THE JOURNAL OF FINANCE • VOL. LIV, NO. 4 • AUGUST 1999

# Market Risk and Model Risk for a Financial Institution Writing Options

T. CLIFTON GREEN and STEPHEN FIGLEWSKI*

**ABSTRACT**

Derivatives valuation and risk management involve heavy use of quantitative models. To develop a quantitative assessment of model risk as it affects the basic option writing strategy that might be followed by a financial institution, we conduct an empirical simulation, with and without hedging, using data from 1976 to 1996. Results indicate that imperfect models and inaccurate volatility forecasts create sizable risk exposure for option writers. We consider to what extent the damage due to model risk can be limited by pricing options using a higher volatility than the best estimate from historical data.

WITH THE REMARKABLE GROWTH OF TRADING in derivative instruments in recent years, derivatives activities of banks and other financial institutions are accelerating. The volume of outstanding contracts has expanded at a rapid pace, and the contracts themselves are more complex, span longer maturities, and cover a broader range of underlying assets.

Much of the growth is in "plain vanilla" instruments like forwards, for which the principles of valuation are well established. However, assessing and managing risk, even for simple products, can entail considerable uncertainty. Derivatives with option features present significantly greater challenges with regard to correct valuation and the design and implementation of risk management strategies. Moreover, options involve the asymmetry between buying and writing, in that the option buyer has liability limited to the amount invested but the option writer is exposed to the risk of losses that can greatly exceed the initial premium received. Not surprisingly, the public prefers to buy options rather than to write them. Since each contract requires a buyer and a writer, if the public wants to be long options, the dealer community must be short options overall. This means that the typical financial institution entering the derivatives business in order to satisfy its customers' demand for options, and (ideally) to earn profits by making markets, is primarily writing contracts. In so doing, it is exposed to a variety of risks.

Derivatives risks have been widely discussed. Often they are classified into several categories that may differ from one discussion to another,[1] but a common taxonomy is the following:

* The authors are from New York University and Emory University, respectively.

[1] See, for example, the Group of Thirty (1993) or U.S. Government Accounting Office (1994).

- *Market risk:* the risk that movements in financial market prices impair a firm's financial condition due to its positions in derivatives.
- *Credit risk:* the risk (broadly defined) that the counterparty to a derivatives contract fails to fulfill its contracted obligations.
- *Operational risk:* the risk of derivatives-related losses from deficient internal controls or information systems.
- *Legal risk:* the risk that derivatives contracts are not legally enforceable.

Recommended procedures for managing these different risks safely have been broadly promulgated.[2] The growth of derivatives activity has also increased awareness of risk exposures generally and has contributed to the development of more formal methods of risk assessment, such as value-at-risk (VaR). Indeed, the Bank for International Settlements (BIS) has suggested regulatory policies for setting capital requirements for banks that are closely related to the VaR methodology, and the system has been adopted by the European Community (EC) and by banking authorities elsewhere around the world.[3]

One important feature of both the growing derivatives trade and the new approaches to risk management is that the derivatives business today depends very heavily on theoretical models for pricing contracts, and for risk assessment and hedging. This introduces an important new type of risk that has not played much of a role in investing before: model risk. Since the derivatives trade is based so much on pricing models, model errors create risk—misvalued contracts may be sold for less than they are actually worth or may be purchased at overvalued prices, incorrectly estimated risk exposures may be greater than anticipated, and hedging strategies may be less effective than they are supposed to be.

Figlewski (1998) discusses several sources of model risk facing derivatives traders. First, there is the risk that a given model may be misspecified. A common problem is that to derive a valuation model, it is necessary to assume a stochastic process for the derivative's underlying asset. The original Black–Scholes (1973) option pricing model assumes that the underlying follows a lognormal diffusion process. This process has a number of virtues, including the fact that (instantaneous) rates of return have a normal (Gaussian) distribution, for which the mathematics are well known, and the time independence of the stochastic component is consistent with the principle of informationally efficient markets. Many subsequent derivatives models have generalized the returns process but continue to assume that the stochastic component remains locally Gaussian. However, empirical investigation almost invariably finds that actual returns are too "fat-tailed" to be lognormal. There are more realizations in the extreme tails (and the extreme values themselves are more extreme) than a lognormal distribution allows for. In

[2] See the Risk Standards Working Group (1996) or the Basle Committee on Banking Supervision (1994).

[3] The Basle Committee on Bank Supervision (1996).

other words, the standard valuation models are based on assumptions about the returns process that are not empirically supported for actual financial markets. Other difficulties with the distributional assumptions, notably parameter instability, create additional errors in the standard models.

A second important source of risk in using a valuation model is that not all of the input parameters are observable. In particular, even if one has a correctly specified model, using it requires knowledge of the volatility of the underlying asset over the entire lifetime of the contract. This creates a formidable forecasting problem, for which neither the "best" estimation procedure nor the model risk characteristics of the resulting theoretical option values are known.

Third, one of the important features of the option pricing paradigm is that the theoretically correct risk management strategy emerges naturally, because the model is derived from an arbitrage trade. Along with the fair option value, a pricing model provides the option's delta, which indicates how to hedge the option's market risk exposure by taking an appropriate offsetting position in the underlying asset. Delta hedging is the mainstay of risk management for options marketmakers. Proper hedging requires that the pricing model be correct, of course, and also that the right volatility input be used. Even so, the delta hedging strategy on which the pricing model is based involves continuous trading to maintain "delta neutrality" at every instant as the underlying asset's price fluctuates. In practice, this is not feasible for a diffusion process. Such a strategy theoretically entails an infinite number of transactions in the underlying asset over an option's life, and therefore infinite transactions costs. Moreover, financial markets do not remain open continuously, meaning that rebalancing cannot always be done when it is needed. In practice, delta hedging is done approximately, with frequent periods between rebalancing trades during which the hedge is somewhat inaccurate. The result is that a delta hedge in practice does not fully eliminate price risk.

These sources of potential model risk are easy to see, but their quantitative impact is not known. For example, we can observe that actual stock returns are not exactly lognormal, but without knowing what the true distribution is, it is difficult to judge how much inaccuracy is introduced by assuming lognormality in building a valuation model.

The object of this paper is to develop a quantitative assessment of the extent to which the different sources of model risk can be expected to affect a basic option writing strategy that might be followed by a bank or another financial institution. We conduct an empirical simulation, with and without hedging, using historical data from several important markets. Specifically, we wish to explore the following problem: If a bank or similar financial institution regularly writes standard European calls and puts and prices them using the appropriate variant of the Black–Scholes model with a volatility forecast computed optimally from historical data, what are the risk and return characteristics of the trade? More generally, what is the market and model risk exposure faced by a bank that performs this transaction repeatedly over time?

The use of historical simulation to examine the performance of option trading strategies has had a long history, beginning with early papers by Galai (1977) and Merton, Scholes, and Gladstein (1978, 1982). Unlike Monte Carlo simulation in which the investigator must posit a form for the returns distribution that may be incorrect, simulation with historical data inherently employs the true distribution of returns as it would have impacted the strategy of an institutional writer of options.

In the next section, we describe the markets and the simulation strategies to be examined in detail. Section II discusses the various possibilities for obtaining volatility forecasts from historical data. One aspect of the research design that we emphasize throughout is that the volatility estimate should make optimal use of historical data, but only of those data that were available at the time the forecast would have been made. We wish to examine trading strategies as they would have been implemented.

Section III presents the simulation results for the strategies of writing options at model prices each period and carrying the positions to expiration, either without hedging or following a delta neutral hedging policy with regular rebalancing of the hedged position. To see how important volatility forecast errors are to the overall result, we also compute returns that would have been obtained had the actual volatility that was realized over the option's remaining lifetime been known at each point. In Section IV we examine the impact of variations in the strategy, including using suboptimal volatility forecasts and writing options that are either in the money or deep out of the money. Our results indicate that the various sources of model error combine to produce substantial risk exposure for the option writer, even when delta hedging of the position is actively followed. In Section V we consider to what extent the bank can limit the damage due to model risk by pricing options using a higher volatility than its best estimate from historical data. Section VI concludes.

## I. Design of the Simulation

We consider a bank that is in the business of writing European calls and puts either every day, for short maturity contracts, or every month, for two- and five-year maturities. The options are priced at model values using standard valuation formulas of the Black–Scholes family. The required volatility input is computed from historical returns data from the market in question, using either the classical estimator for unconditional volatility with a fixed length sample of past returns or an estimator with exponentially declining weights and all available past returns. The exact forms of the estimators are described in the next section. An important point, however, is that the volatility inputs to the models are based only on data that would have been available at the time the calculations were needed, but they attempt to use that data optimally. In strategies involving hedging, the positions are rebalanced at each point using deltas computed from up-to-date volatility estimates for those dates, as would be normal for a bank that manages its risk properly using all of the information available to it.

