NBER WORKING PAPER SERIES

VALUATION OF VARIANCE FORECASTS WITH SIMULATED OPTION MARKETS

Robert F. Engle

Che-Hsiung Hong

Alex Kane

Working Paper No. 3350

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1990

This paper is part of NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3350
May 1990

VALUATION OF VARIANCE FORECASTS WITH SIMULATED OPTION MARKETS

ABSTRACT

An appropriate metric for the success of an algorithm to forecast the variance of the rate of return on a capital asset could be the incremental profit from substituting it for the next best alternative. We propose a framework to assess incremental profits for competing algorithms to forecast the variance of a prespecified asset. The test is based on the return history of the asset in question. A hypothetical insurance market is set up, where competing forecasting algorithms are used. One algorithm is used by each hypothetical agent in an "ex post ante" forecasting exercise, using the available history of the asset returns. The profit differentials across agents (in various groupings) reflect incremental values of the forecasting algorithms.

The technique is demonstrated with the NYSE portfolio, over the period of July 22, 1966 to December 31, 1985. For the limited set of alternative specifications, we find that GARCH(1,1) yields better profits than the 3 competing specifications. The profit from pricing one-day options on the NYSE portfolio significant. The evidence also suggests that using a limited estimation period may be preferable to estimating specification parameters from all available observations. Finally, the hedging activity that requires a variance determined hedge ratio is an important component of the success of a variance forecast-algorithm.

Robert F. Engle
Dept. of Economics
D-008
University of California
La Jolla, CA 92093-0519

Che-Hsiung Hong
Citibank
399 Park Avenue
New York, NY 10043

Alex Kane
Graduate School of
International Relations and
Pacific Studies
University of California
La Jolla, CA 92093-0519

Introduction

When probability distributions of asset rates of return are time varying, forecasts of portfolio variance must be a priority item for investors. Even clients who restrict themselves to passive strategies would need periodic variance forecasts to calibrate efficient asset allocation (e.g., Bodie [1989]).

An algorithm to forecast the variance of the rate of return on a capital asset would include: (i) a specification of the return-generating process, and (ii) a procedure to estimate the parameters and compute implied forecasts, based on available information at the time the forecast is made. A criterion for choosing between any pair of competing algorithms is the incremental profit of a switch from the lesser to the better one.

In this paper we propose a technique to assess incremental profits for a set of competing algorithms, and thereby evaluate algorithms used for forecasting the variance of a given portfolio. The proposed technique resorts to the role of volatility in the pricing of contingent claims. It incorporates into the analysis considerations of hedge positions which, too, depend on variance forecasts. If proven reliable, our technique would be useful for financial decision making and for research in capital markets. We demonstrate the technique with the NYSE

---

1. We thank Robert L. McDonald for helpful discussions and suggestions, and the participants in the UCSD conference on modeling the volatility in asset pricing for comments.

portfolio experience over the period of July 3, 1962 to December 31, 1985.

We estimate incremental profits of algorithms that forecast portfolio variance by setting up a hypothetical insurance market. Each forecasting algorithm is used exclusively by one hypothetical agent in an "ex post ante" forecasting exercise. The hypothetical agents prepare a forecast for every period of the exercise using past observations of the portfolio rates of return. Using the designated algorithm, each agent sets a price for options of one period maturity, on one dollar's worth of the target portfolio. In the beginning of each period, upon examination of the entire set of forecast-induced option prices, each agent buys any option believed to be underpriced. In the "ex post" part of the exercise, actual subsequent portfolio returns are used to settle the agent accounts. The period profits and losses accumulate in the agent accounts over the entire exercise.

Profits and losses that accumulate over the exercise period represent the relative success of agents, and hence of the forecasting algorithms that they represent. In the present experiment, one-day options on $1 shares of the NYSE are priced from the following specifications of the NYSE return process: (i) moving average of squared daily returns, (ii) ordinary least squares: the variance rate is estimated from the standard error of an AR(1) on the daily rate, (iii) ARMA(1,1) on squared errors, and (iv) GARCH(1,1). Each of the four specifications is used by 3 algorithms that differ in the number of past observations in the rolling sample used to estimates the specification parameters. The 3 sample lengths are: (i) 300 days, (ii) 1,000 days, and (iii) the entire available history, up to 4901 observations but no less than 1,000 days.

The main objective of the work presented here is to introduce this new technology. We have no commitment to the algorithms that were (quite arbitrarily) chosen to compete. For this set, results of the experiment with the NYSE portfolio clearly favor the GARCH(1,1) specification. The worst performer of the specification set has been ARMA(1,1). The economic performance of the various specifications in the experiment has been significantly different, suggesting that inferences about the relationship between stock price and changes in volatility, as in Poterba and Summers [1986], should be reexamined with respect to the specification used to estimate the time-path of stock variance. Evidence from the experiment also indicates that some restriction on the length of the rolling sample may lead to improved forecasts.

We allow agents to hedge their variance-forecast driven transactions by taking positions in the NYSE stock. Evidence from the experiment suggests that imperfect hedge ratios affect the risk reduction capacity in an economically significant way. The hedging activity itself helps to distinguish better forecasting algorithms.

Theoretical arguments in favor of the proposed forecasting valuation technique are presented in section I. Section II develops the mechanics of the valuation method, and describes the details of the experiment involving the forecasting of the variance of the NYSE portfolio. Presentation and discussion of the experiment results can be found in Section III. Summary and discussion of future research conclude the paper.

I. Theoretical Underpinnings

In the genre of testing for the choice of best forecasts, the case of the variance of asset returns is unique. While the period variance of the rate of return of an asset is unobservable, its observed, realized rate of return allows us to test variance-driven option prices. The role of variance in "no-arbitrage" pricing of contingent claims suggests that the best forecast algorithm should be decided on performance in correctly pricing contingent claims on the asset in question.

This approach is different from testing the correlation between option-price implied volatility (perhaps in conjunction with other variance forecasts) and some measure of subsequently realized variance. We take the position that variance over any period is unobservable, and that we can do without ever observing it directly. Rather, we ask, which return specification and estimation technique from available data (that may include some transformation of implied volatility from some traded contingent claims) would lead to better pricing of some specified contingent claims. The volatility predicted from that specification, into any future period, is taken to constitute the variance of the rate of return on the asset over the future period.

