0.228 (0.74)	
3	-2292.1	0.489 (0.42)	613.55 (9.80)			0.000 (0.00)
4	-2280.2	0.008 (0.01)	404.80 (2.45)	0.095 (1.76)	0.230 (0.70)	0.157 (0.85)
C: Datapoint (DPT) (23810020)						
1	-2511.9	0.910 (0.52)	492.31 (9.68)			
2	-2482.0	1.338 (0.88)	11.56 (2.06)	0.020 (5.00)	0.969 (138.43)	
3	-2485.6	0.583 (0.59)	0.00 (0.00)			1.629 (9.70)
4	-2475.2	0.562 (0.34)	31.37 (0.39)	0.081 (2.89)	0.373 (1.34)	0.843 (1.86)
D: Federal Express (FDX) (31330910)						
1	-2259.2	0.678 (0.65)	538.12 (20.00)			
2	-2245.3	0.474 (0.46)	40.18 (1.84)	0.056 (2.67)	0.869 (15.52)	
3	-2251.9	0.301 (0.29)	0.0 (0.00)			1.348 (3.62)
4	-2243.7	0.298 (0.28)	0.00 (0.00)	0.057 (2.36)	0.791 (8.15)	0.204 (1.46)
E: National Semiconductor (NSM) (63764010)						
1	-2464.3	1.855 (1.14)	1230.98 (18.40)			
2	-2450.7	1.972 (1.26)	55.20 (1.73)	0.052 (2.60)	0.904 (24.43)	
3	-2451.8	1.489 (0.95)	0.00 (0.00)			1.768 (3.94)
4	-2445.8	1.274 (0.83)	0.00 (0.00)	0.050 (1.32)	0.678 (2.63)	0.480 (1.01)
F: Paradyne (PDN) (69911310)						
1	-2492.8	-0.359 (-0.21)	1443.24 (29.08)			
2	-2476.5	1.599 (0.97)	188.51 (3.08)	0.089 (3.56)	0.783 (12.63)	
3	-2482.3	0.616 (0.35)	321.59 (2.60)			1.402 (7.01)
4	-2477.4	0.625 (0.34)	230.78 (1.96)	0.000 (0.00)	0.194 (0.85)	1.166 (3.28)

**Table 1**  
Continued

Model	$\mathcal{L}$	$\bar{r}$ ( <i>t</i> -stat.)	$C$ ( <i>t</i> -stat.)	$\alpha$ ( <i>t</i> -stat.)	$\beta$ ( <i>t</i> -stat.)	$\gamma$ ( <i>t</i> -stat.)
G: Rockwell (ROK) (77434710)						
1	-2181.0	1.558 (1.73)	392.10 (20.84)			
2	-2171.7	1.590 (1.73)	31.86 (1.33)	0.048 (2.00)	0.871 (11.02)	
3	-2175.5	1.558 (1.77)	161.21 (2.71)			1.011 (3.74)
4	-2167.9	1.427 (1.59)	25.98 (0.94)	0.043 (0.12)	0.742 (3.80)	0.253 (1.28)
H: Storage Technologies (STK) (86211110)						
1	-2443.6	-1.328 (-0.85)	1134.58 (23.09)			
2	-2423.7	-1.548 (-1.09)	96.26 (3.06)	0.079 (3.76)	0.836 (19.90)	
3	-2428.2	-2.015 (-1.34)	0.000 (0.00)			1.505 (5.57)
4	-2421.1	-1.762 (-1.19)	0.00 (0.00)	0.093 (2.16)	0.573 (3.54)	0.513 (2.01)
I: Tandy Corp. (TAN) (87538210)						
1	-2337.2	0.362 (0.30)	738.35 (19.11)			
2	-2311.3	-0.548 (-0.49)	49.97 (2.38)	0.112 (3.50)	0.819 (16.06)	
3	-2325.4	-0.248 (-0.21)	0.00 (0.00)			1.350 (4.35)
4	-2307.9	-0.66 (-0.59)	0.00 (0.00)	0.128 (2.72)	0.656 (4.72)	0.287 (1.39)
J: Toys R US (TOY) (89233510)						
1	-1980.2	2.082 (1.63)	680.39 (20.64)			
2	-1962.2	2.086 (1.80)	8.14 (1.48)	0.040 (3.64)	0.946 (59.12)	
3	-1968.6	1.674 (1.35)	0.00 (0.00)			1.311 (5.06)
4	-1961.9	-1.615 (1.28)	0.00 (0.00)	0.102 (1.57)	0.381 (1.22)	0.682 (1.80)

Model:  $r_t = \bar{r} + \epsilon_t$ 1.  $\epsilon_t \sim N(0, C)$ 2.  $\epsilon_t \sim N(0, b_t)$  $b_t = C + \alpha\epsilon_{t-1}^2 + \beta b_{t-1}$ 3.  $\epsilon_t \sim N(0, b_t)$  $b_t = C + \gamma\zeta_{t-1}$ 4.  $\epsilon_t \sim N(0, b_t)$  $b_t = C + \alpha\epsilon_{t-1}^2 + \beta b_{t-1} + \gamma\zeta_{t-1}$ 

$\zeta_t$  represents the daily implied variance from minimizing the sum-of-squared errors from all option midpoint quotes on day  $t$ , for the nearest to-the-money option, intermediate term to expiration. All returns are daily percentages times 1000. *t*-stat. represents the asymptotic Student's *t* statistic. This may be biased as a result of the departure from normality of  $\epsilon_t b_t^{-0.5}$ .  $\mathcal{L}$  represents the value of the log-likelihood function at its optimum for each model. All models are estimated using daily data from April 19, 1982 through March 30, 1984; except TOY, which starts June 30, 1982 (423 days).

Given these parameter values, which are grounded in the data, the simulation proceeds as follows for each company in the sample. Risk-neutral pricing and (1) and (2) are assumed to be true. Also, the sample value of  $\rho$  is assumed to be true, as are the (time-dependent) values of  $\mu$  and (time-independent) values of  $\xi$  obtained from the GARCH estimates. Finally,  $\phi$  is assumed to be the average daily return on the riskless asset and takes the value 0.000245. Given these parameter values, the stock-return process is simulated 1000 times over the life of the option according to (1) and (2), where we approximate the continuous-time process by dividing the 135-day remaining life of the option into discrete, daily increments.<sup>3</sup> This experiment yields an estimate of the actual price for the (European) call option. Under risk-neutral valuation, this estimate is computed as the mean of the discounted terminal value of the option over the 1000 trials. The method described in Section 2 is then used to impute a variance from the simulated price and the initial value of the stock price, where stock and strike prices are chosen to be at-the-money. The simulations also define the actual mean cumulative variance over the option's remaining life,  $E(\bar{V}_t | I_t)$ . Under the null hypothesis, this conditional mean equals the subjective variance. Bias is then defined as the difference between the simulated actual mean cumulative variance and the implied variance from the model.

If there were no computational limits to the simulation, there would be no reason to generate only 1000 replications of the process. In the limit as the number of realizations approaches infinity, the empirical distribution of  $\bar{V}$  approaches its true distribution. However, given practical constraints, it is impossible to determine how many finite simulations are required to obtain a good estimate of the true distribution. Our strategy is to use 1000 simulations to compute the option price and  $E(\bar{V}_t | I_t)$ , and to repeat this procedure 100 times. With 100 realizations of the option price (and thus the implied variance) we are able to construct a confidence interval around mean bias, which quantifies the numerical bias in the experiment.<sup>4</sup>

As a check on the possible numerical bias from using only 1000 draws, the foregoing procedure is modified so that 10,000 draws of the stock and variance evolution are made (the 10,000-draw simulation is conducted independently from the 1000-draw simulation).

<sup>3</sup> The antithetic variate technique of using each pair of draws four times, discussed in Hull and White (1987), is used in this regard. Equations (11) of that article (see notes to our Table 2), with normally distributed random variates, form the basis for the simulation.

<sup>4</sup> In effect, the simulation experiment utilizes 100,000 simulations (1000 realizations, 100 times) to construct estimates of bias. The comparable alternative strategy—100,000 realizations with no repetitions—yields only one estimate of option price, so no confidence interval can be constructed. Though 100,000 simulations may be sufficient to ensure a precise estimate of the true option price and the distribution of  $\bar{V}$ , there is no way to determine this from the alternative strategy.