In the basic simulation, we consider calls and puts with five different maturities (one month, three months, one year, two years, and five years) and two degrees of "moneyness" (at the money, with the strike set equal to the initial asset price, and out of the money, with the strike set 0.4 standard deviations away from the asset price). Basing the strike price for an out of the money option on the volatility estimate results in strike prices that are further from the current asset price for longer maturity contracts, but with a constant probability of ending up in the money. Under lognormality, the probability that an observation will fall more than 0.4 standard deviations below (or above) the mean is 0.34; in other words, the estimated probability that our out of the money options would finish in the money is slightly greater than one-third. Trading volume for exchange-traded options tends to be greatest for options close to this range of strikes.

The option maturities we analyze span the range over which most options are written. For shorter term contracts of a year or less, we use daily data in the simulations. This includes computing historical volatilities from daily data, writing new contracts every day, and rebalancing hedged positions daily. For two- and five-year maturities, we use monthly data and trading intervals.

We examine four major underlying assets from the most important asset classes on which financial options are actively traded: a stock index (the Standard and Poor's 500), short-term interest rates (three-month LIBOR), long-term interest rates (the yield on 10-year U.S. Treasury bonds), and foreign currencies (the deutsche mark(DM)/dollar exchange rate). Sample periods vary, depending in part on data availability, but all begin in the early 1970s and end at the beginning of 1997. The exact data definitions and data sources are described in Appendix A.

## A. *Pricing Formulas*

Although numerous option models have been developed over the years to extend the Black–Scholes (BS) framework to new markets and attempt to take account of fat tails, stochastic volatility, and the model's other known shortcomings, the basic BS paradigm is still the most widely used approach for practical option valuation. Typically, marketmakers compute BS values, then add the "tweaks" and subjective adjustments they feel are appropriate for particular markets and times. We adopt the following models for our specific markets.

*Equities and exchange rates:* The original BS (1973) model for option pricing applies to a European call option on a non-dividend-paying stock. The model is easily modified as shown in equation (1) to allow a cash payout at a continuous proportional rate $q$ during the life of the option (see Merton (1973)). For stock index options, we set $q$ equal to the (realized) rate of dividend payout on the underlying index portfolio. Although traders do not know the actual future dividends at the time they must implement the model, dividend payout is quite stable and easily predicted over short intervals. However, dividend uncertainty is another source of model risk for longer term contracts.

$$C = Se^{-qT}N(d) - Xe^{-rT}N(d - \sigma\sqrt{T}), \tag{1}$$

where $C$ is the call value, $S$ is the underlying stock index level, $T$ is the time to expiration, $N(\cdot)$ denotes the cumulative normal distribution function, $X$ is the strike level, $r$ is riskless interest at a continuously compounded rate , $\sigma$ is the estimated annualized volatility, and $d$ is given by

$$d = \frac{\ln \dfrac{S}{X} + \left(r - q + \dfrac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}.$$

Garman and Kohlhagen (1983) present a variant of equation (1) for options on foreign currencies. In their model, $S$ is the level of the exchange rate in dollars per unit of foreign currency, $q$ is set equal to $r_f$, the foreign riskless interest rate, and $r$ is the domestic riskless interest rate.

For put options, the continuous payout option value is

$$P = Xe^{-rT}N(-d + \sigma\sqrt{T}) - Se^{-qT}N(-d). \tag{2}$$

The call and put deltas for equity and FX options are given by

$$\text{Call delta} = e^{-qT}N(d)$$

$$\text{Put delta} = -e^{-qT}N(-d). \tag{3}$$

*Interest rates:* Options on interest rates require special treatment because the underlying asset is not a tangible one. Although many more elaborate interest rate models have been developed over the years, for options on a single rate (unlike options on bonds, which require dealing with the whole term structure of interest rates), the variant of the BS model introduced by Black to price options on futures is widely felt to be adequate (see Black (1976)):

$$C = Re^{-rT}N(d) - Xe^{-rT}N(d - \sigma\sqrt{T}). \tag{4}$$

Here, $R$ is the underlying interest rate that the option is based on, $r$ is the short-term riskless interest rate (which is computed from the same $R$ in the case of options on 90-day LIBOR), the other variables have the same meanings as above, and $d$ is now given by

$$d = \frac{\ln \dfrac{R}{X} + \dfrac{\sigma^2}{2}T}{\sigma\sqrt{T}}.$$

For interest rate puts,

$$P = Xe^{-rT}N\left(-d + \sigma\sqrt{T}\right) - Re^{-rT}N(-d). \tag{5}$$

The call and put deltas for interest rate options are given by

$$\text{Call delta} = e^{-rT}N(d)$$

$$\text{Put delta} = -e^{-rT}N(-d). \tag{6}$$

### *B. Trading Strategy*

The trading strategy we examine is to compute the model value for an option with the desired maturity and strike and to write enough contracts to produce a premium inflow of \$100. This allows easy comparison of performance across contracts with different initial prices. The results of the analysis can be interpreted either as the dollar return on a \$100 initial position, or as percentage returns per dollar of option premium. The proceeds are assumed to be invested, earning the current risk free interest rate in each period. For strategies that involve hedging, cash flows from all subsequent purchases and sales of the underlying asset are assumed to come out of the money market account, making the trading strategy self-financing at all times.

Risk management of an options book can range from essentially none to very sophisticated techniques that attempt to insulate the overall position against both small and large market moves, as well as changes in volatilities and other input parameters. There are three basic ways to limit overall derivatives risk. The first is simply through diversification across different markets and over time, as with stock portfolios and other risky assets. The second is through cash flow matching, such as when the bank holds offsetting long and short positions in the same option contract with different counterparties. The third is delta hedging using the valuation model, by taking positions in the underlying asset that have market exposure of the same magnitude and opposite sign as the option portfolio to be hedged.

Of these, cash flow matching offers the most precise hedge, but it is seldom possible to construct fully matched positions all of the time. If option dealers must be short options overall to satisfy the public demand to buy contracts, it is not possible for all of them to be cash-flow matched. This argument also applies to more complex hedges that attempt to offset the effects of convexity (gamma risk), changing volatility (vega risk), and so on.[4]

[4] Hedging the gammas and vegas of short option positions requires buying options. But given that there must always be a seller for every option purchased and dealers in aggregate are going to be short options, the dealers cannot all hedge these risks. Static hedging strategies simply shift one dealer's risk to other dealers. But those dealers do not want to increase their exposure to gamma and vega risks either, so the necessary options are expensive to buy.

We consider the two possibilities of either (i) not hedging at all and hoping that diversification will cause profits and losses on option writing to balance over time or (ii) delta hedging using the valuation model. The first simply treats an option position like any other risk asset held by the bank, and the latter trades the underlying asset as indicated by the model, in order to minimize the variability of return on each written option.

To see how much of the observed model risk is due to incorrect volatility estimates, we also examine pricing and hedging option positions using as the volatility input the actual volatility that is realized over the remaining life of the option. This presents a few difficulties, however. One of them is that when an option approaches expiration, the number of return observations remaining in its life from which to compute realized volatility obviously goes to zero. We deal with this by requiring at least 10 data points for an estimate, which are obtained by extending the period for estimation beyond option maturity when there are fewer than 10 days remaining before expiration. Another problem can occur when the strike for a deep out of the money option is set based on the volatility estimate from historical data, but this turns out to be much larger than the realized volatility over its lifetime: The strike may be set so far out of the money that with the smaller volatility the option value is extremely small. This difficulty is discussed further in Section IV.

## II. Forecasting Volatilities

One of the most important and obvious sources of model risk with options is that even a correct model requires a value for the unknown volatility of the underlying from the present through expiration day. Model error from using a forecasted value in place of the true volatility produces mispricing of derivatives and also inaccurate hedging calculations. To give a fair analysis of model risk exposure, it is important to allow option writers to make optimal use of the information they have available in making their decisions, but not to allow them to peek into the future. This section analyzes alternative procedures for obtaining the most accurate volatility forecasts from historical data.

The most basic estimation procedure is simply to calculate the realized volatility in a sample of recent price data for the underlying and assume that the same value applies over the future life of the derivative one is pricing. Variations on this method involve the choice of how much past data to include, periodicity (e.g., daily data versus monthly data), whether deviations are measured around the sample mean or around an imposed mean value such as zero, and whether to downweight old data.

Figlewski (1997) examines the impact on forecast accuracy of these variations, for volatility of the three month T-bill rate, the 20-year T-bond yield, the S&P 500 index, and the DM/dollar exchange rate. Following is a capsule summary of the findings from that study with regard to the choice of volatility estimator.