In efficient capital markets; when the variance of the rate of return on a target asset is constant over time; and where options on the asset are traded, the option price implies an estimate of the asset variance. Under such ideal circumstances, indeed by the definition of market efficiency, option prices would yield preferred estimates of asset variances. As a practical matter, Schmalensee and Trippi [1978], and others, have shown that stock-option implied variances yield better forecasts of standard deviation than do simple estimates of standard deviation from past returns.

Yet variance forecasts would be particularly valuable when variances do vary over time. When the variance is time varying, option pricing becomes more difficult because the hedge ratio that allows arbitrage pricing depends on the unknown variance. It still is possible to compute the implied <u>expected</u> variance over the life of an efficiently-priced option, if surprise changes in variance during the life of the option would be non-systematic (see Wiggins [1987] and Hull and White [1989]). Such circumstance is not guaranteed, particularly when the asset in question is a well diversified portfolio.

The upshot is that forecasts of volatility must be empirically tested with an appropriate loss function in order to establish the best alternative forecast. Indeed, with time varying variances, we cannot even take for granted that the market is efficient with respect to any variance forecast. Such determination can only be reached when no known variance-forecasting algorithm can be used to profit from trading in the asset and its derivatives. The technique proposed here can also be applied to the question of market efficiency with respect to specific variance forecasts.

Using realized rates of return to test the rationality of option prices so as to evaluate the accuracy of variance forecasts, is really a joint test of the variance forecast and the option-pricing formula that is used. This fate befalls all theory based empirical work.

Mean-variance models of portfolio theory tell us that investor utility from a universe of capital assets will be lower if: (i) uncertainty about the variance of one or more assets increases, and/or (ii) the investor is misinformed about the probability distribution of the variance of one or more of the assets. In the first case, when investors are informed that uncertainty about asset variance is increased, they will shift portfolio demands. New equilibrium prices will prevail, and the issue will become history.

In the second case, when a significant number of investors are misinformed, asset prices may be inefficient with respect to some information about variance. Well diversified portfolios will be underpriced if the variance is overestimated, and vice versa. Small sets of assets (in terms of market value), will be mispriced in the same way if the information asymmetry involves covariance with a well diversified portfolio.

As information asymmetry about asset variance develops, investors with superior information will attempt to gain from it. If options on the mispriced asset are traded, then a natural strategy would be to set up a hedged position that includes the asset and options. This position cannot be made entirely risk-free, because the hedge ratio also would depend on the variance; furthermore, an exact arbitrage price might be unknown if unanticipated changes in variance, over the life of the option, are systematic. Still, a strategy that consistently employs superior variance information to rebalance an active portfolio of variance-driven hedged positions, may be expected to yield a superior risk-reward ratio.

Even when options on an asset are not regularly traded, a developing asymmetry in variance information could induce such trade. Put another way, the degree of heterogeneity in expectations of asset variance is a factor in the demand for a market in options on the asset. Moreover, we may expect consistently better "variance analysts" to be gainfully employed by traders of contingent claims on the asset that they specialize in.

In sum, when it comes to time-varying variance, the profit potential of a better informed investor will depend on: (i) the degree of asymmetry of variance information, (ii) the option-pricing formula that is implicit in the market place (perhaps unknown), with the sensitivity of implied option prices and hedge ratios to the variance of the underlying asset.

We chose for our first experiment short-term variance forecasts, utilized to price short-lived options. We realize that tests of longer-range forecasts, to be used for pricing longer-maturity options, may be of greater economic value. The experiment and analysis in this paper, being mainly concerned with the technique (rather than with identifying optimal variance forecasts), relates short (one-day) variance forecasts to profit opportunities from pricing one-day options on the NYSE portfolio.

Estimates of the relative profitability of alternative variance-forecast algorithms depend on which option-pricing formula is used. At the same time, the validity of the experiment does not depend on whether the user knows with certainty that the formula used is actually implicit in market prices; only that it is the one he would bet on if and when he had to write contingent claims on the asset. It is understood, however, that further experimentation will be required whenever a different option-pricing formula comes into favor.

Another issue of methodology is transitivity in the dominance relationship among variance forecasts. When the experiment is performed with more than two hypothetical agents, representing more than two competing forecast algorithms, the winner is not guaranteed to dominate every agent in a different grouping, or in pair-wise competition. This issue has to be empirically addressed.

Finally, the Jensen Inequality presents itself in this valuation technique on two levels. First, the square root of a conditionally unbiased variance forecast, will be conditionally downward biased. If a specification of a return generating process produces conditionally unbiased forecasts of variance, then an insurance underwriter who uses forecasts of standard deviation in pricing contingent claims will have to adjust the square root of that forecast. Another reason to adjust the forecast (upwards) arises from the non linearity (concavity) of the option-pricing formula in standard deviation.

In order to adjust a forecast for Jensen's Inequality, one needs to estimate the variance of the forecasting error. Estimating the variance of a variance forecast-error presents a problem, since the realized variance is unobservable. We elected not to include Jensen Inequality adjustments in this experiment, since our focus here is on the technique[2].

II. Application of the Valuation Method to the NYSE Portfolio

2. A crude correction for Jensen's inequality resulted in very little change in the relative profitability of the forecast algorithms presented below. This is not sufficient evidence, however, to discard the problem.

It is best to describe the valuation methodology as it applies to a specific experiment. Here we apply it to the NYSE portfolio, using daily returns from July 3, 1962, to December 31, 1985. We begin by describing the competing variance-forecast algorithms.

II.1 The set of competing variance forecasts

We test 4 specifications for the return generating process of the NYSE portfolio, and 3 alternative lengths of rolling samples from past data to update estimates of the specification parameters. The 4 specifications are detailed below with the following notation: Z_t denotes the portfolio return; a subscript t dates observations and forecasts; a subscript n on a forecast, or parameter estimate, refers to the length of the rolling sample that is used every day to produce the variance forecast for the next day; finally, σ^2 denotes a variance forecast.