Here 1,000,000 draws are taken for each stock. The effect of the size of the simulation is seen by comparing the results as we go from 1000 to 10,000 draws.

The results from the Monte Carlo experiments are contained in Table 2. Sample means and standard errors are reported over the 100 repetitions of the simulated mean cumulative variance  $E(\bar{V}_t | I_t)$ , the implied variance, and inherent bias. These values are computed and reported for three values of initial variance  $[V(0)]$ . The middle value of  $V(0)$  is the unconditional variance,  $C$ , taken from Table 1, specification 1. Simulations using 10,000 draws are only conducted for the middle value of  $V(0)$  and are reported last for each company in the table.

The results show that mean bias inherent in the analytic approximation using the option pricing model is never more than 1.3 percent of the actual variance. The variance nested in the Monte Carlo analysis appears to be trivial because standard errors of the bias are also small: two-standard-error confidence intervals around the mean bias never include 3 percent in absolute value. When 10,000 trials are repeated, the two-standard-error bounds are always less than 1.5 percent in absolute value. Note also that the mean implied volatility from the 10,000-draw simulations is always within two numerical standard errors of the 1000-draw simulation.

From the simulation results obtained by Hull and White (1987), we might expect large bias for cases in which  $\alpha$  is large. For example, they show that for  $\xi = 3$  on an annual basis, which approximately corresponds to  $\alpha = .11$  (as for CSC and TAN in our data), the bias of the implied variance is 20 percent. However, their result holds for  $\mu = 0$ ; that is, the variance process follows a random walk. Our findings of low bias are likely due to mean reversion in the estimated discrete-time variance process. Also, for companies like DPT and TOY that have strong persistence in variance, estimated  $\alpha$  values are relatively small.

As a sensitivity check, the experiments for several companies were repeated with much higher values of  $\rho$  in absolute value. For instance, FDX with  $V(0)$  of 538.12 and  $\rho$  of  $-.8$  generates a mean bias of 0.59 percent (two-standard-error confidence interval:  $[-.69, 1.86]$ ); when  $\rho$  is set to  $.8$ , the mean bias becomes  $-1.28$  percent (two-standard-error confidence interval:  $[-2.78, 0.22]$ ). This result is representative of those for all 10 stocks. We conclude that under the HW model with risk-neutral probabilities, our variance extraction procedure appears to be insensitive to the nonlinearity assumption and skewness in the context of our data. We will refer to these results to ascertain the potential effects of the inherent bias on inferences from the forecast-based tests of the orthogonality restrictions.

**Table 2**  
**Bias inherent in analytic approximation**

Draws	V(0)	Simulated variance (std. err.)	Implied variance (std. err.)	Bias (%) (2 std. err. range)
A: CSC, $\rho = 0.2962$				
1000	438.56	589.49 (0.01)	593.86 (4.12)	-0.74 (-2.14, 0.66)
1000	590.52	591.02 (0.01)	596.54 (4.14)	-0.93 (-2.33, 0.47)
1000	750.52	592.84 (0.01)	599.52 (4.17)	-1.13 (-2.53, 0.28)
10,000	590.52	591.04 (0.004)	592.50 (1.26)	-0.25 (-0.67, 0.18)
B: DEC, $\rho = -0.1872$				
1000	354.71	614.06 (0.01)	617.19 (4.29)	-0.51 (0.89, -1.91)
1000	614.71	614.42 (0.01)	619.43 (4.14)	-0.81 (-2.15, 0.53)
1000	834.71	615.36 (0.01)	621.98 (4.17)	-1.07 (-2.43, 0.27)
10,000	614.71	614.39 (0.003)	614.98 (1.30)	-0.10 (-0.52, 0.33)
C: DPT, $\rho = -0.1059$				
1000	332.31	681.07 (0.03)	679.87 (4.81)	0.18 (-1.24, 1.59)
1000	492.31	763.05 (0.03)	762.97 (5.49)	0.01 (-1.43, 1.45)
1000	642.31	840.19 (0.04)	841.14 (6.13)	-0.11 (-1.57, 1.35)
10,000	492.31	763.09 (0.01)	757.56 (1.66)	0.72 (0.29, 1.16)
D: FDX, $\rho = 0.0493$				
1000	378.12	522.16 (0.02)	524.68 (3.59)	-0.48 (-1.86, 0.89)
1000	538.12	536.66 (0.03)	540.46 (3.72)	-0.71 (-2.10, 0.68)
1000	798.12	560.59 (0.03)	566.48 (3.91)	-1.05 (-2.44, 0.34)
10,000	538.12	536.67 (0.01)	536.79 (1.12)	-0.02 (-0.44, 0.39)
E: NSM, $\rho = 0.2085$				
1000	880.98	1196.00 (0.10)	1205.98 (9.58)	-0.83 (-2.44, 0.76)
1000	1230.98	1251.91 (0.11)	1265.30 (10.12)	-1.07 (-2.69, 0.55)
1000	1580.98	1308.22 (0.11)	1325.03 (10.67)	-1.28 (-2.92, 0.35)
10,000	1230.98	1252.03 (0.03)	1254.77 (3.10)	-0.20 (-0.70, 0.28)
F: PDN, $\rho = -0.0408$				
1000	1043.24	1457.01 (0.09)	1465.23 (11.58)	-0.56 (-2.15, 1.02)
1000	1443.24	1476.86 (0.09)	1488.33 (10.20)	-0.78 (-2.16, 0.60)
1000	1843.24	1497.19 (0.09)	1511.89 (12.05)	-0.98 (-2.59, 0.63)

**Table 2**  
Continued

Draws	V(0)	Simulated variance (std. err.)	Implied variance (std. err.)	Bias (%) (2 std. err. range)
10,000	1443.24	1476.97 (0.03)	1475.35 (3.64)	0.11 (-0.40, 0.60)
G: ROK, $\rho = 0.0360$				
1000	262.10	382.81 (0.01)	384.55 (2.54)	-0.45 (-1.78, 0.87)
1000	392.10	393.62 (0.01)	396.39 (2.63)	-0.70 (-2.04, 0.63)
1000	522.10	404.61 (0.01)	408.40 (2.72)	-0.94 (-2.28, 0.41)
10,000	392.10	393.60 (0.005)	393.83 (0.79)	-0.06 (-0.46, 0.34)
H: STK, $\rho = -0.2735$				
1000	884.58	1115.95 (0.09)	1114.47 (8.19)	0.13 (-1.34, 1.60)
1000	1134.58	1135.74 (0.09)	1136.07 (8.38)	-0.03 (-1.50, 1.45)
1000	1484.58	1163.81 (0.09)	1166.63 (8.64)	-0.24 (-1.73, 1.24)
10,000	1134.58	1135.84 (0.03)	1127.21 (2.56)	0.76 (0.31, 1.21)
I: TAN, $\rho = -0.0563$				
1000	518.35	707.59 (0.16)	704.63 (4.82)	0.42 (-0.94, 1.78)
1000	738.35	729.40 (0.17)	727.98 (5.01)	0.19 (-1.18, 1.57)
1000	1008.35	756.52 (0.18)	756.91 (5.23)	-0.05 (-1.43, 1.33)
10,000	738.35	729.59 (0.06)	722.75 (1.53)	0.94 (0.005, 1.36)
J: TOY, $\rho = 0.1518$				
1000	530.39	558.84 (0.09)	559.12 (3.99)	-0.05 (-1.48, 1.38)
1000	680.39	625.41 (0.10)	626.52 (4.53)	-0.18 (-1.63, 1.27)
1000	880.39	714.29 (0.13)	716.49 (5.26)	-0.31 (-1.78, 1.16)
10,000	680.39	625.34 (0.03)	622.22 (1.38)	0.50 (0.06, 0.94)

$$\text{Model: } dS = \phi S dt + \sqrt{V} S dw, \tag{1}$$

$$dV = \mu V dt + \xi V dz. \tag{2}$$

S is the stock price. The Brownian motions dw and dz have an instantaneous correlation of  $\rho$ . Following Hull and White (1987), risk-neutral option valuation is performed by simulating the following two equations:

$$S_t = S_{t-1} \exp[(r_f - V_{t-1}/2)\Delta t + u_t \sqrt{V_{t-1}} \Delta \tilde{t}], \tag{1'}$$

$$V_t = V_{t-1} \exp[(\mu - \xi^2/2)\Delta t + \rho u_t \xi \sqrt{\Delta \tilde{t}} + \sqrt{1 - \rho^2} v_t \xi \sqrt{\Delta \tilde{t}}]. \tag{2'}$$

The relevant time increment ( $\Delta t$ ) is taken to be one day. In all cases the option expires in 135 days. (1') and (2') are simulated 1000 times (draws), using the four-way control variate technique described in Hull and White (1987). [The simulation is repeated using 10,000 simulations, 100 times for the middle value of V(0). The results from this experiment are reported for each company—using the middle value of V(0)—below the first horizontal line.] This provides a single

## 2. Tests of the Orthogonality Restriction

### 2.1 Tests and data

We test the orthogonality restrictions in two basic ways. The first set of tests, in Section 2.2, analyzes the marginal predictive power of the past behavior of the return process, given the implied variance, using in-sample regressions and classical statistical tests. Section 2.3 contains results from out-of-sample tests of forecastability. We investigate the ability of implied variance to predict actual volatility out-of-sample to models using past information by comparing root mean square errors and by estimating encompassing regressions.