- Out of sample forecast errors of even the best method tend to be quite large.
- The forecast error for longer horizons tends to be lower than for short horizons.[5]
- Although in practice it is common to estimate volatility from short historical samples, using a larger amount of past data (several times the length of the forecast horizon or more) generally gives considerably greater accuracy, except when forecasting over the very shortest horizons (e.g., less than three months).
- Estimating from daily data improves accuracy for short horizons (six months or less), but for longer horizons, monthly data give better results, probably because they are not affected as much by transient high frequency noise in market prices.
- Since the statistical properties of the sample mean make it a very inaccurate estimate of the true mean, taking deviations around zero rather than around the sample mean typically increases forecast accuracy.

Another class of volatility estimators based on past data tries to model a time-varying volatility process. The most common approach uses the generalized autoregressive conditional heteroskedasticity (GARCH) framework (see Bollerslev (1986), and Bollerslev, Chou, and Kroner (1992)). GARCH and its variants comprise a broad set of volatility models, but GARCH has some serious shortcomings as a forecasting tool. One is that the parameters must be estimated from past data, and this frequently requires quite a large data set. Another is that each period's volatility depends on the disturbance in the immediately preceding period, data that are not available when projections must be made many periods ahead. These problems are discussed in more detail in Figlewski (1997). We do not consider GARCH-based estimators here.[6]

The other major type of volatility estimator is the implied volatility (IV) derived from market prices of traded options. IV is felt by many practitioners, and academics as well, to be the best estimate, even though empirical research does not establish that IV is highly accurate as an estimate of true

[5] As an illustration, for each month from January 1952 through December 1990, estimating volatility of the S&P 500 stock index from five years of past monthly data and using it to forecast volatility over the next two years gives a root mean squared error (RMSE) of 4.17 percent when the realized volatility averages 14.25 percent; RMSE for three- and five-year forecasts, respectively, are 3.62 percent and 3.10 percent.

[6] Figlewski (1997) examined the forecasting performance of the GARCH(1,1) estimator, perhaps the simplest member of the family and the one most widely used for financial data. GARCH was found to be useful primarily for short-term forecasting of stock returns volatility with daily data; for longer horizons and in different markets it did not work as well as historical volatility. Moreover, for the cases in which GARCH was the most accurate estimator, the errors were still very large. For example, RMSE in forecasting daily volatility of the S&P 500 index over a three-month horizon was 5.37 percent relative to the average realized daily volatility of 13.29 percent.

volatility.[7] Pricing options based on IV simply amounts to pricing them the same way the market currently does. This can be useful to a marketmaker regardless of whether IV is an accurate forecast of future price variability. However, to make IV into a viable forecasting tool requires a broad range of prices for traded options with different strike prices and maturities. Lacking such data, we do not examine IV here.

We therefore consider pricing and hedging options using simple volatility estimates drawn only from past returns data that would have been available at the time the forecast was made. This still leaves quite a bit of leeway regarding how much data to include in the estimate. It is reasonable to suppose that the length of the data sample that gives the most accurate forecast is a function of the forecast horizon, so that a good one-month forecast may be calculated from a few months of recent returns data, but the best five-year forecast requires looking much further back in time.

Figlewski (1997) examines a forecast procedure that at each date computes the unconditional volatility (i.e., as if it were a fixed constant) but optimizes over the amount of past data included, so as to achieve the minimum RMSE out of sample predictions. In the spirit of ARCH-family nomenclature, this technique is dubbed OUCH, for "optimized unconditional conditional heteroskedasticity."

We also consider an "All available" estimator that uses all past data back to the beginning of the sample as of each date. Thus, the amount of historical data in this volatility estimate grows over time, from five years on the first date to 20 years or more on the last.

The nature of the estimation error in predicting volatility from our historical data samples with an OUCH model is illustrated in Table I. We examine daily and monthly data on the S&P 500 stock index, three-month U.S. dollar LIBOR, the 10-year T-bond yield, and the DM/dollar exchange rate. For each market, several different forecasting horizons are examined, and for each, we try three historical sample lengths and also a forecast using exponentially declining weights. The details of our procedure for constructing out of sample volatility forecasts are described in Appendix B.

We note that the forecast errors are not serially independent because the volatilities calculated for consecutive dates are computed from almost exactly the same data points.[8] Lack of independence produces serial correla-

[7] One large study, by Canina and Figlewski (1993), found that IVs from the S&P 100 stock index options market (one of the most active in the United States) appeared to contain no information at all about future realized volatilities. Figlewski (1997) reviews a number of empirical studies of the forecast accuracy of implied volatilities. The typical study shows that IV does contain information about future volatility, and generally more than the particular variant of historical volatility to which it is compared, but that IV is biased and forecast errors are substantial.

[8] Lack of independence does not bias the estimated RMSEs (although it does cause inconsistency in an estimate of the standard error of the RMSE). We think of this procedure as a way to assess the impact of estimation error for a financial institution that writes options every period (i.e., daily or monthly), basing pricing and hedging strategy on volatility values estimated from past data in a standard way.

tion in the pricing errors, which shows up in the form of runs of losing and winning months. To give an idea of how this impacts performance, in later tables we report both the worst single month and the worst full year for the option writing strategy.

An OUCH estimate weights each past observation equally, but it is often felt that recent observations are more meaningful than ones from the distant past. A relatively easy way to take account of this is to weight each data point in inverse proportion to its age. The exponentially declining weight forecast is computed as in equation (7), where $0 < w \leq 1$ is the weighting factor and $R_t$ is the log price relative from date $t$.

$$\sigma_{\text{forecast at } T} = \left( \frac{\sum_{t=1}^{t=T} w^{T-t} R_t^2}{\sum_{t=1}^{t=T} w^{T-t}} \right)^{1/2}. \tag{7}$$

In the limit, equation (7) includes an infinite number of past observations—but with those from the distant past having infinitesimally small weights. Thus, depending on the value of $w$, the weighted forecast may be largely determined by data from the recent past. One way to measure the effect of this weighting scheme is by the mean lag, which is the weighted average date for the sample points in the estimate. The mean lag can be computed simply as

$$\text{Mean lag} = \frac{1}{1 - w}. \tag{8}$$

For example, the average $w$ value for forecasting LIBOR daily volatility over one month is shown in Table I, Panel A, as $w = 0.973$. From equation (8) this $w$ corresponds to a mean lag of 37 trading days. Note that the mean lag figures shown in Panel B are in months.

Table I, Panel A, presents forecast results for daily volatility over 1-, 3-, and 12-month horizons using daily historical data. Panel B does the same thing for monthly observations to forecast over two- and five-year horizons. In the daily table, for convenience in estimation, a "month" is defined to be 21 trading days in all cases. Notice that the samples span slightly different periods. In particular, it is only necessary to hold out 12 months of daily data at the end of the sample for postsample forecasting, rather than five years as in the monthly table.

In each column, the shaded figure indicates the minimum RMSE historical sample size (or, in some cases, the exponentially weighted forecast). In the next sections, we assume that the volatility input to the option pricing model on each date is computed using this method. This is the one place where we have not strictly adhered to the requirement that all decisions be based only on information that would have been available at the time: The analysis shown in these tables indicating which horizon is best uses data

**Table I**

**Forecast Accuracy of Volatilities Estimated from Historical Data**

The table shows the root mean squared forecast error for annualized volatility calculated from daily data around a mean of zero for different forecast horizons and historical sample lengths. Also reported is the performance of exponential weighting across all past data, in which the optimal weight on each date is the one that would have minimized the forecast error for that forecast horizon over the historical sample available at the time. A "month" is defined as 21 trading days. Shading indicates the minimum RMSE estimation method.

Panel A: Daily Data

| | Forecast Horizon (months) | | | | Forecast Horizon (months) | | |
|---|---|---|---|---|---|---|---|
| | 1 | 3 | 12 | | 1 | 3 | 12 |
| S&P 500 Stock Index | | | | US$ London Inter Bank 3-Month Rate | | | |
| Dec. 30, 1975–Jan. 3, 1996 | | | | Nov. 1, 1979–Jan. 12, 1996 | | | |
| 3 | 7.2% | 7.2% | 7.1% | 3 | 10.2% | 9.3% | 8.4% |
| 12 | 7.5% | 7.1% | 6.8% | 12 | 10.2% | 8.8% | 8.1% |
| 60 | 7.8% | 7.2% | 6.5% | 60 | 13.0% | 11.5% | 10.3% |
| Exponential wgts | 7.0% | 6.7% | 6.5% | Exponential wgts | 9.8% | 8.7% | 8.0% |
| Average weight | 0.979 | 0.986 | 0.990 | Average weight | 0.973 | 0.984 | 0.987 |
| Mean lag (days) | 47 | 73 | 103 | Mean lag (days) | 37 | 64 | 80 |
| Average Realized | 12.9% | 13.2% | 13.7% | Average Realized | 20.8% | 21.5% | 21.5% |
| 10-Year Treasury Bond Yield | | | | Deutsche mark Exchange Rate | | | |
| Jan. 26, 1976–Jan. 5, 1996 | | | | Jan. 22, 1976–Jan. 10, 1996 | | | |
| 3 | 4.6% | 4.4% | 4.4% | 3 | 3.9% | 3.6% | 3.3% |
| 12 | 4.8% | 4.2% | 4.0% | 12 | 3.8% | 3.2% | 2.8% |
| 60 | 5.7% | 5.1% | 4.6% | 60 | 4.0% | 3.3% | 2.5% |
| Exponential wgts | 4.7% | 4.6% | 4.7% | Exponential wgts | 4.0% | 3.3% | 2.8% |
| Average weight | 0.976 | 0.986 | 0.992 | Average weight | 0.981 | 0.994 | 0.995 |
| Mean lag (days) | 41 | 70 | 132 | Mean lag (days) | 52 | 160 | 203 |
| Average Realized | 12.7% | 13.0% | 13.5% | Average Realized | 10.0% | 10.3% | 10.5% |