(A) Moving Average Variance: the "MA" model

The forecast for the next-day variance, using the most recent n observations is:

$$\sigma^2_{n,t+1} = \frac{1}{n-1} \sum_{t-n+1}^{t} (Z_t - \bar{Z}_{nt})^2 , \qquad (1)$$

where

$$\bar{Z}_{nt} = \frac{1}{n} \sum_{t-n+1}^{t} Z_t . \qquad (2)$$

(B) Ordinary Least Square: the "OLS" model

The rate of return is assumed to follow an AR(1),

$$Z_t = a_{n,t} + b_{n,t}Z_{t-1} + e_t \ , \qquad (3)$$

and the forecast is:

$$\sigma^2_{n,t+1} = \frac{1}{n-1} \sum_{t-n+1}^{t} e_t^2 \ . \qquad (4)$$

(C) ARMA(1,1) in the Squared Residual: the "AR" model

As in the OLS model,

$$Z_t = a_{n,t} + b_{n,t}Z_{t-1} + e_t \ , \qquad (1)$$

except that the squared residual follows:

$$e_t^2 = w_{n,t} + v_{n,t}e_{t-1}^2 + u_t - d_{n,t}u_{t-1} \ . \qquad (5)$$

The inclusion of this process was motivated by the work of Poterba and Summers [1986], where the sample variance of the residual was assumed to follow an AR(1). The variance forecast will be:

$$\sigma^2_{t+1} = w_{n,t} + v_{n,t}e_t^2 - d_{n,t}v_t \ . \qquad (6)$$

(D) GARCH(1,1): the "ARCH" model

The ARCH (auto regressive conditional heteroskedasticity) family of specifications was first proposed by Engle [1982]. A sample of its increasing use for rates of return on capital assets can be found in Bollerslev, Chou, Jayaraman and Kroner [1990], Bollerslev Engle and Wooldridge [1988], Chou [1988], French, Schwert and Stanbaugh, Hong [1987], and references cited there.

As in the OLS model,

$$Z_t = a_{n,t} + b_{n,t} Z_{t-1} + e_t \ , \qquad (1)$$

and the forecast will be:

$$\sigma^2_{t+1} = w_{n,t} + v_{n,t} e^2_t + d_{n,t}(h^2_t - e^2_t) \ . \qquad (7)$$

The parameters of (1) and (7) are estimated by maximum likelihood, assuming e_t to be conditionally normal. We note that (6) and (7) provide identical forecast equations; they differ only in parameter estimation where (6) uses least squares, while (7) is maximum likelihood which optimally weights the observations.

We used 3 alternative sample lengths for n. In the first, we let n take the entire available set of past observations, but no less than 1,000. Thus, the forecasting experiment begins at June 22, 1966 for all 3 algorithms, leaving 4,902 test forecasts. The second and third alternatives assume that the parameters of the true specification change over time. Ideally, one would search for the best sample length. We chose, quite arbitrarily, 300 and 1,000, respectively, as "short" and "intermediate" alternatives to the "long" sample which varies from 1,000 to 5,901 observations.

The 4 specifications and 3 sample lengths produce 12 competing daily forecasts. To these we added 3 more daily forecasts. The 13-th is a simple average of all daily forecasts, and the 14-th and 15-th are the daily maximum and minimum forecasts.

A simple average of n equal-quality, conditionally independent forecasts, will rapidly converge to a perfect forecast (see Kane and Lee [1984]). Hence, failure of the average forecast indicates economically significant divergence in quality and mutual dependence of the 12 forecasts.

We added the maximum and the minimum of the daily forecasts to the set as a check for any forecast bias that affects profits significantly. If, for instance, a downward bias is indeed present and significant, then the maximum forecast will beat the minimum forecast, and any of the individual forecasts that are more severely biased. If some forecasts are sufficiently upward biased, then the minimum forecast will overcome the diversification effect of the average forecast and show a better cumulative profit. At the same time, both the minimum and maximum of the daily forecasts, are expected to show greater profit volatility than the average forecast. The two extreme forecasts will hereafter be referred to as MAXIMUM and MINIMUM, respectively.

II.2 The option pricing formula and hedge ratio

The accuracy of the proposed valuation technique may be compromised if an inferior option-pricing formula is used in the experiment. The robustness of the technique with respect to the accuracy of the formula cannot be a-priori assessed, and is left for future research. But there is reason to think that the rank order of the forecasting algorithm profitability will not be affected when the formula is not too far off the mark. We have opted to use the Black-Scholes formula on the assumption that for maturities of only one business day (hence, no longer than 4 calendar days), it will be sufficiently accurate.

Agents in the experiment trade one-day options on a $1 share of the NYSE portfolio. The exercise price of the options is taken to be $1 plus the risk-free rate. (Merton [1981] used this characterization and resultant option-pricing formula for his valuation of market timing ability.) The one-day excess rate of return on the NYSE portfolio, over the risk-free rate, is assumed to satisfy the required distributional assumptions. Under these conditions the Black-Scholes call (or put) option price reduces to:

$$P_t = 2N[.5\sigma_t] - 1 , \qquad (8)$$

where P_t is the call or put price, $N[\cdot]$ is the cumulative normal density, and σ the standard deviation of the daily rate of return. For small σ the price is linear in σ.

With this pricing formula, ignoring uncertainty about the variance, the hedge ratio for a riskless position involving the stock and call option (the number of shares per call), is given by: $H_t = -N[.5\sigma_t]$, (see also fn 4, and Smith [1976, p.21]).

II.3 The daily round of trade among hypothetical agents

Step 1: agents set their call (= put) price. Every day of the exercise, each agent applies his designated algorithm to the sample of past observations (with the appropriate length), and computes a forecast for the variance of the NYSE index of the next day. From this forecast, using (8), the agent determines his price for a one-day call/put on a $1 share of the NYSE portfolio.

The price of an option is increasing in the underlying asset variance. As a result, agents with low variance forecasts will believe that agents with high variance forecasts are overpricing call and put options, and vice versa[3].

Step 2: agents execute trades. The following trades are executed each day:

(i) every agent buys one call and one put from every agent that offers them at a lower price. The transaction is executed at the

---

3. In the experiments we had agents calculate option prices from (8), which assumes an exercise price of $(1 + risk-free rate). Yet in the settlement of the accounts from subsequent NYSE returns, we actually used $1 for the exercise price instead. This corner was cut in order to avoid the need to determine a source, and then observe daily risk-free rates of interest.

Reducing the exercise price of options (on $1 shares), by an amount equal to the daily risk-free rate, will slightly increase the expected transaction profits to buyers of calls and sellers of puts. In each transaction of call or put, the buyer is the agent with the higher variance forecast than the seller's. Thus, by lowering the settlement exercise price we have been slightly favoring upward-biased forecasts for call transactions, and downward-biased forecasts for put transactions.