The implied variances used in this study are constructed from option and contemporaneous stock-price data for 10 individual stocks with publicly traded options on the Chicago Board Options Exchange (CBOE) for the period April 19, 1982, through March 31, 1984. On the floor of the CBOE there are multiple competing market makers for each of these options. A clerk records every time one of the market makers quotes a bid price that is higher than the bids of the other market makers. Similarly, the clerk records all ask quotes that are lower than extant asks. Both the bid and ask quotes at such points are posted—time-stamped to the nearest second—along with the most recent stock price to have crossed the ticker. We refer to the highest bid and lowest ask quotes as the inside spread. The best bid and best ask quotes are not necessarily from the same market maker. Our data, taken from the Berkeley options database, consist of every inside spread during each day in our sample. Each option inside spread is paired with the most recent stock price from the ticker, recorded by the clerk.

The sample of data and extraction techniques used in this study are chosen to avoid as much as possible errors in the measurement of the implied variance. The stocks in our sample paid no cash dividends from 1981 through 1985; therefore, the options are essentially European. The sample period has no special significance. It starts nine years following the inception of public trading in listed options, so market makers should be adroit at their job. It also allows isolation of non-dividend-paying stocks and is a relatively calm period (com-

---

(market) option price (and market cumulative variance forecast) from which an implied volatility is computed analytically, following Black-Scholes. This will be exactly correct in those cases where Black-Scholes is exactly linear in  $V$  and where  $\rho = 0$ . This simulation to obtain an implied volatility is repeated 100 times to provide the mean and standard errors of the inherent bias (the source of variation is numeric, not sampling error).  $u$  and  $v$  are independent normal  $(0, 1)$  random deviates.  $S(0) = Xe^{-rt}$ ,  $r_f = 0.000245$ . The values of  $\mu$  and  $\xi$  are derived from the GARCH (1, 1) estimates of the variance equation for each company (see Table 1), following Nelson (1990). The value of  $\rho$  is computed as the sample correlation between the GARCH (1, 1) conditional variance on day  $t$  (given information up to day  $t-1$ ) and the return on day  $t$ . All variances are daily times 1 million.



pared to, say, late 1987), so the ticker should report up-to-the-minute stock transaction prices.

Note that we are not using transaction prices from the options market. At any point in time when the market is open, the option price is assumed to be the midpoint of the inside bid-ask spread. Although actual transaction prices may include price pressure effects, and (latest) transaction prices from both markets at a fixed point in time (such as closing) will always be asynchronous, our data suffer from neither of these problems.

However, the daily return series may embody noise due to the bid-ask spread “bounce.” The subset of stocks with listed options in 1982 consisted of actively traded and generally large stocks in terms of market value of equity. For these stocks, relative spreads tend to be small, and the observed price process should be an excellent instrumental variable for the latent “true” price. By the same token, the most recent stock price matched to the option from the Berkeley tapes may contain some noise. Again, to the extent that this noise is well-behaved, the procedure used to imply a single variance from the entire day’s data should serve to trivialize the errors-in-variables problem.

Only those options that were closest to being at-the-money are used. Furthermore, of the three expiration dates available for most companies and most days, the intermediate-term option is used throughout the analysis. The number of quotes in a day varies significantly across the sample, but the average number of quotes per day per company is about 50 for the at-the-money, intermediate-term call option.

We construct time series of implied variances for each stock in the sample as follows. On each day, at-the-money options are isolated by choosing those options with the closest discounted (at the risk-free rate) strike price to stock price at option market close. Of these, for those options that expire at the intermediate term, a single daily implied variance is computed by minimizing over variance the sum of squared errors from actual midpoint quotes to the Black-Scholes model value at that variance.<sup>5</sup> The model value is taken to be a function of the yield on the U.S. Treasury bill maturing as closely as possible to the intermediate-term expiration date and the most recent transaction price of the stock from the NYSE, which is the simultaneous stock price reported on the Berkeley tape.<sup>6</sup> As noted, there are

<sup>5</sup> A quadratic hill-climbing algorithm is used that has good convergence properties. This technique was suggested by Whaley (1982) and is more efficient than that used by Brenner and Galai (1987), namely, implying a variance from each option and using the daily average.

<sup>6</sup> We collected daily Treasury-bill yields from the *Wall Street Journal*. For options that mature in six months or less, there is always a one-day difference in the maturity date of the Treasury bill and

on average 50 quote pairs used to define a single implied variance per day. Options quote midpoints and the corresponding stock prices that violate the Black–Scholes lower boundary condition are discarded. Otherwise, all qualifying options within the day are treated equally. This procedure assumes that the variance is constant within a day.

## 2.2 Regression-based tests

In this subsection we test the orthogonality restrictions by examining the significance of the GARCH coefficients ( $\alpha$  and  $\beta$ ) in the conditional variance specification after accounting for option market forecasts of volatility. Hence, we define  $\zeta_{t-1}$  in (6) to be the implied variance given information at time  $t - 1$ , and we allow  $\gamma$  to be a free parameter. From the discussion in Section 1.2, the conditional variance  $h_t$  corresponds to the instantaneous variance of the diffusion,  $V_t$ . If the life of the option  $T$  is one day, then  $\bar{V}_t$  and  $V_t$  are identical for the discrete-time approximation. The general conditional variance equation in (6) then can be interpreted as a regression of  $\bar{V}_t$  on the subjective variance, as measured by the implied variance, and past information. The orthogonality restriction of the joint hypothesis implies that the GARCH coefficients in (6) are zero (i.e., that the GARCH variables have no marginal predictive power).

This test is subject to an obvious and important criticism. Whereas  $h_t$  is the conditional variance of daily returns, in this study  $\zeta_{t-1}$  is the implied variance from an option that matures at a horizon greater than a day (between 64 and 129 trading days). Under the HW model, the implied variance represents the market's prediction of average daily volatility over the remaining life of the option. Thus, given this specification,  $\zeta_t$  is not an exact predictor of the dependent variable  $h_t$ . Day and Lewis (1992) use weekly return horizons with index options that mature in a month and are therefore subject to the same criticism. We conduct these tests as a preliminary characterization of the orthogonality restrictions and account for this maturity mismatch problem with out-of-sample tests in Section 2.2 below.

The results of the in-sample, regression-based tests are shown in Table 1. The table contains the estimation results for the unrestricted model in (4) to (6) (specification 4) and three restricted specifications. Specification 1 is a homoskedastic model where the restriction

---

the option expiration date. Since 12-month bills are only auctioned every four weeks, there is sometimes an eight-day difference between the maturity of the applicable Treasury bill and the expiration of the option. There are three days in the sample period where the Treasury-bill market was closed, but the NYSE and CBOE were open. In these cases, the previous day's rate data were substituted. The Berkeley database is missing data for the dates July 1, 1983, and December 23, 1983. For these dates, we use the options data from the previous day to replace the missing values.