Panel B: Monthly Data

| | Forecast Horizon | | | Forecast Horizon | |
|---|---|---|---|---|---|
| | 24 | 60 | | 24 | 60 |
| S&P 500 Stock Index | | | US$ London Inter-Bank 3-Month Rate | | |
| January 1976–December 1991 | | | January 1976–December 1991 | | |
| 24 | 5.9% | 5.4% | 24 | 13.3% | 11.9% |
| 60 | 4.9% | 4.5% | 60 | 12.6% | 10.5% |
| All available | 4.0% | 3.2% | All available | 15.2% | 13.5% |
| Exponential wgts | 4.2% | 4.1% | Exponential wgts | 13.7% | 12.8% |
| Average weight | 0.976 | 0.987 | Average weight | 0.898 | 0.942 |
| Mean lag (months) | 42 | 79 | Mean lag (months) | 10 | 17 |
| Average Realized | 15.4% | 15.2% | Average Realized | 25.4% | 25.3% |
| 10-Year Treasury Bond Yield | | | Deutsche mark Exchange Rate | | |
| January 1976–December 1991 | | | January 1976–December 1991 | | |
| 24 | 5.4% | 5.0% | 24 | 3.5% | 2.8% |
| 60 | 4.7% | 4.2% | 60 | 2.5% | 1.3% |
| All available | 4.4% | 3.4% | All available | 2.4% | 1.4% |
| Exponential wgts | 5.3% | 3.7% | Exponential wgts | 2.7% | 1.5% |
| Average weight | 0.971 | 0.991 | Average weight | 0.977 | 0.987 |
| Mean lag (months) | 35 | 107 | Mean lag (months) | 44 | 77 |
| Average Realized | 15.2% | 15.7% | Average Realized | 12.3% | 12.3% |

from the entire sample. Thus, early in the sample, without knowledge of subsequent returns behavior, an actual bank following the procedure we have outlined might not have used the minimum RMSE historical sample length, and would have made less accurate forecasts than what we are assuming here.

One striking result is how large the forecasting errors are. For example, on average the forecast of three-month LIBOR volatility over the next 24 months based on the preceding 24 months has an RMSE of 13.3 percent. This is very large relative to a mean realized volatility of 25.4 percent. Roughly speaking, this means that about one-third of the time, the predicted volatility would be more than 50 percent above or below the true value. LIBOR is actually the worst case, but forecast errors are substantial for all of these markets.

These results are consistent with those reported in Figlewski (1997) for different sample periods and different markets. For monthly data, in three of the four cases the best estimates come from using the largest possible historical sample, but for daily observations and shorter forecasting horizons, performance is better if the historical sample is several times as long as the horizon, but not too long. Surprisingly, forecasting accurately over longer horizons seems to be easier than over shorter ones.[9]

One interesting feature of these results is that (annualized) volatility for daily data appears to be substantially lower than for monthly data. This may be due in part to the somewhat different sample periods. It also may be a result of short-term positive serial correlation in the daily data series. If price changes are not independent over time, estimated volatility is affected. Positive autocorrelation, which occurs when observed prices adjust to new information with a lag over short intervals, reduces estimated volatility. This problem largely disappears with longer differencing intervals, which is the reason to use monthly rather than daily data for longer horizon forecasting.

## III. The Risk and Returns to Option Writing Using Forecasted Volatilities

The various simulations involving different markets, option classes, holding periods, hedging techniques, and forecast methods produce a tremendous amount of output. We wish to make the full range of results available to those who may have a particular interest in some specific subset of them. However, the complete set of tables examining both calls and puts struck at the money and slightly out of the money (0.4 standard deviations away from the current asset price in terms of the estimated volatility) with greater detail on performance using realized volatility, goes well beyond what is necessary to see the nature of our findings. We therefore present our results

[9] The result that volatility forecasting errors do not grow larger for longer horizons would suggest the existence of mean reversion in volatility. However, mean reversion alone does not explain why forecasting over the immediate future is *less* accurate than over longer horizons.

in this article by extracting representative subsamples for close examination and describing results for the others in general terms. The full set of simulation results is available to be downloaded from the Internet.[10]

### *A. Managing Option Risk by Diversification*

Table II examines the impact of volatility forecasting errors on a bank or financial institution that writes options each period (either every day or every month). In each case, enough options are sold to produce $100 of premium income, so the performance figures may be interpreted either as dollar amounts or as percentages of the initial option price. The top portion of the table presents results for at the money calls for the five daily and monthly forecast horizons examined in Table I. For comparison, at the money puts are examined in the lower panel for just two horizons: Three months with daily observations and two years with monthly observations.

Table II looks at the strategy of selling the options at their model values, investing the proceeds at the risk-free interest rate, and simply holding the short position until maturity without hedging. A bank may consider an option as just another kind of risky asset and try to deal with option risk exposure essentially through diversification rather than hedging, by simply adding options positions into its overall portfolio of risky assets. This is the strategy examined here.

The table uses the lowest RMSE forecast method for each horizon from Table I, so the results reflect model error due both to inaccurate volatility inputs and to errors in the option pricing models (such as a failure to fit tail probabilities accurately). The rightmost two columns show the mean return and standard deviation if the same options are priced using the realized volatility over the option's life as the input to the model. This removes the volatility estimation error, leaving only the effect of inaccuracy in the valuation model itself, along with random errors due to the finite sample.

For each market and option maturity, the first two columns give the mean return and standard deviation of the strategy. These are reported on a "per trade" basis—that is, not annualized. For example, writing $100 worth of one-month S&P 500 Stock Index calls lost on average $$-21.52$ over the month, with a standard deviation of $158.57 across months.

Since there is no hedging, the theoretical mean return to option writing should be a function of the expected value of the change in the underlying asset. For options on the S&P stock index, this should be positive and well above the risk-free interest rate for calls, and negative for puts. Over the long run, stocks have averaged returns between eight and nine percent above Treasury bills, and as leveraged instruments, options should have larger risk premia than the underlying stock index. Thus, a call-writing strategy for the S&P 500 index should lose money on average, as it does here, and

[10] An extended set of tables, along with the text of this article, can be obtained as an Adobe Acrobat file from Stephen Figlewski's website. The URL is ⟨http://www.stern.nyu.edu/~sfiglews⟩.

**Table II**

**Return and Risk in Writing Options without Hedging, Using Estimated Volatilities, for At the Money Calls and Puts**

The top portion of the table reports the performance of a strategy of writing unhedged at the money call options each period for several maturities. The period is daily (D) or monthly (M). Option prices are computed using the the minimum RMSE volatility estimate with historical data. The rightmost columns give the results using the realized volatility over the option's lifetime as the volatility input. The strategy always writes enough options to produce $100 of option premium. The bottom portion of the table shows comparable results for at the money put options with three month and two year maturities.