As far as overall profits go, since every pair of agents always trade a straddle (one call and one put, as described in step 2), the inaccuracy in the settlement exercise price is exactly offset.

average of the bid (seller's) and ask (buyer's) price. For every trade, the agent uses his own variance forecast to determine the appropriate hedge ratio. Each trade, call or put, is then separately hedged by each agent, taking the appropriate position in the stock of the NYSE index.

(ii) The trades in (i) are repeated 2 more times with different transaction prices: instead of averaging the bid-ask spread, the first duplicate transactions are executed at the lower (seller's) price, and the second at the higher (buyer's) price.

Transacting at the lower price creates an asymmetry, where upward-biased forecasts are preferred to downward-biased forecasts. The reverse is true when transactions are executed at buyer's (higher) price. A separate account of these transactions serves two purposes: first, the difference in the relative profits of the duplicate transactions will help smoke out the algorithms that produce forecasts with a bias that is economically significant. Second, from these separate transactions we can compose a category that we call "trade at own price." Here, every agent transacts both "buy" and "sell" orders at his own price. This category is less sensitive to the price differentials between agents, and more sensitive to the rank order of forecasts, when compared with the "average bid-ask price" transactions. Each of these identical trades (at different prices) is also hedged (with the same ratio).

<u>Step 3</u>: settle end-of-day accounts, and accumulate profit/loss in individual agent accounts. At the end of each forecast day, the actual daily rate of return on the NYSE index is used to compute the profit/loss of each trade. For the purpose of future analysis, agent accounts are separated to subaccounts. There are 24 separate sub

accounts in all, to distinguish the following 4 categories: (1) transaction price of the trade: bid, ask, average; (2) days of positive and negative return on the NYSE index; (3) unhedged and hedged trades; and (4) trades in puts and calls. Finally, each agent sub account is further partitioned by customer/competitor agent. This partition can be used to tests any subgroup of competing forecasting agents.

For each agent, profits from all trades for the day, are totaled in each subaccount. Because each pair of agents trades one straddle, absolute profits depend on the number of agents (algorithms) in the experiment. In computing agent profits we wish to eliminate the effect of the volume of transactions that is proportional to the number of participating agents. Hence, the total daily profit in each subaccount is averaged over the number of trades/competitors, dividing it by k-1, where k is the number of participating agents.

III Relative Profitability of the NYSE Variance Forecasts, 6/1966-12/1985

We report results of the experiment along the categories described in section II.3 Step 3, beginning with some forecast statistics and the rank order of profitability for the entire experiment.

III.1 Summary statistics of the variance forecasts

Over the experiment period, June 22, 1966 to December 31, 1985, the annual rate of return, compounded daily, on the NYSE portfolio has been 9.78%, with a standard deviation of 12.86%. The ratio of daily standard deviation to average return was 20.8.

The annualized average daily forecasts of the standard deviation and the standard deviation of these forecasts, by algorithm, are presented in Table 1(a). The first column shows that all algorithms' average forecasts of standard deviation were lower than the in-sample standard deviation of the NYSE return.

The second column of Table 1(a) shows that the variance of the forecasts was very different across specifications. For example, the standard deviation of forecasts from the OLS (1,000-5,000) is only 1.24%, indicating a smooth measure of variance, while that of the ARMA (1,000-5,000) is 7.92%, revealing a widely varying forecast.

The standard deviation of forecasts also varies quite significantly across sample lengths within each specification, and is not always smallest for the longest sample length. Within the 3 ARCH specifications, the shortest sample (300) has the smallest standard deviation of forecasts. Within the ARMA specifications, the intermediate sample length (1,000) results in the least variable forecast.

The right hand column in Table 1(a) displays the betas of the forecasts of standard deviation. They are estimated from regressions of the daily forecasts of standard deviation on the subsequent NYSE returns. The betas are practically zero, indicating that changes in the forecasts of standard deviation are non systematic.

As a proxy for the accuracy of the variance forecasts, a forecast error is often defined as the difference between the squared return and the variance forecast. (With daily observations, taking the squared deviation from a daily mean would make no difference.)

The first column in Table 1(b) shows the mean forecast error of the variance, as a percent of the sample variance for the entire period. Here, too, the variance forecasts appear to have been too low, on average. The magnitude does not tell us much about the Jensen inequality, however. In 5 out of the 12 algorithms, the percent difference between the in-sample standard deviation and the average forecast of the standard deviation, from the first column of Table 1(a), is smaller than the percent forecast error in the first column of Table 1(b). This measure again describes the level of the forecast on average, without indicating whether they vary too much or too little.

The second column of Table 1(b) will be small when large, subsequent, squared returns are associated with large variance forecasts, indicating the variability of the forecast. It is not a direct measure of forecast accuracy, however, since the squared returns are not the true variance, and since the implied loss function is linear. Thus, for example, there is no penalty for negative variance forecasts.

The last column in Table 1(b) shows the beta for the forecast error. The magnitude of the positive betas suggests that large forecasting errors (as defined here) are more likely in "up markets," days when the NYSE return is positive.

The statistics in Tables 1(a) and 1(b) do not suggest a clear choice of specification. On the basis of mean square error, MA(300) and all the three ARMA specifications look best. But the variance of forecast error (as proxied here) is smallest for the ARCH specifications. The standard deviation of the ARCH forecasts themselves, in Table 1(a), is about average.

Table 1(c) introduces the simulation of option valuation that we use to transform differences in variance forecasts to differences in pricing options on the NYSE index portfolio. First, the daily variance forecast of each algorithm is transformed to a one-day call price (equal to put price). The first column in Table 1(c) shows the average call price in cents per one-day call on a $1 share on the NYSE, with an exercise price of $1 plus the daily risk free rate. Next, we pretend that every day, each forecaster buys one call and one put at his own price, based on his variance forecast. Holding this straddle with an exercise price of $1 (the risk-free rate is actually neglected here), the end-of-day payoff to each forecaster is identical, and equal to the absolute rate of return on a 1$ share of the NYSE over the forecast day. The day t profit to forecaster i equals: $|r_t|-2P_{it}$, where r_t is the daily NYSE return and P_{it} is the (i^{th} forecast-driven) call price for day t, because put and call prices are equal. This profit varies across forecasters as their forecasts differ.