$\alpha = \beta = \gamma = 0$  is imposed. Specification 2 is standard GARCH(1, 1) where  $\gamma$  is restricted to equal 0 (and is discussed in Section 1.2). Specification 3 restricts the conditional heteroskedasticity to be entirely manifest in the option market's variance forecast. A likelihood ratio test (LRT) on the restrictions imposed on specification 2 by specification 1 suggests that the null hypothesis of no GARCH ( $\alpha = \beta = 0$ ) can be rejected at the 1 percent level for all 10 companies, assuming conditional normality, which confirms the inferences in Section 1.2. As discussed there, nontrivial variance clustering is an important characteristic of most of the companies.

From specification 3, we can reject the null that  $\zeta$  has no explanatory power for actual daily variance for all companies except DEC, even given the maturity mismatch problem. In all other cases, this coefficient is significant and positive. The coefficient on  $\zeta$  exceeds unity in eight cases.

The joint null hypothesis of informational efficiency and the HW model can be tested against the alternative that allows GARCH terms to have incremental predictive ability by comparing specifications 3 and 4. Using the LRT, we reject the null hypothesis at standard significance levels for 7 of the 10 companies: CSC, FDX, NSM, ROK, STK, TAN, and TOY.<sup>7</sup> Statistical inference for the remaining three companies is hindered by the fact that the nonnegativity constraints imposed in estimation are binding in the variance equation.<sup>8</sup>

Despite different sample periods, assets, horizons, and so forth, these results are consistent with those of Day and Lewis (1992)—past information improves the market forecast of volatility. However, this result must be interpreted with caution since the maturity mismatch problem may bias the test against the implied variance. In the next subsection we make adjustments for this problem in the context of out-of-sample forecast comparisons.

### 2.3 Out-of-sample tests

The incompatibility of forecast horizons that arises in the regression-based tests is eliminated in this subsection by transforming GARCH forecasts of daily variance to forecasts of average daily variance over the remaining life of the corresponding option. On any day  $t$  in the sample ( $t = 1, \dots, 495$ ), we can construct a forecast of  $h_t$  by using the fitted value of Equation (6) with  $\gamma$  equal to zero. By recursive substitution of this GARCH equation, the forecast for  $h_{t+k}$  can be con-

<sup>7</sup> The LRT is more appropriate than Wald tests given the potential collinearity between  $\epsilon_{t-1}^2$ ,  $h_{t-1}$ , and  $\zeta_{t-1}$ .

<sup>8</sup> These constraints bind in some other cases (e.g., FDX), but here the constraint is binding in both specifications 3 and 4.

structured for any  $k > 0$ , given information at  $t$ . To obtain a GARCH forecast that is directly comparable to our interpretation of the implied variances, we construct the GARCH forecasts for  $h_{t+1}, h_{t+2}, \dots, h_{t+N}$ , where  $N$  is the number of days left in the life of the intermediate-term option on day  $t$ . Denote the mean over these  $N$  forecasts by  $G_t$ . The forecast horizons for  $G_t$  and the implied variance are identical by construction. The joint null hypotheses imply that  $G_t$  is not a better predictor of  $N$ -step ahead realized return volatility than the implied variance; that is, the orthogonality restriction means that  $G_t$  cannot be used to improve the forecast in the implied variance. In this subsection, we compare the forecast performance of implied variance with the GARCH forecast  $G$  and a naive forecasting model.

To ensure that forecast comparisons are not biased in favor of the time-series model, we are careful to construct the GARCH forecasts by using information available to traders at the time the forecasts are made. Thus, forecast comparisons are made out-of-sample. We estimate both rolling and updating GARCH models. The rolling structure uses a constant sample size of 300 observations, adding the return on day  $t - 1$  and deleting the return on day  $t - 301$  from the sample used to estimate GARCH on each day  $t$ . The updating procedure simply adds information as time progresses to construct an updated forecast. Because the GARCH model is estimated only from stock-return data (and is therefore not tied down to the options sample), the first sample begins 301 trading days before April 19, 1982, the first day of our implied variance sample. The GARCH model is reestimated 495 times, for each procedure, to construct out-of-sample forecasts that are up-to-date. In addition to GARCH, we also consider the updated sample variance of past returns as a naive forecast of variance:

$$H_t = \frac{1}{t} \sum_{i=1}^t \hat{\epsilon}_i^2,$$

where  $\hat{\epsilon}_i^2$  is the estimated residual from Equation (4), with  $\alpha = \beta = \gamma = 0$ .

Forecasting performance is judged by comparing the ability of the forecasts to predict the out-of-sample mean of the squared return residuals from Equation (4) over the remaining life of the intermediate-term option. To be precise, assume that this option on day  $t$  has  $N$  days to maturity. Then the realized volatility over this period is given by

$$z_t = \frac{1}{N} \sum_{i=1}^N \hat{\epsilon}_{t+i}^2.$$

Note that  $z_t$  is constructed to be compatible with the interpretation

of the implied variance and  $G_t$ . Comparisons are based upon the out-of-sample mean error

$$\text{ME} = \frac{1}{495} \sum_{t=1}^{495} (z_t - x_t),$$

where  $x_t$  is alternatively the implied variance at time  $t$ , the two GARCH forecasts, and  $H_t$ , mean absolute error

$$\text{MAE} = \frac{1}{495} \sum_{t=1}^{495} |z_t - x_t|,$$

and root mean square error

$$\text{RMSE} = \left[ \frac{1}{495} \sum_{t=1}^{495} (z_t - x_t)^2 \right]^{1/2}.$$

The results of this exercise are contained in Table 3. The implied variance has the smallest RMSE for only two companies (DPT and TOY). For the remaining companies, the GARCH and the naive forecasts have lower RMSE than the implied variance. Thus, for these companies, using past information on the stock-return process can improve the market's forecast. The following points are also evident from the table:

1. The updating GARCH outperforms rolling GARCH for all 10 companies under RMSE criterion.<sup>9</sup>
2. The updated sample variance has the lowest RMSE in 5 of 10 cases. This relative forecasting performance is at odds with the results of Akgiray (1989). Akgiray finds that GARCH variance forecasts are convincingly superior to historical variance as a forecast using the RMSE criterion for stock index data. However, Akgiray uses a forecast horizon of only 20 days. We replicated his analysis with a 100-day horizon (representative of the average number of calendar days in the horizon in this study) and found that the relative rankings of historical variance and GARCH were overturned, which is consistent with our results.
3. The ME for implied variance is positive for all companies, which indicates that the implied variance is systematically lower than the actual volatility in this period. This characteristic of the forecasting performance of implied volatilities is readily visible in Figures 1 and 2. Here we plot the realized mean squared returns over the option's life, the updated sample variance, and the implied variance for two

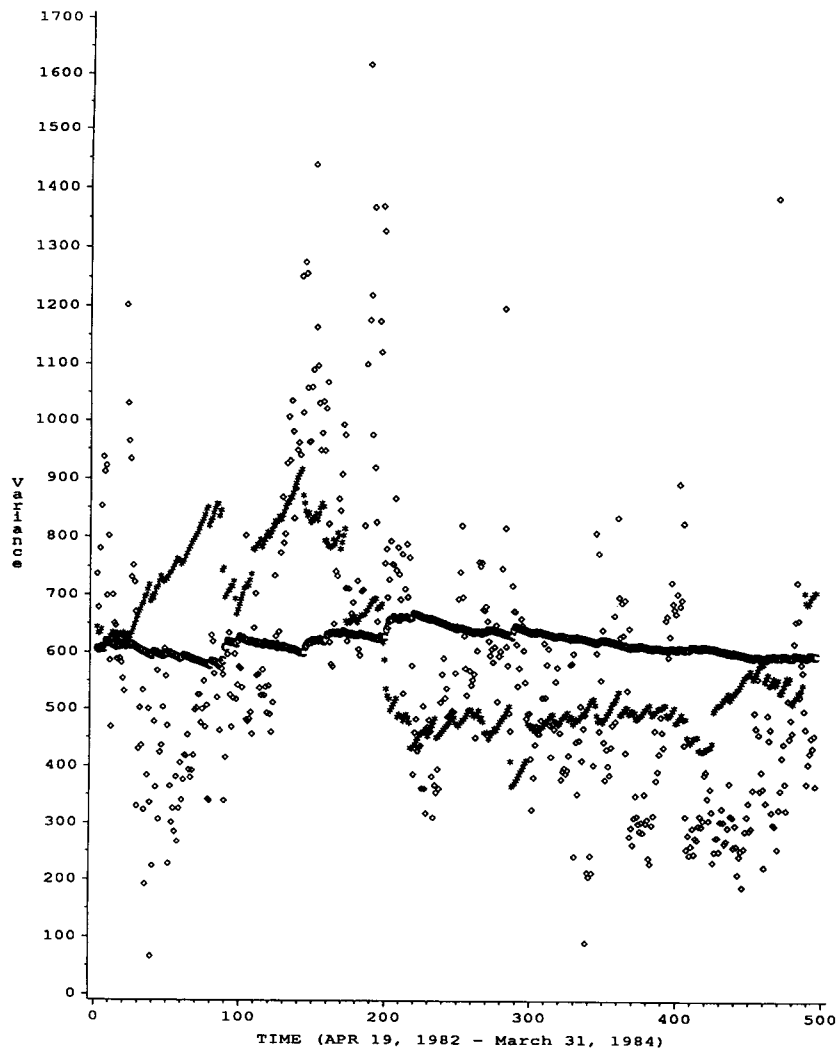
<sup>9</sup> In a study that examines the forecast efficiency of GARCH (in terms of expected utility), West et al. (1990) find that rolling GARCH is superior to updating. The data in that paper are weekly foreign exchange rates.