| | | | Minimum RMSE Volatility Forecast Using Historical Data | | | | | | | Realized Volatility | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Percentage | Worst Case | | Worst Full Year | | | |
| Underlying | Maturity | Freq | Mean Return | Std. Dev. | In-the-money | Date | Return | Year | Mean | Mean | Std. Dev. |
| At-the-money calls | | | | | | | | | | | |
| S&P 500 stock index | 1 month | D | −21.52 | 158.57 | 61% | 8/12/82 | −1035.20 | 1975 | −525.52 | −29.17 | 157.54 |
| | 3 months | D | −33.57 | 159.51 | 68% | 8/11/82 | −1077.77 | 1975 | −334.20 | −38.47 | 155.55 |
| | 1 year | D | −56.22 | 164.20 | 76% | 7/28/82 | −688.87 | 1982 | −344.38 | −78.25 | 173.38 |
| | 2 years | M | −59.96 | 142.93 | 92% | Sep-85 | −602.51 | 1985 | −369.70 | −75.49 | 148.31 |
| | 5 years | M | −172.46 | 182.41 | 100% | Jul-82 | −883.74 | 1982 | −568.82 | −174.06 | 182.77 |
| 3-month US$ LIBOR | 1 month | D | 2.97 | 163.37 | 43% | 2/1/94 | −1140.84 | 1994 | −184.31 | 3.33 | 144.69 |
| | 3 months | D | 5.12 | 164.57 | 43% | 2/14/94 | −837.47 | 1994 | −307.20 | −6.42 | 174.47 |
| | 1 year | D | −20.80 | 247.74 | 38% | 12/17/93 | −1268.64 | 1993 | −456.56 | −42.47 | 288.61 |
| | 2 years | M | −16.14 | 250.64 | 36% | Feb-93 | −836.53 | 1993 | −749.55 | −24.97 | 261.14 |
| | 5 years | M | 144.01 | 9.74 | 5% | May-85 | 108.51 | 1990 | 129.04 | 143.24 | 12.53 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10-year Treasury yield | 1 month | D | −30.89 | 215.31 | 50% | 9/21/79 | −2189.14 | 1977 | −218.78 | −22.94 | 167.62 |
| | 3 months | D | −40.16 | 221.98 | 53% | 8/7/79 | −1473.74 | 1979 | −386.28 | −35.09 | 188.74 |
| | 1 year | D | −70.44 | 274.19 | 50% | 2/20/79 | −1860.33 | 1979 | −676.18 | −47.16 | 214.96 |
| | 2 years | M | −60.28 | 307.71 | 43% | Sep-79 | −1292.01 | 1979 | −862.62 | −13.62 | 204.94 |
| | 5 years | M | 37.54 | 250.87 | 29% | Aug-77 | −1096.42 | 1977 | −668.10 | 80.17 | 155.03 |
| Deutsche mark exchgange rate | 1 month | D | 5.36 | 155.32 | 47% | 2/26/91 | −1132.42 | 1991 | −97.77 | 1.96 | 146.66 |
| | 3 months | D | 6.91 | 162.82 | 45% | 2/8/91 | −1184.26 | 1991 | −168.60 | 6.56 | 147.97 |
| | 1 year | D | 15.77 | 146.65 | 46% | 8/21/92 | −704.32 | 1980 | −327.21 | 20.35 | 131.42 |
| | 2 years | M | 37.05 | 121.94 | 43% | Dec-91 | −401.84 | 1980 | −221.30 | 38.79 | 119.19 |
| | 5 years | M | 88.49 | 135.57 | 27% | Feb-80 | −426.01 | 1980 | −213.79 | 89.96 | 133.10 |
| At-the-money puts | | | | | | | | | | | |
| S&P 500 stock index | 3 months | D | 23.94 | 178.72 | 32% | 8/31/87 | −1210.56 | 1987 | −156.42 | 39.51 | 127.34 |
| | 2 years | M | 108.81 | 50.13 | 8% | Jul-80 | −443.22 | 1976 | 19.51 | 104.11 | 75.67 |
| 3-month US$ LIBOR | 3 months | D | −30.39 | 196.33 | 54% | 10/11/84 | −1122.51 | 1991 | −164.07 | −25.38 | 161.60 |
| | 2 years | M | −58.94 | 212.47 | 64% | Sep-90 | −601.90 | 1990 | −507.21 | −80.90 | 198.20 |
| 10-year Treasury yield | 3 months | D | −24.86 | 190.39 | 46% | 1/14/86 | −879.41 | 1985 | −210.06 | −18.22 | 170.36 |
| | 2 years | M | −22.48 | 174.96 | 56% | Jun-84 | −553.65 | 1984 | −440.56 | −12.36 | 151.94 |
| Deutsche mark exchange rate | 3 months | D | −60.26 | 236.14 | 55% | 9/28/77 | −1725.11 | 1977 | −482.25 | −56.32 | 204.22 |
| | 2 years | M | −158.80 | 373.94 | 57% | Sep-77 | −1180.70 | 1977 | −889.15 | −103.96 | 287.31 |

writing puts should be profitable because they have negative betas. For the other three markets, there is no presumption that the expected change is either positive or negative, so our prior expectation is that both call and put writing with reinvestment of the proceeds at the riskless rate should break even on average.

One thing that is clear in these results is that without hedging, standard deviations are very large, and it does not make much difference whether the volatility is known or just forecasted. Indeed, since the strategy simply amounts to taking a bet that the underlying does not move too far in the wrong direction over the option's lifetime, the results are dominated by what these markets actually did during this period of approximately 20 years. Although there were obviously both up and down movements in each market, the overall trend for stocks was strongly upward. It was downward for short-term interest rates like LIBOR, and fairly steady to slightly down for 10-year Treasury yields and for the deutsche mark exchange rate.

Thus, the at the money stock index calls written have a strong likelihood of ending up in the money which, without hedging, produces large losses of −20 percent to −50 percent on average for the shorter term contracts and more than −170 percent of the premium received for the longest term calls (*all* of which finish in the money). Knowledge of the true volatility that would be realized over the option's life does not improve profitability. Indeed, mean returns for index calls are a little worse and the standard deviations are about the same. For other markets, the "write and hold" strategy would have produced losses on average in some cases and profits in others, but risk exposure as measured by the standard deviation is quite large in every case, generally well over 100 percent of the initial premiums. Not surprisingly, we find substantially greater standard deviations for out of the money options (not shown here).

Note that we do not consider the rather large mean losses reported in Table II to be necessarily indicative that the true expected return to the option writing strategy is negative. Rather, the sampling error is very large, as indicated by the large standard deviations, so that observed mean returns may deviate far from zero over extended periods, even if the options are being priced correctly on average.

Though it is customary to measure risk exposure in terms of standard deviation, the returns to writing options are negatively skewed, so standard deviation does not tell the complete story. The most the strategy can earn is 100 (plus riskless interest) if all written options end up out of the money, but there is no limit to the potential loss from writing contracts that go deep in the money. Another dimension of risk exposure is the single worst loss registered by a given strategy. The Worst Case columns in Table II give those figures and show that writing options unhedged can produce extremely serious losses. In most cases, the worst single trade for a given market loses more than eight times the amount of premium income received. The most disastrous outcome of all those shown here is from writing calls on the 10-year Treasury yield, which produces a loss about 22 times larger than the initial premium received for one-month options written on September 21, 1979. Distinctly *worse* worst cases are found when we look at writing out of the money contracts, such as one-year puts on the DM written on September 28, 1977, that are out of the money at the outset by only $0.4\sigma$, but produce a loss of more than \$29,000 per \$100 initial premium received. Unlike the mean and standard deviation figures, in many cases the worst outcomes are less bad when the true volatility is known than when the historical estimate is used.

A problem alluded to above is that volatility forecasts estimated from long historical data samples do not change rapidly from day to day or from month to month. This means that there may be strings of losing trades, in which the options written over a number of periods all go bad in sequence. Along with the worst single trade, we also show the worst calendar year and the mean returns for options written in that year. For example, five-year, at the money S&P call options written every month during calendar year 1982 lose on average 568 percent of the initial premium. In other words, for the \$1200 of premium received from writing calls each month in 1982, without hedging the writer would have lost a total of $12 \times -568.68 = -\$6825.84$ in the year they matured.

These results show clearly that the strategy of writing and holding option positions without hedging entails very large risk exposure, even when pursued consistently over a long period. Diversification of the outcomes over time is not sufficient to control risk to reasonable levels. Results for the other cases not shown here are consistent with these. Out of the money options have somewhat higher standard deviations; for example, if the figure is 200 for the at the money option, it may be 300 for options struck $0.4\sigma$ out of the money. As Table II indicates, the risks in writing puts do not seem to be much different from those for calls.

### *B. Managing Option Risk by Delta Hedging*

Table III gives the results when the written options positions are delta hedged over their full lifetimes. For options written on the S&P 500 or on the DM, we assume hedging is done using the underlying cash instrument—either the S&P stock portfolio or the appropriate number of deutsche marks. The hedge ratio adopted is the model delta from equation (3), and the quantity traded is set to produce "dollar equivalence" in the hedge.[11] The delta is recomputed and the hedge rebalanced each period (daily or monthly) using the current volatility estimate from data available as of that date.

For the two interest rate contracts, a different hedging approach is required. For instance, it is not feasible to hedge changes in future LIBOR (i.e., the level of LIBOR on the date the option's payoff is determined) by trading actual securities. Instead, we assume the interest rate options would

[11] "Dollar equivalence" means that when the price of the underlying asset changes, the quantity of the hedge instrument held (theoretically) changes in value by an amount exactly equal in dollar magnitude and opposite in sign to the change in the value of the item being hedged.