Note that the objective here is to identify the best option-pricing forecast algorithm. Hence, we are looking for the agent with the smallest cumulative absolute profit from holding the straddle, and with the smallest variance of profit.

The first column of Table 1(c) shows the average daily price that agents would have assigned to a call (put) on a $1 share of the NYSE, with an exercise price of $1 plus the daily risk-free rate, over the entire forecast period. It is of an order of .28 cents as reflected by the average price of the AVERAGE forecast.

The second column in Table 1(c) shows the average daily profit from holding one straddle a day, over the sample forecast period. The last column shows the standard deviation of the daily profit. The bottom panel in Table 1(c) introduces the additional forecasters, the MAXIMUM, MINIMUM and AVERAGE forecasters.

Using option prices in this way, too, does not indicate a winning forecast algorithm. The smallest absolute profit goes to the OLS(300) and OLS(1,000) algorithms. The average loss of the first, amounts to .12% of the average price of a straddle per day. For comparison, the average profit from the most accurate (by this standard) ARCH specification, using a sample length of 1,000, is 1.60%. Yet the standard deviation of profits of OLS agents is not the lowest. The difference between the standard deviation of OLS(1,000) and ARCH(1,000) is 6.5% of the average straddle price. Hence, we have to conclude that at this level of aggregation, the data is insufficient to tell the entire story of how each of these contingent-claim pricing agents would fare if they had to compete with each other in the contingent-claims market.

The AVERAGE forecast fails to dominate in either average profit or standard deviation, implying that the forecasts are significantly different in quality and not conditionally independent. MAXIMUM is a better forecast than MINIMUM as a result of the prevailing sample downward bias of the variance forecasts and perhaps, to some extent, due to the Jensen Inequality.

It is interesting that while all agents' variance forecasts are (on average) too low, the average profit from holding the straddle is negative for 4 out of 12 forecast algorithms. This lack of effect of the forecast downward bias has occurred despite the Jensen Inequality that is expected to exaggerate the degree of underpricing due to concavity of the option formula. Further, since the straddle makes for a risky position, we would expect it to earn a positive risk premium, and hence such effect would also go the other way[4].

The probable explanation lies in the properties of the sample, namely, in the frequency of "up" and "down" markets. Recall that we use $1 as the exercise price for a $1 share of the NYSE, while using a pricing formula that assumes the exercise price to be $1 plus the risk-free rate. This would not affect the price of a straddle in any significant way, because the derivatives of the call and put prices with respect to the exercise price differ only by the risk-free rate to maturity and offset each other. However, as far as realizations go, if the frequency of negative realized returns on the stock is greater than expected, then the deficiency of returns on the put will exceed the extra returns on the call, accumulating a bias toward smaller profits. Indeed, the average excess return on the NYSE for the experiment period (less than 5%) has been below the historical average (of more than 8%), giving credence to this explanation.

III.2 Rank order by overall profit

The overall result of the experiment is given in Table 2. It summarizes agent profits from all hedged trades over the entire period, from transactions at the average of the bid and ask prices.

Turning first to the question of whether the algorithms are of similar quality with independent pricing errors we note from Table 2

4. The hedge ratio of a put, $1-N(d_1)$, is here slightly smaller in absolute value than the call's: $-N(d_1)$. With a one-period maturity, $d_1=[\log(S/X)+r_f+.5\sigma^2]/\sigma$, is the well known term from the Black-Scholes formula, where S and X are the current stock and exercise price, respectively. With an exercise price of $S(1+r_f)$, d_1 reduces to $.5\sigma$. So here, the hedge ratio for a straddle, the sum of the hedge ratios of the put and call, is slightly negative. Hence, the unhedged straddle has a positive beta and should earn a positive risk premium.

that the AVERAGE forecast is inferior to ARCH(1,000), and only marginally superior to the other two ARCH specifications. This indicates that the 3 ARCH forecasts are distinctly better than the remainder of the 12. As for conditional independence of the forecasts, note that the standard deviation of the average daily profit of the AVERAGE forecast is by far the smallest. This indicates that the pricing errors of the competing agents offset one another for the large part, so that the average pricing error is small. For this to happen, the pricing errors of the agents have to be quite independent.

The losses of the MAXIMUM forecast are significantly worse than the MINIMUM. This despite the indication from Tables 1(a) and 1(b) that all forecasts are biased downward, on average. One would expect that a reason may be that the MINIMUM option price (and forecast) is non negative, while MAXIMUM forecasts and option prices are unbounded, hence inherently more variable. However, the standard deviation of the daily profits of the MAXIMUM and MINIMUM forecasts are similar (and by far the largest, as expected).

Overall, the results indicate that there are significant incremental profits from switching among variance-forecast algorithms in a competitive asset market. These profit (which appear to be non systematic) are mostly driven by the variance of the variance forecasting-errors, a property that is difficult to asses with standard statistics. This should be stated, however, in the context of the small, arbitrary, set of specifications that are part of the experiment. (See Pagan and Hong [1988], Pagan and Ullah [1988], and Pagan and Schwert [1989] for discussion of the quality of variance estimators.)

More specific indications of the quality of forecasts from Table 2 are:

(i) The ARMA specification appears as the experiment's worst performer. Poterba-Summers [1986] made a powerful argument that variance shocks should matter little to stock prices. Using an ARMA specification, they show that shocks to the variance of stocks do not persist and hence, should have little effect on value. Our experiment suggests that in the context of asset prices vis-a-vis changes in variance, persistence needs to be examined with better forecast algorithms.

(ii) The results suggest that some restriction on the length of forecasting sample may be profitable. A straight forward application of the technique proposed here will allow identification of a preferred sample length for the appropriate specification. The question of whether business cycles or shifts in monetary policies affect the parameters (particularly in terms of persistence) of various return process specifications can be addressed using the technique proposed here.

(iii) The dominance of the ARCH specification is quite striking. It justifies the new interest in testing asset pricing models with ARCH estimates in order to account for rational forecasts of market variance and asset covariances.