**Table 3**  
**Comparisons of out-of-sample variance forecasts**

	Implied variance	Updated $G_t$	Rolling $G_t$	$H_t$
A: CSC				
ME	33.48	-11.04	-12.14	-20.32
MAE	187.26	132.05	147.85	133.67
RMSE	232.15	150.61	169.84	150.32
B: DEC				
ME	257.00	188.80	107.23	239.22
MAE	308.37	306.09	331.32	300.41
RMSE	392.84	375.86	389.23	300.72
C: DPT				
ME	281.63	-829.87	-899.04	-327.47
MAE	402.04	905.79	1051.97	593.85
RMSE	541.84	1537.71	1719.35	670.95
D: FDX				
ME	143.41	-22.91	-4.89	-25.89
MAE	179.12	185.86	211.14	184.84
RMSE	241.21	213.12	233.74	212.61
E: NSM				
ME	547.96	115.72	-5.66	141.25
MAE	557.17	367.14	447.95	381.81
RMSE	693.59	535.82	592.99	536.92
F: PDN				
ME	644.95	372.79	136.16	386.43
MAE	689.63	571.96	674.26	584.19
RMSE	917.84	751.40	812.22	760.52
G: ROK				
ME	177.74	-4.38	0.25	-12.99
MAE	203.28	94.51	115.09	99.67
RMSE	243.78	114.91	135.86	114.93
H: STK				
ME	434.22	115.36	59.76	-125.06
MAE	496.36	431.79	520.75	414.72
RMSE	635.48	538.00	616.10	513.18
I: TAN				
ME	203.32	-50.24	-51.02	-53.56
MAE	275.50	279.86	307.21	283.17
RMSE	389.52	350.04	393.33	343.31
J: TOY				
ME	71.15	-79.78	-136.79	-60.62
MAE	241.51	266.82	275.33	256.89
RMSE	296.57	330.84	363.50	319.96

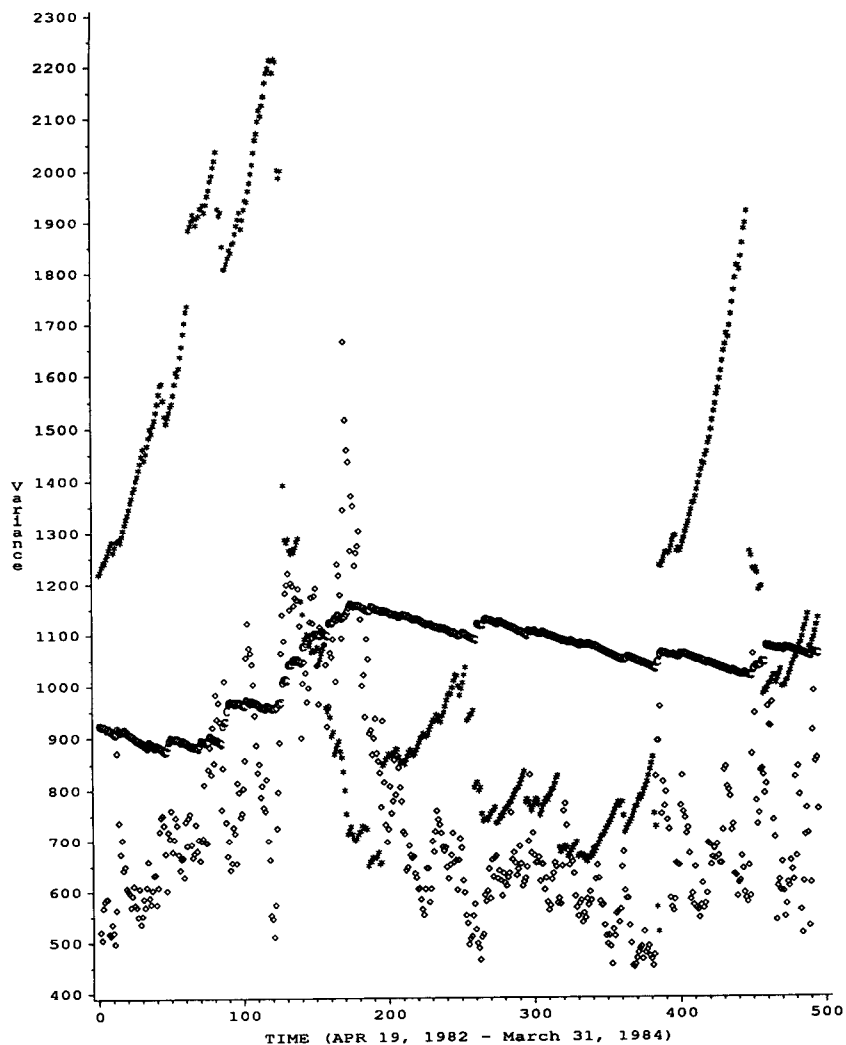
GARCH ( $G$ ), historical ( $H$ ), and implied variance are each being used to forecast the mean of the daily variance over the remaining life of the option. For each day in the sample, each forecast is compared to the actual mean of the daily variance. In this table, only those options with days to maturity of between 90 and 180 days are used. Only call options that are closest-to-the-money are used. The realized variable is measured as the sample average of  $\epsilon^2 = (r_t - \bar{r})^2$  over the remaining life of the option, where  $\bar{r}$  is the unconditional mean of the return process.

As defined in the text, ME refers to mean forecast error, MAE refers to mean absolute error, and RMSE refers to root mean square error. Rolling GARCH forecasts use 300 days prior to the day from which the forecast is being made. Updated GARCH adds an additional observation for each forecast. All returns are daily percentages times 1000.



**Figure 1**  
**Daily variance measures for CSC (April 19, 1982 through March 31, 1984)**

Two forecasts of the average stock-return variance over horizon  $t + N$  are compared with the actual average variance over that horizon for each day in the sample period. We plot the updated sample variance (using at least the past 300 daily returns) and the variance implied from simultaneous stock price and option quotes over the day on the closest-to-the money call option that matures in  $N$  (trading) days ( $64 \leq N \leq 129$ ). Both measures can be thought of as predictors of the average variance over the remaining life of the option. The actual variance of the stock return (which was, in fact, realized) over the forecast horizon is also plotted. The three variance measures plotted are (CCC) the updated sample variance; ( $\diamond \diamond \diamond$ ) the implied variance; and (\*\*\*) the actual variance over the forecast horizon. If the joint null hypothesis of the paper were true, then the implied variance would be an unbiased predictor of the actual realized variance, and the historical sample variance would be orthogonal to the prediction error.