**Table III**

**Return and Risk in Writing Options with Delta Hedging, Using Estimated Volatilities, for At the Money Calls and Puts**

The top portion of the table reports the performance of a strategy of writing at the money call options each period for several maturities and delta hedging the position. The period is daily (D) or monthly (M). Option prices are computed using the minimum RMSE volatility estimate with historical data. The rightmost columns give the results using the realized volatility over the option's lifetime as the volatility input. The strategy always writes enough options to produce $100 of option premium. Hedge rebalancing is done each period using the updated volatility estimate. The bottom portion of the table shows comparable results for at the money put options with three-month and two-year maturities.

| | | | Minimum RMSE Volatility Forecast using Historical Data | | | | | | | Realized Volatility | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Percentage | Worst Case | | Worst Full Year | | | |
| Underlying | Maturity | Freq | Mean Return | Std. Dev. | In-the-money | Date | Return | Year | Mean | Mean | Std. Dev. |
| At the Money Calls | | | | | | | | | | | |
| S&P 500 stock index | 1 month | D | −0.07 | 33.38 | 61% | 10/16/87 | −536.22 | 1987 | −23.99 | −3.18 | 12.41 |
| | 3 months | D | 1.01 | 24.74 | 68% | 10/16/87 | −329.05 | 1987 | −23.72 | −2.51 | 5.14 |
| | 1 year | D | 3.26 | 42.40 | 76% | 1/9/87 | −290.53 | 1987 | −99.22 | −4.66 | 3.52 |
| | 2 years | M | −0.36 | 22.35 | 92% | Jun-86 | −57.17 | 1986 | −43.87 | −6.71 | 6.81 |
| | 5 years | M | −11.93 | 10.11 | 100% | Jun-86 | −28.89 | 1986 | −26.02 | −12.38 | 5.59 |
| 3-month US$ LIBOR | 1 month | D | −13.71 | 99.11 | 43% | 5/16/91 | −521.86 | 1994 | −95.86 | −24.40 | 103.42 |
| | 3 months | D | −32.06 | 115.14 | 43% | 4/30/84 | −682.32 | 1994 | −236.75 | −49.87 | 124.99 |
| | 1 year | D | −108.34 | 149.64 | 38% | 4/4/94 | −810.39 | 1994 | −403.44 | −132.56 | 179.89 |
| | 2 years | M | −116.02 | 184.29 | 36% | Feb-93 | −768.90 | 1993 | −612.18 | −158.12 | 168.47 |
| | 5 years | M | −75.44 | 71.07 | 5% | Aug-87 | −251.02 | 1987 | −174.99 | −163.20 | 109.18 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10-year Treasury yield | 1 month | D | −24.68 | 74.18 | 50% | 9/20/79 | −929.23 | 1978 | −59.72 | −26.56 | 56.26 |
| | 3 months | D | −34.54 | 71.12 | 53% | 7/23/79 | −567.52 | 1978 | −109.04 | −41.05 | 64.17 |
| | 1 year | D | −59.11 | 92.25 | 50% | 10/17/78 | −553.46 | 1993 | −244.71 | −60.49 | 89.22 |
| | 2 years | M | −60.72 | 82.64 | 43% | Jan-87 | −303.88 | 1993 | −193.81 | −55.86 | 83.03 |
| | 5 years | M | −64.93 | 68.35 | 29% | Sep-79 | −259.36 | 1979 | −179.73 | −52.37 | 69.16 |
| Deutsche mark exchange rate | 1 month | D | 3.80 | 36.25 | 47% | 12/19/77 | −372.73 | 1991 | −18.84 | −2.54 | 12.20 |
| | 3 months | D | 2.75 | 27.16 | 45% | 3/6/95 | −147.93 | 1991 | −19.74 | −1.67 | 4.99 |
| | 1 year | D | −0.43 | 24.21 | 46% | 3/11/91 | −96.95 | 1991 | −31.47 | −1.28 | 4.39 |
| | 2 years | M | −2.72 | 25.39 | 43% | Nov-90 | −109.04 | 1978 | −33.13 | −1.24 | 10.53 |
| | 5 years | M | −2.94 | 23.79 | 27% | Aug-78 | −40.02 | 1978 | −31.11 | −2.65 | 8.68 |
| At the Money Puts | | | | | | | | | | | |
| S&P 500 stock index | 3 months | D | 5.00 | 34.72 | 32% | 10/16/87 | −451.05 | 1987 | −30.87 | 2.23 | 7.73 |
| | 2 years | M | 14.22 | 39.23 | 8% | Jun-86 | −101.59 | 1986 | −75.34 | 9.30 | 23.33 |
| 3-month US$ LIBOR | 3 months | D | 30.42 | 118.75 | 54% | 12/19/95 | −687.48 | 1989 | −87.83 | 31.30 | 113.41 |
| | 2 years | M | 136.15 | 118.36 | 64% | Jan-89 | −45.23 | 1989 | 18.24 | 123.96 | 112.47 |
| 10-year Treasury yield | 3 months | D | 10.04 | 57.24 | 46% | 10/22/79 | −383.25 | 1982 | −33.87 | 7.07 | 46.39 |
| | 2 years | M | 48.70 | 61.23 | 56% | Jan-81 | −130.54 | 1979 | −63.62 | 49.89 | 44.74 |
| Deutsche mark exchange rate | 3 months | D | 6.31 | 38.49 | 55% | 12/14/77 | −368.30 | 1977 | −23.55 | 1.12 | 7.12 |
| | 2 years | M | −4.70 | 49.12 | 57% | Aug-78 | −271.28 | 1978 | −137.30 | 5.81 | 18.04 |

be hedged using futures contracts: 90-day Eurodollar futures traded on the Chicago Mercantile Exchange, and 20-year Treasury bond futures traded at the Chicago Board of Trade. The T-bond contract was first introduced in 1977 and the Eurodollar contract did not begin trading until 1982, so the sample had to be trimmed at the beginning.

Hedging long-term option contracts with futures presents several important questions of hedging tactics. Ideally, the futures expiration should be set equal to the maturity date of the option, but this was not feasible for longer term options, because futures with long enough maturities did not always exist, and when they did exist, they were often highly illiquid. Also, in the case of the T-bond contract, any open long position (such as would be taken to hedge written put options on the Treasury yield) that was carried into the contract's expiration month could potentially result in undesired delivery of the underlying bonds. We therefore adopt the following procedures.

Options on the Treasury bond yield are hedged in all cases with the nearest-to-expiration future (which has the greatest liquidity). These positions are assumed to be closed out and rolled over into the next expiring contract on the last trading day prior to the nearby contract's expiration month. Over time, the Eurodollar futures contract has developed liquidity at more distant maturities and trading has extended to quite long-term contracts. This was not true at first, however. We have therefore assumed that an option with less than one year to expiration is hedged with the nearest future that matures after its expiration date, but longer term options are always hedged with the one-year maturity future. This can translate to a large maturity mismatch for two- and five-year options, and may partly account for the poor hedging performance seen here.

Comparing Table III with Table II, it is clear that delta hedging greatly reduces risk exposure for these written options. For example, the standard deviation of the return on three-month, at the money S&P 500 calls drops from 159.51 to 24.74. In virtually every case, hedging reduces the standard deviation and the magnitude of the worst cases. Nevertheless, the standard deviations of the returns remain sizable—more than 100 percent in some cases. The worst losses still are many times the initial premium received, a result that is particularly true for the out of the money contracts.

The interest rate options hedged with futures contracts retain more of their variability than do the S&P and DM options. Several factors might be contributing to this. First, there is the basis risk that affects futures prices. Futures prices are tied to the price behavior of their underlying assets, but the relationship is flexible so mispricing can arise and persist for some time. The fact that futures may be mispriced at the time a hedge must be rebalanced adds variability to the overall return. Such basis risk is likely to be greater when there is a mismatch between the item being hedged and the future. This occurs here because of the need to use futures with different maturities from the options they are hedging, and, in the case of Treasury bonds, because the future is based on a 20-year bond (or, more precisely, on the cheapest to deliver bond, which might have up to 30 years to maturity) but it is the 10-year rate that determines the option payoff. Thus, changes in the slope of the yield curve produce basis risk in the hedged position. Finally, there is a large amount of empirical evidence that interest rates follow more complex processes than the standard diffusions that the Black model allows for, suggesting that model inadequacy may well be greater for the interest rate options.[12]

One factor that could possibly play a role in causing the high standard errors of hedge returns we observe when using forecasted volatilities is the nonlinearity in the relationships among the estimated variance, the volatility, and the model values for an option's price and delta.[13] The customary procedure, which we follow, is first to use past data to form a statistical estimate of the returns variance, then take the square root of the point estimate (a nonlinear transformation) to get the estimated volatility, and finally use this as an input into the (nonlinear) option pricing and delta equations. Even if the variance point estimate is the expected value of the true variance, the square root of the estimated variance is not equal to the expected value of the volatility, because of the nonlinearity. Moreover, because the option model is nonlinear, the expectations of the option value and delta taken over the probability distribution for the volatility parameter are not the option value and delta evaluated at the expected volatility.