The significance of potential profits from better forecasting can be judged from the annualized average profit of the agents. From table 1(c), the average investment in one straddle (one call and one put on $1 share of the NYSE) was in the order of .6 cents per day. Table 2 shows the annualized profit from an average daily trade from such average investment (actually less, because the profit is stated per trade and agents could be taking offsetting positions). For the best 3 forecasters in Table 2, the total annual profit (from 250 daily sessions per year) was in the order of 10 cents, with an annualized daily standard deviation of about 2 cents. Such large profits are obviously due in part to the even larger losses of the big losers. Still, bear in mind that none of the algorithms in the experiment has been condemned in the literature as inadequate, and empirical studies have made extensive use of these and similar ones. We shall show results from pairwise trades below.

III.3 Trade at your own price and ranking within option categories

Table 3 reports agent relative performance in a different way. It reports profitability in the 4 separate option-trade categories: sell/buy; call/put. The table reports profits from hedged trades at their own prices: the "sell" transactions are executed at the low (seller's) price. The "buy" transactions are executed at the high (buyer's) price. Thus, each transaction between two agents is reported at a different price (with a different profit) in each trader account.

If the net position of an agent were "long one straddle" (at own price) each day, then his average profit would be identical to that reported in Table 1(c). The difference in profits as reported in Table 3 results from the frequency of long and short trades with the various agents. It is not affected by the difference in prices across agents, and hence profits cannot be driven by the magnitude of the forecast differences across agents. Rather, profits will be driven by the frequency and timing of above and below-average forecasts. This measure of profitability is very sensitive to agent bias vis-a-vis other agents.

Average profits from all transactions at own price, as reported in Table 3, must be worse than those at average of bid-ask price in Table 2. When transacting at "own price," as compared with "average of bid-ask prices," every transactor loses half of the bid-ask spread.

Another reason why, at own prices, agents profits are less than those reported in Table 1(c) is that now, each agent gets to execute more net transactions on days when his forecast is more extreme relative to other agents. At "own price," same-day transaction prices are same for both, "buy" or "sell," transactions. Hence, long positions exactly offset short ones. With an extreme forecast, the agent will be on the same side of most or all transactions. If more often than not, extreme forecasts are biased even for the better forecasters, then profits must be smaller in Table 3 from those in Table 1(a).

The relative performance at own price, as reported in Table 3, is quite similar to profit from trading at bid-ask prices, as reported in Table 2. Moving from average bid-ask to own price, we find a switch in ranks at the top, where the AVERAGE forecast now ranks 1, and displaces ARCH(1,000) that ranks 2. The ranking from 3 down to 8 remains unchanged, and there is only little change farther down the line. This implies that the rank order of profitability is not mostly driven by the magnitude of the forecasting error of the big losers, but also by the number of positions "on the correct side" that agents take. Further evidence of that will be given in smaller groups and pairwise comparisons.

The individual transaction categories in Table 3 reveal that the profit differentials between better and worse performers are larger for "buy" than for "sell" transactions, but not so for "put" versus "call" transactions. This suggests that the relative quality of forecasts across competing algorithms does not differ in up market from down market days[5]. At the same time, it appears that the algorithms are more distinct in quality when they forecast higher variances, compared with days when they forecast lower ones.

III.4 The effect of hedging variance-driven put-option trades

In general, the overall portfolio positions of information traders will be affected by their variance forecasts in two ways: the desired (short vs. long) position, and the hedge ratio of the position.

Agents with more accurate variance forecast, will also be better hedged. It is therefore interesting to see how the hedging activity affects the profitability of trading on variance forecasts.

Table 4 isolates the effect of the hedging activity with put trades (the effect of hedging straddles is negligible). The results are quite striking. For unhedged trades, the AVERAGE forecast is a clear winner. Its average profit is 15% higher than the next (ARCH1000) competitor. Otherwise, profitability ranking is similar to that in the other tables. When the hedge activity is incremented into the agent accounts, all 3 of the ARCH specifications fare better than the AVERAGE (ARCH1000 by 33%). In general, the better forecasts gained in profitability relative to the

5. We also split the sample by the sign of the NYSE realized daily return and recalculated relative profits. There was no noticeable difference between "up" and "down" days.

inferior ones, pushing down the relative profitability of the AVERAGE forecast.

Note that the hedging activity, unlike trading in options, is not a zero sum game. An agent with a variance forecast that is too high will buy puts at a high price. In addition, he will take a larger than called for, long position in the stock. The entire portfolio will amount to a net long position in the stock which, for its part, has a positive expected rate of return[6]. Results were similar for trades in call options.

While hedging tended to exaggerate the relative profits from put trades, it also served its original purpose of reducing risk. The standard deviation of all profits/losses decreased significantly, but more so for the successful forecasts with better hedge ratios.

III.5 Transitivity in the relative-profit relation

There is no reason to expect that the relative profit relation across forecasts is transitive. This issue is investigated in Table 5.

The first 6 columns of Table 5 show the ranking of agents within the forecasting group, as the least profitable are dropped out. The first column in Table 5 shows the ranks of all 15 forecasts. When the bottom three forecasts are dropped, there are some changes in the ranks. Along the way, dropping the 3 worst forecasters at a time, slight changes take place, in the form of switches between forecasts that perform similarly. The expected change occurs in the rank of the AVERAGE, which

6. The portfolio will still have a less than adequate risk premium, perhaps even a negative expected excess return because of the overpricing of the put option.

deteriorates as the group becomes smaller and inferior forecasts dropped..

The last two columns in Table 5 show how the best forecast in the tested group (ARCH 1,000) fares in pairwise trades with all other forecasts, both when traded at the high and low price. These comparisons show that ARCH 1,000 is preferred to all other forecasts in the group. It is apparent that this forecast is downward bias, though. When trading at the low price, ARCH 1,000 loses to MAXIMUM which is always on the buy side. At the same time, when trading at the high price, ARCH1000 still wins against MINIMUM which is always on the sell side.

The magnitude of the gains of ARCH 1,000 against all other forecasts is striking. Even against the next best specification on a pairwise basis, OLS 300, the gain at the average of bid-ask prices would be 6.31 cents per year on an average investment of less than .6 of one cent.

Summary

The paper proposes a technique to compare variance forecast algorithms using an economic value criterion. It takes advantage of the role of variance in pricing contingent claims. This is accomplished by setting up a hypothetical option market, where each agent represents a competing variance forecast-algorithm.

The technique has been demonstrated with one-day options on the NYSE portfolio over the period 1966-1985. We show that the economic value to improved accuracy of variance forecasts, from switching across some widely used forecast algorithms, is large by any standard.