**Figure 2**  
**Daily variance measures for STK (April 19, 1982 through March 31, 1984)**  
 Two forecasts of the average stock-return variance over the period  $t + 1$  through  $t + N$  are compared with the actual average variance over that period. On day  $t$  in the sample, we plot the updated sample variance (using at least the past 300 daily returns) and the variance implied from simultaneous stock price and option quotes over the day on the closest-to-the money call option that matures in  $N$  (trading) days ( $64 \leq N \leq 129$ ). Both measures can be thought of as predictors of the average variance over the remaining life of the option. The actual variance of the stock return (which was, in fact, realized) over the forecast horizon is also plotted. The three variance measures plotted are (CCC) the updated sample variance; ( $\diamond \diamond \diamond$ ) the implied variance; and (\*\*\*) the actual variance over the forecast horizon. If the joint null hypothesis of the paper were true, then the implied variance would be an unbiased predictor of the actual realized variance, and the historical sample variance would be orthogonal to the prediction error.



representative companies, CSC (with relatively low persistence in variance) and STK (with relatively high persistence).

As noted by Fair and Shiller (1990, pp. 375–376), simply comparing out-of-sample forecasts using RMSE has limitations. Further insights into the nature of the different forecast models can be obtained by regressing the realized mean squared residuals on the three alternative out-of-sample forecasts:

$$z_t = \beta_0 + \beta_1 \zeta_t + \beta_2 G_t + \beta_3 H_t + u_t. \quad (7)$$

All variables are as defined before, except note that  $G_t$  is the updated GARCH forecast, conditional on information available at time  $t$ , of the mean variance; rolling GARCH is not used here because updating GARCH dominates it under RMSE, and these two are highly correlated.

This regression is in the spirit of the encompassing literature [Hendry and Richard (1982)]. If a forecast contains no useful information regarding the evolution of the dependent variable, we would expect the coefficient on that forecast to be insignificantly different from zero. The orthogonality restriction implies that the alternative time-series models contain no information not incorporated in the implied variance that can be used to predict realized volatility. Thus, the encompassing regressions are closely related to the in-sample regression tests. However, this design avoids the maturity mismatch problem, and the forecasts are conditioned on information available at period  $t$ . As pointed out by Fair and Shiller (1990), this test also avoids the inherent ambiguity of RMSE comparisons.<sup>10</sup>

Ordinary least squares (OLS) is a consistent estimator of these regression coefficients. However, because the forecast horizon exceeds the frequency of the available data, the error term will be a moving average process. Since generalized least squares is inconsistent (as the forecast errors are not strictly exogenous), we take the approach of obtaining a consistent estimator of the variance covariance matrix of the OLS estimators by generalized method of moments (GMM). An early example of this procedure is Hansen and Hodrick (1980).

To construct this consistent estimator, we use the Bartlett kernel approximation to the spectral density at frequency 0 of the residuals to weight lagged values as suggested by Newey and West (1987). As noted by Andrews (1991), the asymptotic theory for this estimator

<sup>10</sup> Fair and Shiller (1990) use this empirical strategy to infer the information content of alternative models of real GNP growth. In particular, they examine the forecasting performance of a structural model and various time-series models of aggregate output. Although the economic issue in our article differs from Fair and Shiller, it is clear that the questions we ask are analogous to theirs. Thus, their methods of inference are appropriate for our analysis.

**Table 4**  
**Encompassing tests of variance forecasts**

$\beta_0$ (GMM <i>t</i> )	$\beta_1$ (GMM <i>t</i> )	$\beta_2$ (GMM <i>t</i> )	$\beta_3$ (GMM <i>t</i> )	$r^2$ <i>M</i> *
A: CSC				
2014.269 (5.39)	0.276 (3.81)	-2.580 (-4.32)		0.357 50
2950.717 (4.72)	0.272 (3.81)	0.522 (0.98)	-4.566 (-3.47)	0.451 54
B: DEC				
981.366 (7.52)	0.382 (1.60)	-1.012 (-3.96)		0.222 42
1069.920 (3.26)	0.361 (1.16)	-0.890 (-2.49)	-0.327 (-0.30)	0.227 66
C: DPT				
445.920 (3.18)	0.531 (1.82)	0.129 (2.92)		0.418 60
3692.963 (2.64)	0.970 (3.97)	0.093 (2.12)	-2.360 (-2.49)	0.575 51
D: FDX				
1585.291 (3.38)	1.116 (2.35)	-2.641 (-2.71)		0.304 75
3161.79 (3.87)	0.831 (2.99)	1.736 (1.29)	-6.938 (-2.70)	0.508 85
E: NSM				
2084.680 (3.65)	1.462 (4.28)	-1.659 (-2.80)		0.387 59
4955.058 (9.73)	0.490 (2.21)	0.366 (1.57)	-4.093 (-7.17)	0.839 73
F: PDN				
1997.741 (2.96)	0.118 (0.52)	-0.632 (-0.93)		0.030 184
2388.188 (2.63)	0.212 (0.79)	1.074 (1.94)	-2.175 (-2.34)	0.086 196
G: ROK				
1988.798 (6.69)	-0.452 (-0.48)	-3.775 (-5.02)		0.356 98
2141.949 (4.81)	-0.100 (-1.24)	-1.556 (-0.80)	-2.504 (-1.02)	0.375 119
H: STK				
2491.690 (4.24)	1.300 (3.01)	-2.166 (-3.27)		0.402 47
5073.759 (3.41)	0.619 (2.28)	-0.389 (0.96)	-3.792 (-2.54)	0.637 86
I: TAN				
1858.705 (4.24)	0.748 (3.01)	-1.907 (-3.27)		0.215 47
5030.08 (3.41)	0.390 (2.28)	0.730 (0.96)	-6.351 (-2.54)	0.488 86
J: TOY				
824.120 (4.23)	1.171 (2.56)	-1.273 (-2.85)		0.227 59
4053.326 (4.91)	0.660 (2.57)	0.028 (-0.02)	-5.999 (-4.32)	0.555 87

Model:  $z_t = \beta_0 + \beta_1 \zeta_t + \beta_2 G_t + \beta_3 H_t + u_t$ .

requires that the lag length used to accumulate the variance go to infinity with the sample size. Therefore, selection of lag length is not obvious. We use the method of Andrews (1991) to estimate optimally this lag length.<sup>11</sup>

Table 4 contains the results of the optimal forecast weighting tests. Two regressions are reported. In the first,  $\beta_3$  is constrained to be 0, and the second is the unrestricted model. In the table, we report the automatic bandwidth obtained from the Andrews procedure,  $M^*$ ; the  $t$ -statistics are computed by using  $M^*$  in the Newey–West weighting scheme. All tests were conducted with a bandwidth of 32, 66, and 132, as well. None of the following inferences are sensitive to the choice of bandwidth.

In general, the optimal out-of-sample forecast of mean realized volatility places a statistically significant positive weight on the implied variance from the options market, no significant weight on the GARCH forecast, and a large significant negative weight on the updated sample variance. DEC, PDN, and ROK are exceptions to this pattern. Only DEC has a significantly negative coefficient on the GARCH forecast. In all 10 cases the intercept is positive and statistically significant. This result is consistent with earlier evidence that during this period variance forecasts are biased downward.

These regression results suggest the importance of the point made by Fair and Shiller (1990). RMSE comparisons suppress a large amount of information about the problem of constructing an optimal forecast. Despite having the lowest RMSE in only two cases, implied variance has significant forecast weight in seven cases. Symmetrically, note that in the two cases where implied variance had the lowest RMSE

<sup>11</sup> Andrews (1991) has developed the asymptotic theory to estimate the optimal lag length or bandwidth for a given sample of size  $T$  as a function of the autocovariance structure of the matrix  $V$ , where  $V = \epsilon'X$ ,  $\epsilon$  is a  $T \times 1$  vector of the OLS residuals, and  $X$  is the  $T \times K$  matrix of regressors. In our case, the first column is a vector of 1. We estimate a univariate AR(1) process for each column of  $V$ . The  $K$  AR(1) coefficients and residual variances are then used to compute  $\hat{\alpha}(1)$  by using Equation (6.4) in Andrews (1991, p. 835). We use weights of 1 for  $p = 2, \dots, K$  and 0 for  $p = 1$  (which Andrews notes yields a scale-invariant covariance matrix, page 834). Since we are using the triangular (Bartlett) kernel estimator as in Newey and West (1987), we obtain the optimal bandwidth by plugging  $\hat{\alpha}(1)$  into Andrews' Equation (6.2) (page 834). The result of this equation is equal to  $M^* + 1$ , where  $M^*$  is the optimal bandwidth used to compute the variance-covariance matrix, as in Newey and West.  $M^*$  increases with the cubed root of  $T$ .