Normally, the bias introduced by these nonlinearities is ignored, because the model is not actually very nonlinear in the volatility input for at the money options. The effect is larger for out of the money options, however, so it is worth considering to what extent the interaction between sampling error and the nonlinearities in the problem could account for some of our results.

We do not attempt an in-depth analysis, but a quick examination of these nonlinearities in the context of a classical variance estimation suggests that the impact is probably quite small for the range of volatilities and option moneyness we consider. Under the assumption of constant variance and zero mean, the sampling distribution of the variance estimate is chi-squared with degrees of freedom equal to the number of observations. We computed true option values and deltas by a numerical integration under the appropriate chi-squared distributions for the range of sample sizes and model parameters in our problem. In most cases, the true expected volatility is higher than the square root of the variance point estimate but differs only in the fourth decimal place, for example 0.1504 versus 0.1500 with one year of daily data. The option values are off by a few cents, and the deltas deviate only in the fourth decimal place. We conclude, therefore, that nonlinearity has an effect on the problem, but not enough to account for any appreciable amount of the hedging risk that we observe.

[12] To look further at the effect of basis risk between the interest rate futures and the underlying rates, we consider (hypothetically) hedging the options as if we could take the appropriate hedge positions directly in the underlying interest rates, rather than the futures. The results look much like those for the S&P 500 and the DM, suggesting that the need to hedge with futures is a primary cause of the poor hedging performance seen here.

[13] We thank Pedro Santa Clara for pointing out this issue.

Table III shows how much estimation error on the volatility input affects the variability of hedge returns. For both S&P and DM options, knowing the volatility would substantially reduce the standard deviation. This does not necessarily hold true for the interest rate contracts, however. Although hedging with futures does reduce position risk for LIBOR and T-bond yield options relative to not hedging at all, the improvement is generally rather limited.

The results illustrate that writing options and delta hedging the positions using the most common valuation models involves exposure to considerable risk, much of it "model" risk. Hedging is much better than not hedging, but a great deal of risk still remains. Note also that we have made no attempt to take into account any transactions costs of running a hedge strategy over long periods. In practice, there would be brokerage fees and market impact costs (e.g., the bid-ask spread) that would increase with the amount of rebalancing needed to maintain delta neutrality. On the other hand, we have also not included any "markup" on the written option contracts. A bank writing options to satisfy its customers should not price them exactly at their model values because in theory the bank should only break even at those prices. It would be normal to raise the price above the model values, which would cover many of the mean losses in these tables. We look at that strategy in more detail in Section V.

## IV. Variations on the Theme

The previous section looked at writing at the money and slightly out of the money options, which we expect to be the ones with the greatest activity, and we assume that the volatility input to the model is the historical estimate with the best forecasting power. We now briefly explore some "variations on the theme," writing deep out of the money or in the money contracts, and using suboptimal volatilities.

We define an "out of the money" contract to be one with a strike price 0.4 standard deviations above the current asset price for a call (or below, for a put). For a lognormal distribution without drift, there is about 0.34 probability that such an option finishes in the money. Continuing this idea, we now define "in the money" and "deep out of the money" options as being 0.4 standard deviations in the money, or 1.0 standard deviations out of the money, respectively. The equivalent probabilities of finishing in the money are about 66 percent and 16 percent, respectively.

Volatility forecasting for option traders is widely thought to be an art as much as a science, with different traders preferring different approaches (generally without submitting their choices to any rigorous testing). There is a common feeling that because volatility changes over time, it is better to eliminate old data from the calculation. Volatility estimates provided by online derivatives services are often computed from three or six months of past daily returns, and often from as little as one month. Using exponentially declining weighted forecasts also has the effect of placing most of the weight on the recent past. For example, the daily volatilities presented by J.P. Morgan's RiskMetrics system originally used data from only the last 25 trading days, with an exponential decay factor of $w = 0.94$. From equation (8), this implies that the mean lag in the estimator was only about 17 days and it contained no information more than 25 days old.[14]

The results shown in Table I, however, suggest that using such a short historical sample leads to larger RMSE forecast errors than would be achieved with more past data. For the one-month horizon, the most accurate volatilities come from at least several months of data; for the longer horizons examined in Table I, Panel B, the most accurate volatilities use all available historical data in three markets out of four, and the fourth still gets the most accurate estimates using 60 past months. This suggests that many option writers may be pricing and hedging their positions using volatilities calculated from less than the optimal amount of historical data.

Results from classical statistics show that when the underlying asset follows a diffusion process, much greater accuracy in estimating volatility can be obtained by using high-frequency data. This suggests that if both daily and monthly data are available, it is much better to use daily data. On the other hand, because the data are prices recorded in financial markets, they may actually deviate quite far from true diffusions at very short intervals. In markets with a bid-ask spread, transaction prices tend to show high negative autocorrelation due to bid-ask bounce. Stock indexes observed daily may show spurious positive serial correlation because of infrequent trading of less liquid stocks. And in still other cases, such as LIBOR, changes at the shortest intervals are not diffusions at all; the series either remains constant or jumps instantaneously by a relatively large discrete amount (e.g., 1/16th or 1/8th of a percent).

These observations suggest that unless the forecast horizon is very short, such that the essentially spurious price effects due to trading noise may nevertheless be relevant determinants of option payoffs, a volatility estimator based on daily data may actually give less accurate values than one with a longer observation interval. In other words, if one is writing two-year contracts, it may well be better to estimate volatility from monthly data than from daily data over the same historical period. We do not provide evidence of this principle here, but we have incorporated it into the analysis by using monthly data for the longer horizon contracts and daily data for the shorter ones.[15]

Table IV presents results on these issues, using three alternative volatility estimators to implement a strategy, for the markets we have been examining, of writing and delta hedging calls and puts with four different strikes and two maturities. The mean and standard deviation of the per trade return are shown, along with the average return per trade in the single worst year for the strategy.

[14] See *RiskMetrics* (1995).

[15] Figlewski (1997) finds that in using five years of historical data to forecast volatility over 24 months, daily observations increase RMSE by approximately five to 10 percent above that achieved with monthly data.

## Table IV
## Return and Risk in Writing Options with Alternative Volatility Estimates and Different Degrees of Moneyness

The table reports the performance of a strategy of writing and hedging call and put options each day for one-month options, or each month for two-year options. For deep out of the money, out of the money, at the money, and in the money contracts, the strike prices are set, respectively, +1.0, +0.4, 0.0, and −0.4 standard deviations above or below the current spot price. S&P 500 and deutsche mark options are hedged with the underlying asset; the interest rate options are hedged using futures, as described in the text. Alternative option values and hedge ratios are computed using the realized volatility, the minimum RMSE forecast, or a volatility estimate from a suboptimal historical data sample. For exponentially weighted estimates, the mean lag in days is shown in parentheses. The hedge is rebalanced to be delta neutral every period, using the current volatility estimate. The strategy always writes enough options to produce $100 of option premium.