For the limited set of specifications tested here, it has been demonstrated that ARCH specifications appear the most suitable, and that some restriction on the length of the rolling sample is preferred. We have shown that variance-forecast accuracy bears significant economic profits.

Examination of the relative forecast profits with the proposed technique reveals the economic significance of various statistical properties of the forecast errors. In particular, it appears that differences in the variance of the forecast errors may be generating most of the large economic losses from using inadequate variance forecasts.

In subsequent research we intend to apply this technique to longer-term options. There, the issue of the appropriate option-pricing formula will have to be reevaluated. Testing of the technique will include the introduction of new assets, such as currency and bond options. In addition, the importance of adjustment for the Jensen Inequality problems will be examined.

REFERENCES

Bodie, Zvi (1988). "Inflation, Index-Linked Bonds and Asset Allocation," NBER WP #2793, December 1988.

Bollerslev, Tim, Robert Engle and James Wooldridge (1988). "A Capital Asset Pricing Model with Time Varying Covariances," Journal of Political Economy, 96, 116-131.

Bollerslev, Tim, Ray Chou, Narayaman Jayaraman and Kenneth Kroner (1990), "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence, with Suggestions for Future Research," Working paper, March 1990, presented at the UCSD conference on Modeling Volatility in Financial Markets, April 6-7 1990.

Chou, Ray (1988). "Volatility Persistence and Stock Valuation" Journal of Applied Econometrics 3, November-December, pp.279-294.

Cox, John and Mark Rubinstein (1985). "Option Markets," Prentice Hall, Englewood Cliffs, New Jersey.

Engle, Robert (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," Econometrica,50,987-1008.

French, Kenneth, William Schwert and Robert Stambaugh (1987), "Expected Stock Returns and Volatility," Journal of Financial Economics, September 1987, pp. 3-29.

Hong, Ted (1987). "The Integrated Generalized Conditional Heteroskedastic Model: The Process, Estimation and Monte Carlo Experiments," University of California, San Diego Discussion Paper #87-32.

Hull, John and Alan White (1987). "The Pricing of Options on Assets with Stochastic Volatility," Journal of Finance, 42(June 1987), 281-300.

Merton, Robert (1981). "On Market Timing and Investment Performance I: An Equilibrium Theory of Value for Market Forecasts," Journal of Business, 54, 302-406.

Pagan, Adrian and Y. Hong (1988). "Non Parametric Estimation and the Risk Premium," University of Rochester mimeo.

Pagan, Adrian and Aman Ullah (1988). "The Econometric Analysis of Models with Risk Items," Journal of Applied Econometrics3, 87-105.

Pagan, Adrian and William Schwert (1989). "Alternative Models for Conditional Stock Volatility," University of Rochester Mimeo

Poterba, James and Lawrence Summers (1986). "The Persistence of Volatility and Stock Market Fluctuations," The American Economic Review, 76, 1142-1141.

Schmalensee, Richard and Robert Trippi (1978). "Common Stock Volatility Expectations Implied by Option Premia," Journal of Finance 33, No 1, 129-147

Smith, Clifford (1976), "Option Pricing," Journal of Financial Economics 3, 3-51.

Wiggins, James (1987), "The Pricing of Options with Stochastic Volatility " Journal of Financial Economics 19, 351-372.

Table 1(a)

Summary Statistics of the Daily Forecasts of Standard Deviation (Annualized Percents)

Algorithm		Daily Standard Deviation		
Specification	Sample Length (n)	Average	SD	Beta
ARCH	300	11.72	4.09	.01
ARCH	1,000	11.77	4.11	.01
ARCH	1,000-5,000	11.76	4.59	.01
MA	300	12.41	3.17	.01
MA	1,000	12.36	2.28	.00
MA	1,000-5,000	10.93	1.32	.00
OLS	300	12.06	3.11	.01
OLS	1,000	11.97	2.50	.00
OLS	1,000-5,000	10.59	1.24	.00
ARMA	300	9.55	8.58	.01
ARMA	1,000	10.67	6.70	.01
ARMA	1,000-5,000	9.95	7.92	.00

Annualized Sample SD of the NYSE daily rate of return = 12.86

Table 1(b)

Summary Statistics of the Daily Variance Forecast Errors

Specification	Sample Length (n)	MSE/AV[(a)]	SD(SE)/AV[(b)]	Beta[(c)]
ARCH	300	.0696	1.9640	.22
ARCH	1,000	.0626	1.9531	.22
ARCH	1,000-5,000	.0420	1.9631	.22
MA	300	.0097	2.0492	.24
MA	1,000	.0466	2.0934	.24
MA	1,000-5,000	.2682	2.0758	.25
OLS	300	.0642	2.0469	.24
OLS	1,000	.1045	2.0921	.24
OLS	1,000-5,000	.3136	2.0760	.25
ARMA	300	.0048	2.6773	.18
ARMA	1,000	.0420	2.3929	.20
ARMA	1,000-5,000	.0235	2.6001	.19

(a) The ratio of squared realized return minus the variance forecast divided by the sample variance

(b) Standard deviation of the daily forecast error divided by the sample standard deviation.

(c) The beta of the variance forecasting error, i.e., the regression coefficient of the variance forecast error (sqared realized return less the variance forecast) on the realized market-index return.

Table 1(c)

Summary Statistics of the Simulated Option Prices and Average Realized Profits[a] from Holding Straddles[b]

Specification and Sample Length		Average Option Price (cents)	Average Profit from Straddle	SD of Profit
ARCH	300	.2958	.0118	.5152
ARCH	1,000	.2969	.0095	.5124
ARCH	1,000-5,000	.2960	.0114	.5146
MA	300	.3132	-.0229	.5369
MA	1,000	.3118	-.0202	.5514
MA	1,000-5,000	.2759	.0517	.5456
OLS	300	.3043	-.0051	.5368
OLS	1,000	.3021	-.0007	.5512
OLS	1,000-5,000	.2673	.0688	.5457
ARMA	300	.2411	.1212	.6374
ARMA	1,000	.2693	.0647	.5860
ARMA	1,000-5,000	.2513	.1008	.6201
MAXIMUM		.4045	-.2056	.5515
MINIMUM		.1623	.2789	5554
AVERAGE		.2854	.0326	.5250

(a) The calls and put are on $1 share of the NYSE. Prices and profits are reported in cents.
(b) The profit from holding a straddle on a $1 share with and exercise price of $1 is equal to the absolute rate of return less the price of the straddle that is here equal to twice the value of the call.