←

$\beta_3$  is restricted to be 0 in the first model and a free parameter in the second.  $z_t$  is the realized mean  $(r_t - \bar{r})^2$  during the period  $t + 1$  through  $t + N$ , where  $t + N$  is the maturity date of the option.  $\zeta_t$  is the implied variance from all quotes on the at-the-money, intermediate-term options (that expire at time  $t + N$ ) on day  $t$ .  $G_t$  is  $N$ -step ahead (updated) GARCH (1, 1) variance forecast (conditional on information available at time  $t$ ).  $H_t$  is the updated sample variance measured through day  $t$ .  $M^*$  is the optimal bandwidth used to estimate the Bartlett kernel in the Newey–West variance-covariance matrix [following Andrews (1991)]. GMM  $t$  refers to the Student's  $t$ -statistic computed using this matrix. All returns are daily percentages times 1000. All variances are on a daily basis.

**Table 5**  
Implied volatility forecast "errors" and the level of variance

Company	$\beta_0$ (GMM $t$ )	$\beta_1$ (GMM $t$ )	$r^2$ $M^*$
	3601.40	-5.77	0.26
CSC	(4.36)	(-4.32)	24
	905.16	-1.54	0.12
DEC	(6.63)	(-4.39)	38
	3557.52	-2.16	0.28
DPT	(2.35)	(-2.18)	54
	2869.53	-4.81	0.48
FDX	(4.41)	(-4.43)	75
	4622.75	-3.73	0.82
NSM	(10.62)	(-9.61)	46
	2381.30	-1.67	0.18
PDN	(2.57)	(-2.11)	101
	1721.33	-3.64	0.11
ROK	(1.75)	(-1.54)	47
	5092.60	-4.47	0.62
STK	(5.68)	(-5.28)	77
	4790.36	-5.74	0.48
TAN	(5.45)	(-5.37)	29
	3987.13	-6.13	0.51
TOY	(5.56)	(-5.64)	69

Model:  $z_t - \zeta_t = \beta_0 + \beta_1 H_t + u_t$ .

$z_t$  is the realized mean  $(r_t - \bar{r})^2$  during the period  $t + 1$  through  $t + N$ , where  $t + N$  is the maturity date of the option.  $\zeta_t$  is the implied variance from all quotes on the at-the-money, intermediate-term options that expire at time  $t + N$  on day  $t$ .  $H_t$  is the updated sample variance measured through day  $t$ .  $M^*$  is the optimal bandwidth used to estimate the Barlett kernel in the Newey-West variance-covariance matrix [following Andrews (1991)]. GMM  $t$  refers to the Student's  $t$ -statistic computed using this matrix. All returns are daily percentages times 1000. All variances are on a daily basis.

(DPT and TOY), the historical variance has significant forecasting power.

The results reported in this table suggest that the joint null hypothesis of market efficiency and the HW model is rejected at standard significance levels. However, it is not the case that filtering the data with the HW model is uninformative. To an agent confronted with deriving an optimal variance forecast, the norm here is to exploit information in both the historical path of stock prices and contemporaneous option and stock prices.

An alternative regression is run, and the results are provided in Table 5. Here the dependent variable is defined as the variance-measured model error:  $z_t - \zeta_t$ . The regressor is the updated sample variance estimate at time  $t$ . Test statistics are derived from the GMM covariance matrix as in Table 4. Except for ROK (also an outlier in the optimal forecast analysis, with an insignificant negative weight on  $\zeta_t$ ), the intercept in this regression is positive and significant. The coefficient on the current variance estimate is negative and significant. Note that the order of magnitude of the intercept is the same as the dependent variable itself. The  $r^2$  values reported in Table 5 are high.

In half of the cases, more than 45 percent of the variance-measured pricing error is explained by the evolution of the current variance.

#### 2.4 Interpretation of results

The results of the different experiments conducted are in general inconsistent with the joint hypothesis of the stochastic variance option pricing model and informational efficiency. This inference is robust across the in- and out-of-sample experimental designs. The tests uncover two facts for the 10 companies in the sample:

1. Implied variance tends to underpredict realized variance (see mean error in Table 3 and the significantly positive intercept in Table 5).
2. Forecasts of variance from past returns contain relevant information not contained in the market forecast constructed under the HW assumptions. The optimal weight placed on these forecasts of realized variance is negative (see Tables 4 and 5).

We refer to the results in Table 2 and ask whether these characteristics can be due solely to the bias inherent in our procedure for implying volatilities under the assumptions of zero correlation between the instantaneous rate of change in the stock price and the instantaneous variance and linearity of the pricing model. The smallest percentage bias in the data is for CSC. Here the mean error of implied variance from Table 3 is 33.48, and the level of variance is 590.52 (see Table 1, specification 1), a percentage error of 5.67. From Table 2, under the joint null, we would expect a percentage error of  $-0.25$  percent; if the estimated percentage error fell between  $-0.67$  and  $0.18$  percent, we would be unable to reject the null at an approximate 5 percent level of significance.<sup>12</sup> For Tandy, as a more representative company, the expected bias under the null is 0.94 percent, with a two-standard-error range of 0 to 1.36 percent. The bias in the data is 27.54 percent. In all 10 cases, the out-of-sample forecasting error lies well outside the calibrated two-standard-error confidence interval. Thus, the documented underprediction of implied variances cannot be attributed to the two potential sources of bias.

Also, from Table 2, we ask whether the empirical relationship between the forecast error of the implied variance and the current variance is solely a result of the approximation procedure used to imply variances. It is possible that the negative coefficient on  $H_t$  (or  $G_t$ ) in the encompassing regressions picks up a tendency for the bias

<sup>12</sup> Recognizing that the unconditional variance is estimated with sampling error, add two standard errors to this estimate, and the percentage error is 5.1. The inference is unchanged.

in implied variance, though small, to change with the level of variance. Note from Table 2 that in all 10 cases the inherent percentage bias is inversely related to the level of variance,  $V(0)$ . However, the magnitude of this phenomenon is very different from that manifest in the data. Federal Express is a typical case. From the three sets of values in Table 2 (generated with 1000 draws, 100 times), the change in absolute bias divided by the change in level of variance is  $-0.008$ . From Table 5, the change in the forecast error relative to a change in historical variance is  $-4.81$ .<sup>13</sup> Therefore, whereas we expect a negative coefficient on historical variance, the size of this coefficient computed from actual data is too large to be the result of the inherent biases in the experimental design.

Since the biases inherent in our extraction procedure are unlikely to be the reason for the rejection of the null hypothesis, we can speculate as to the potential causes. One interpretation of the statistically significant negative effect of the updated sample variance is that market participants totally ignore the information contained in past realizations of returns. An alternative interpretation is that the market overreacts to recent volatility shocks: too much weight is placed on the recent past of the variance process. A current increase to volatility raises the sample variance, but the negative  $\beta_3$  suggests that this shock is temporary. However, options traders impute a permanence to the shock, leading to an underprediction of variance [Stein (1989)].

Given informational efficiency, our results can be explained by the existence of a risk premium applied to the nontraded variance process. Recall that an assumption underlying the use of the implied variance from the model as an instrument for the market's forecast of variance is that volatility risk is unpriced. The option price is independent of risk preferences if agents are risk-neutral, or if the instantaneous variance is uncorrelated with aggregate consumption and, therefore, is uncorrelated with marginal utility of wealth. If this assumption is false, then observed option prices will include a risk premium. For example, if variance uncertainty gives negative utility to traders, the observed option price will be lower than the risk-neutral price, *ceteris paribus*. When the observed price is applied to the Black-Scholes formula, the implied variance will be correspondingly lower than actual variance. As noted, we document this underprediction by the implied variance for all stocks in the sample. The negative coefficient on the current level of variance means that

<sup>13</sup> The regression coefficient reported in Table 5 is a first derivative analogous in its interpretation to the ratio of changes reported. Although the magnitude of change in the variance level is large in Table 2, the size and sign of the pseudoderivative are virtually identical across companies and independent of the size of change in  $V(0)$ .

the risk premium is time dependent. Specifically, the implied variance rises relative to the future (realized) volatility as the stock's variance rises. Thus the variance risk premium embedded in option prices diminishes as the stock's variance increases.<sup>14</sup>

Melino and Turnbull (1990) use numerical methods to evaluate a partial differential equation in the spirit of Garman (1976), and they find evidence to support the notion that a nonzero risk premium on the variance process exists in the Canadian dollar–U.S. dollar exchange rate option market. They restricted the price of variance risk to be a constant. Our results suggest that such a risk premium is time-varying in the stock market.

Our empirical analysis is confined to constructing an optimal variance forecast. Since we reject the joint null hypothesis of market efficiency and stochastic variance option pricing model-based pricing of these options, these results do not indicate that one could take the optimal variance forecasts and use them to “beat” the option market. It is plausible that the market's forecast as embodied in option prices is optimal, but that filtering the prices through the simple option pricing model distorts the forecast. Thus, if one used the “optimal forecast” from the regressions, the profits generated by that strategy would be offset by volatility risk if this trader's utility function is the same as the market's.

Although we have motivated this experiment as a test of the restrictions of an asset pricing model, the tests also provide insights into the nature of the variance process for these individual stocks. The poor out-of-sample performance of GARCH at the 90- to 180-calendar-day horizons appears inconsistent with its good in-sample fit and good forecast performance at short horizons [see Akgiray (1989)]. This set of results is consistent with a variance process that is subject to highly persistent shocks at low frequencies and quickly dampening shocks at high frequencies. This property of stock-return variances and its relationship to GARCH has been suggested by Lamoureux and Lastrapes (1990). Theoretical analysis of this phenomenon is provided by Nelson (1992). GARCH treats all innovations equally; therefore, it overstates the persistence of high-frequency shocks, which leads to excellent short-term forecasts but poor long-term forecasts.

As a final note, Back (1993) has developed an equilibrium model of informed trader behavior in the spirit of Kyle (1985), where the monopolistically informed trader may trade with an uninformed market maker in either the stock or option market. Back shows that the only equilibrium that does not generate arbitrage opportunities for

<sup>14</sup> There is no reason to suspect a priori that this risk premium would be the same across stocks. It is a function of the correlation between the variance process and marginal utility. To the extent that stock variances include common factors, the risk premiums will be related.

the informed trader is one in which the stock's variance is stochastic and in which the option price contains information about the future variance that is not publicly available elsewhere. Thus, the option allows a finer partition of information than is available from an information set that excludes the current option price. Empirically, this theoretical result could be rejected if, for example in our analysis, we found that the optimal forecast weight on implied variance could not be statistically distinguished from zero. For 7 of the 10 companies, the implied variance has a statistically positive weight. This result is consistent with Back's model.

### 3. Conclusions

We examine the joint hypothesis of a class of stochastic volatility option pricing models and informational efficiency in the options markets using a criterion function based on the variance of the underlying stock returns. To examine this hypothesis, we represent the subjective variance of the market as the implied variance from the data and the option pricing model. By utilizing discrete-time simulations of the continuous-time return process, calibrated to our data, we show that potential biases between these two variance concepts are small and do not affect the inferences of our tests of the joint hypothesis.

Using both in-sample and out-of-sample tests, we reject the implication of the hypothesis that available information cannot be used to improve the market's variance forecast embedded in observable prices *as measured by this class of models*. This result is robust across the different test designs and to inherent measurement bias and is consistent with the results of Melino and Turnbull (1990) and Day and Lewis (1992). If the market is efficient, our results suggest that the fundamental assumptions of the model are not sufficient to account for the properties of the data. In particular, equilibrium models of option pricing that do not assume investor indifference to volatility risk appear necessary to reconcile the theory and data. The data suggest that the market premium on variance risk is time varying; it is a decreasing function of the level of the stock's variance. Results concerning the nature of this risk premium are uniform (and strikingly similar) across the 10 stocks used in our analysis.

Although the option pricing model is rejected as the price-determining market mechanism, filtering the data with the simple model does contain useful information that is not contained in the historical price process of the underlying stock for forecasting the stock's variance over a 180- to 90-calendar-day horizon. This result has normative implications for optimal variance forecast rules, even in the absence



of the “correct” equilibrium option pricing model capable of explaining the data.

#### References

- Akgiray, V., 1989, “Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts,” *Journal of Business*, 62, 55–80.
- Andrews, D. W. K., 1991, “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation,” *Econometrica*, 59, 817–858.
- Back, K., 1993, “Asymmetric Information and Options,” *Review of Financial Studies*, 6, forthcoming.
- Biegler, L. J., and J. E. Cuthrell, 1985, “Improved Feasible Path Optimization for Sequential Modular Simulators. II: The Optimization Algorithm,” *Computers and Chemical Engineering*, 9, 257–267.
- Black, F., and M. S. Scholes, 1973, “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 81, 637–654.
- Bollerslev, T., 1986, “Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics*, 31, 307–327.
- Brenner, M., and D. Galai, 1987, “On the Prediction of the Implied Standard Deviation,” in *Advances in Futures and Options Research*, JAI Press, Greenwich, Conn.
- Cox, J. C., and M. Rubinstein, 1985, *Options Markets*, Prentice-Hall, Englewood Cliffs, NJ.
- Day, T. E., and C. M. Lewis, 1992, “Stock Market Volatility and the Information Content of Stock Index Options,” *Journal of Econometrics*, 52, 267–287.
- Engle, R. F., 1982, “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation,” *Econometrica*, 50, 987–1007.
- Fair, R. C., and R. J. Shiller, 1990, “Comparing Information in Forecasts from Econometric Models,” *American Economic Review*, 80, 375–389.
- Feinstein, S. P., 1989, “A Theoretical and Empirical Investigation of the Black–Scholes Implied Volatility,” unpublished thesis, Yale University.
- Garman, M., 1976, “A General Theory of Asset Valuation under Diffusion State Processes,” working paper, University of California, Berkeley.
- Hansen, L. P., and R. J. Hodrick, 1980, “Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis,” *Journal of Political Economy*, 88, 829–853.
- Hendry, D. F., and J. F. Richard, 1982, “On the Formulation of Empirical Models in Dynamic Economics,” *Journal of Econometrics*, 20, 3–33.
- Heston, S., 1993, “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options,” *Review of Financial Studies*, 6, 327–343.
- Hull, J., and A. White, 1987, “The Pricing of Options on Assets with Stochastic Volatilities,” *Journal of Finance*, 42, 281–300.
- Johnson, H., and D. Shanno, 1987, “Option Pricing When the Variance is Changing,” *Journal of Financial and Quantitative Analysis*, 22, 143–151.
- Kyle, A. S., 1985, “Continuous Auctions and Insider Trading,” *Econometrica*, 53, 1315–1335.
- Lamoureux, C. G., and W. D. Lastrapes, 1990, “Persistence in Variance, Structural Change, and the GARCH Model,” *Journal of Business and Economic Statistics*, 8, 225–234.
- Melino, A., and S. Turnbull, 1990, “Pricing Foreign Currency Options with Stochastic Volatility,” *Journal of Econometrics*, 45, 239–265.

- Nelson, D. B., 1990, "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7-38.
- Nelson, D. B., 1992, "Filtering and Forecasting with Misspecified ARCH Models. I: Getting the Right Variance with the Wrong Model," *Journal of Econometrics*, 52, 61-90.
- Newey, W. K., and K. D. West, 1987, "A Simple, Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-708.
- Scott, L. O., 1987, "Option Pricing When the Variance Changes Randomly," *Journal of Financial and Quantitative Analysis*, 22, 419-438.
- Stein, E. M., and J. C. Stein, 1991, "Stock Price Distributions with Stochastic Volatility: An Analytic Approach," *Review of Financial Studies*, 4, 727-752.
- Stein, J. C., 1989, "Overreactions in the Options Market," *Journal of Finance*, 44, 1011-1023.
- West, K. D., H. J. Edison, and D. Cho, 1990, "A Utility Based Comparison of Some Models of Exchange Rate Volatility," working paper, University of Wisconsin.
- Whaley, R. E., 1982, "Valuation of American Call Options on Dividend Paying Stocks: Empirical Tests," *Journal of Financial Economics*, 10, 29-58.
- Wiggins, J. B., 1987, "Option Values under Stochastic Volatility: Theory and Empirical Estimates," *Journal of Financial Economics*, 19, 351-372.