Table 2

Annualized Daily Profits[a] From All Trades[b], Over the Entire Forecasting Period, By Algorithm

Specification		Average	SD	Beta	Rank
ARCH	300	9.67	2.01	.014	3
ARCH	1,000	10.75	2.04	.015	1
ARCH	1,000-5,000	9.52	2.46	.012	4
MA	300	- .61	2.50	.005	9
MA	1,000	- 3.96	2.43	.009	11
MA	1,000-5,000	4.61	2.03	-.009	6
OLS	300	3.83	2.10	.000	7
OLS	1,000	.70	2.17	.003	8
OLS	1,000-5,000	6.14	2.46	-.014	5
ARMA	300	-14.16	3.50	-.012	14
ARMA	1,000	- 2.25	2.60	-.003	10
ARMA	1,000-5,000	- 5.68	2.92	-.017	13
MAXIMUM		-24.81	3.95	.025	15
MINIMUM		- 4.15	3.95	-.029	12
AVERAGE		10.39	1.18	-.001	2

(a) The profits are per-competitor, i.e, the daily profit from all transactions is divided by k-1 (where k is the number of agents/algorithms).

(b) Trades are at the average of the bid-ask prices, and hedged.

Table 3

Annualized Daily Profits[a] From Trades at "Own Price," Over the Entire Forecasting Period, By Type of Transaction and Algorithm

Specification and Length (n)		Sell[b] Call	Sell[b] Put	Buy[c] Call	Buy[c] Put	Total	Rank
ARCH	300	- .87	- 1.90	-10.29	- 6.88	-19.94	3
ARCH	1000	.07	- 1.26	- 9.65	- 6.89	-17.73	2
ARCH	1000-5000	- 1.10	- 2.34	-10.52	- 7.60	-21.56	4
MA	300	- 1.72	- 2.46	-18.11	-14.77	-37.06	8
MA	1000	- 3.82	- 4.99	-19.24	-16.77	-44.82	10
MA	1000-5000	- 6.15	- 7.57	- 9.04	- 7.21	-29.97	6
OLS	300	- 1.93	- 3.15	-14.46	-13.48	-33.02	7
OLS	1000	- 3.93	- 5.47	-15.47	-13.48	-38.35	9
OLS	1000-5000	- 7.87	- 9.52	- 6.92	- 5.53	-29.84	5
ARMA	300	-23.25	-24.06	-18.27	-15.87	-81.45	14
ARMA	1000	-11.54	-12.97	-13.30	-11.28	-49.09	11
ARMA	1000-5000	-17.32	-18.96	-13.73	-12.41	-62.42	12
MAXIMUM		.0	.0	-47.18	-41.37	-88.55	15
MINIMUM		-31.73	-33.01	.0	.0	-64.74	13
AVERAGE		- 1.67	- 3.41	- 7.02	- 5.16	-17.26	1

(a) The profits are per-competitor, i.e, the total daily profit is divided by k-1 (where k is the number of agents/algorithms).
(b) Trades are at the agent's (seller) lower price, and hedged.
(c) Trades are of type at the agent's (buyer) higher price, and hedged.

Table 4

The Effect of Hedging Variance-Driven Put-Option Trades[a]

Spec. And Sample Length (n)		Hedged Transactions Profit (1)	SD	Beta	Unhedged Tranactions Profit (2)	SD	Beta	(1)-(2)
ARCH	300	10.76	2.66	.07	6.42	3.51	-.06	4.34
ARCH	1,000	11.34	2.76	.08	9.05	3.58	-.08	2.29
ARCH	1,000-5,000	10.23	3.37	.09	8.03	4.30	-.09	2.20
MA	300	.62	3.08	.08	1.59	4.21	-.16	- .97
MA	1,000	- 3.44	2.93	.05	- .66	3.90	-.07	-2.78
MA	1,000-5,000	4.66	2.36	-.04	2.70	3.56	.08	1.96
OLS	300	4.49	2.53	.03	3.34	3.58	-.05	1.15
OLS	1,000	.77	2.59	-.00	1.29	3.60	.03	- .52
OLS	1,000-5,000	5.83	2.88	-.08	1.96	4.34	.18	3.87
ARMA	300	-13.45	4.47	.05	-13.06	5.97	.10	- .39
ARMA	1,000	- 2.11	3.40	.04	- 1.82	4.39	.04	- .29
ARMA	1,000-5,000	- 6.02	3.68	.03	- 3.89	4.84	.07	-2.13
MAXIMUM		-21.91	5.74	.30	-12.42	6.84	-.44	-9.49
MINIMIM		- 4.93	4.31	-.13	-11.37	6.83	.44	6.44
AVERAGE		10.26	1.42	.00	8.83	1.92	.01	1.43

(a) Trades are at bid-ask spread. Profits are in cents per competing agent per year from daily trade of one put on $1 of the NYSE portfolio.

Table 5

Transitivity in Relative Profits Across Forecasts

Specification and Sample Size	Rank in Groups[b] by Size: 15	12	9	6	3	2	Av. Annualized Daily[a] Profit of ARCH 1,000 in Pairwise Trades With: Low Price[c]	High Price[d]
ARCH 300	3	2	2	2	3		3.11	3.07
ARCH 1000	1	1	1	1	1	1	0	0
ARCH 5000	4	3	3	3	2	2	1.98	1.74
MA 300	9	6	5	5			4.61	8.66
MA 1000	11	10					8.44	12.17
MA 5000	6	7	7				10.58	5.31
OLS 300	7	5	4	4			5.39	7.23
OLS 1000	8	9	8				9.57	10.86
OLS 5000	5	8	9				12.22	4.80
ARMA 300	14						27.53	13.57
ARMA 1000	10	11					17.51	10.60
ARMA 5000	13	12					9.72	12.36
MAXIMUM	15						-1.16	25.73
MINIMUM	12						34.74	1.07
AVERAGE	2	4	6	6			9.72	6.33

(a) Profits are stated in cents
(b) Transactions are at average of bid-ask prices
(c) Transactions are executed at the seller's price
(d) Transactions are executed at the buyer's price