| | | | Deep Out of the Money | | | Out of the Money | | | At the Money | | | In the Money | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Underlying | Maturity | Volatility Estimate | Mean | Std. Dev. | Worst Year | Mean | Std. Dev. | Worst Year | Mean | Std. Dev. | Worst Year | Mean | Std. Dev. | Worst Year |
| | | | | | | Panel A: Call Options | | | | | | | | |
| S&P 500 Stock Index | 1 Month | Realized | 8.81 | 75.79 | −10.78 | −3.19 | 21.95 | −7.27 | −3.18 | 12.41 | −5.07 | −2.00 | 7.60 | −4.17 |
| | | Exp. wgt (47) | 8.48 | 113.92 | −71.53 | 2.91 | 55.26 | −32.20 | −0.07 | 33.38 | −23.99 | −1.64 | 20.40 | −18.24 |
| | | 5 years daily | 22.88 | 176.82 | −192.21 | 17.56 | 56.27 | −57.92 | 11.78 | 34.21 | −32.65 | 7.18 | 21.92 | −20.47 |
| | 2 Years | Realized | −2.78 | 51.41 | −34.97 | −7.20 | 14.77 | −21.34 | −6.71 | 6.81 | −14.24 | −5.27 | 4.42 | −9.82 |
| | | All available | 7.51 | 120.97 | −330.32 | 7.82 | 47.30 | −87.20 | −0.36 | 22.35 | −43.87 | −3.22 | 10.80 | −20.59 |
| | | 3 months daily | −90.50 | 254.15 | −816.31 | −23.13 | 58.12 | −106.65 | −14.30 | 27.90 | −56.17 | −9.34 | 14.67 | −26.06 |
| 3-month US$ LIBOR | 1 Month | Realized | −35.72 | 278.25 | −223.13 | −30.53 | 142.42 | −146.74 | −24.40 | 103.42 | −117.20 | −18.60 | 77.69 | −87.49 |
| | | Exp. wgt (37) | −28.25 | 255.65 | −218.52 | −15.26 | 134.43 | −125.76 | −13.71 | 99.11 | −95.86 | −14.44 | 75.80 | −74.92 |
| | | 5 years daily | 3.09 | 266.34 | −225.07 | 12.74 | 122.98 | −127.36 | 8.99 | 89.30 | −97.46 | 3.49 | 68.77 | −76.19 |
| | 2 Years | Realized | −388.36 | 636.53 | −2524.14 | −219.28 | 264.66 | −1031.12 | −158.12 | 168.47 | −631.91 | −120.50 | 119.64 | −427.06 |
| | | 60 months | −240.75 | 588.81 | −2232.65 | −145.79 | 278.84 | −971.22 | −116.02 | 184.29 | −612.18 | −97.74 | 131.18 | −419.01 |
| | | 3 months daily | −1082.95 | 3696.30 | −9214.60 | −293.49 | 492.43 | −1697.72 | −181.09 | 222.24 | −822.71 | −129.48 | 135.39 | −486.96 |
| 10-year Treasury yield | 1 Month | Realized | −58.87 | 195.14 | −235.12 | −40.18 | 84.35 | −105.61 | −26.56 | 56.26 | −67.87 | −17.42 | 40.90 | −46.05 |
| | | 3 months daily | −74.54 | 283.18 | −230.63 | −36.52 | 117.86 | −87.42 | −24.68 | 74.18 | −59.72 | −17.58 | 50.46 | −46.69 |
| | | 5 years daily | −226.52 | 1210.96 | −3560.60 | −46.85 | 170.09 | −438.89 | −22.13 | 80.93 | −141.50 | −12.64 | 50.37 | −54.98 |
| | 2 Years | Realized | −97.93 | 187.87 | −602.18 | −71.74 | 116.47 | −299.59 | −55.86 | 83.03 | −187.08 | −42.91 | 61.24 | −146.38 |
| | | All available | −161.00 | 296.65 | −908.10 | −83.62 | 123.41 | −300.95 | −60.72 | 82.64 | −193.81 | −46.67 | 59.84 | −149.48 |
| | | 3 months daily | −1549.76 | 5428.99 | −11648.51 | −200.30 | 381.52 | −1054.93 | −95.89 | 132.91 | −332.96 | −59.50 | 70.43 | −171.98 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Deutsche mark exchange rate | 1 Month | Realized | 2.18 | 90.90 | −17.51 | −3.15 | 22.87 | −6.90 | −2.54 | 12.20 | −4.42 | −1.38 | 7.55 | −4.29 |
| | | 12 months daily | −3.56 | 136.59 | −125.48 | 5.26 | 60.88 | −43.14 | 3.80 | 36.25 | −18.84 | 1.17 | 22.26 | −11.04 |
| | | 5 years daily | −11.20 | 172.55 | −265.69 | 6.32 | 65.20 | −69.39 | 6.34 | 36.91 | −31.33 | 4.07 | 23.15 | −15.87 |
| | 2 Years | Realized | −3.26 | 25.89 | −32.37 | −1.24 | 17.77 | −20.23 | −1.24 | 10.53 | −12.21 | −1.27 | 7.55 | −8.31 |
| | | All available | −2.58 | 67.94 | −81.04 | 0.59 | 36.82 | −44.83 | −2.72 | 25.39 | −33.13 | −4.71 | 15.66 | −29.62 |
| | | 3 months daily | −67.20 | 277.09 | −484.03 | −10.16 | 83.47 | −160.55 | −6.34 | 45.45 | −67.63 | −6.20 | 28.84 | −69.11 |
| Panel B: Put Options | | | | | | | | | | | | | | |
| S&P 500 Stock Index | 1 Month | Realized | −1.93 | 63.39 | −20.04 | −2.13 | 27.17 | −9.54 | −0.76 | 15.38 | −4.11 | 0.73 | 9.56 | −1.35 |
| | | Exp. wgt (47) | −23.33 | 170.23 | −181.91 | −2.24 | 69.47 | −56.06 | 1.80 | 41.59 | −26.54 | 2.67 | 25.24 | −12.71 |
| | | 5 years daily | −6.74 | 231.68 | −340.40 | 13.92 | 70.70 | −91.61 | 14.63 | 40.68 | −43.94 | 12.33 | 24.73 | −20.63 |
| | 2 Years | Realized | −5.03 | 68.08 | −162.95 | 2.61 | 39.34 | −51.19 | 9.30 | 23.33 | −10.75 | 9.43 | 15.12 | −2.69 |
| | | All available | −9.91 | 93.13 | −183.17 | 4.69 | 55.63 | −91.88 | 14.22 | 39.23 | −75.34 | 15.62 | 29.38 | −56.11 |
| | | 3 months daily | −896.66 | 6052.10 | −10952.68 | −166.12 | 496.74 | −1220.67 | −51.51 | 145.15 | −403.05 | −13.43 | 61.01 | −155.18 |
| 3-month US$ LIBOR | 1 Month | Realized | 67.87 | 366.22 | −81.16 | 36.28 | 157.34 | −41.73 | 27.26 | 107.77 | −37.09 | 20.82 | 77.71 | −27.00 |
| | | Exp. wgt (37) | 19.03 | 268.85 | −143.51 | 29.97 | 148.64 | −48.36 | 25.95 | 106.86 | −33.70 | 19.21 | 79.25 | −24.93 |
| | | 5 years daily | 49.11 | 298.10 | −101.54 | 52.04 | 127.38 | −24.23 | 43.32 | 88.81 | −8.08 | 32.66 | 66.54 | −3.19 |
| | 2 Years | Realized | 294.55 | 401.37 | 3.32 | 178.98 | 183.05 | −32.28 | 123.96 | 112.47 | −16.42 | 87.52 | 76.59 | −4.66 |
| | | 60 months | 337.69 | 378.13 | 63.56 | 196.77 | 181.17 | 29.65 | 136.15 | 118.36 | 18.24 | 96.62 | 83.14 | 11.61 |
| | | 3 months daily | 625.05 | 1914.42 | −507.85 | 197.64 | 255.80 | −137.08 | 132.72 | 136.61 | −18.28 | 95.08 | 90.76 | 4.64 |
| 10-year Treasury yield | 1 Month | Realized | −6.32 | 143.43 | −73.00 | −1.62 | 70.26 | −42.16 | 1.87 | 47.93 | −33.16 | 4.51 | 36.10 | −21.25 |
| | | 3 months daily | −14.96 | 181.82 | −133.34 | 0.93 | 87.93 | −42.73 | 3.51 | 58.71 | −28.46 | 4.41 | 41.58 | −20.77 |
| | | 5 years daily | −78.57 | 1010.81 | −1458.34 | −3.81 | 122.61 | −194.30 | 4.36 | 66.53 | −85.41 | 6.88 | 43.81 | −42.90 |
| | 2 Years | Realized | 106.10 | 107.68 | 11.51 | 68.30 | 61.79 | 1.53 | 49.89 | 44.74 | −4.08 | 40.50 | 34.41 | −5.92 |
| | | All available | 100.13 | 133.05 | −78.67 | 67.28 | 84.36 | −70.52 | 48.70 | 61.23 | −63.62 | 36.72 | 46.90 | −54.83 |
| | | 3 months daily | 105.59 | 1012.07 | −1397.03 | 58.22 | 141.66 | −273.85 | 44.86 | 77.94 | −139.18 | 36.65 | 51.79 | −80.53 |
| Deutsche mark exchange rate | 1 Month | Realized | 6.66 | 119.99 | −15.56 | −2.16 | 30.30 | −8.04 | −1.10 | 14.58 | −4.65 | 0.03 | 8.86 | −3.41 |
| | | 12 months daily | −15.78 | 248.61 | −279.24 | 4.60 | 84.57 | −55.78 | 5.41 | 47.78 | −18.81 | 3.09 | 28.13 | −14.05 |
| | | 5 years daily | −13.26 | 246.22 | −252.03 | 8.28 | 77.58 | −71.15 | 8.73 | 44.82 | −139.18 | 36.65 | 51.79 | −23.97 |
| | 2 Years | Realized | 19.23 | 113.34 | −47.94 | 4.02 | 31.94 | −58.31 | 5.81 | 18.04 | −24.17 | 5.22 | 11.27 | −8.13 |
| | | All available | −118.91 | 327.25 | −844.20 | −26.37 | 126.30 | −404.27 | −4.70 | 49.12 | −137.30 | −0.03 | 25.54 | −50.81 |
| | | 3 months daily | −15454.70 | 90416.03 | −301145.3 | −439.76 | 1616.61 | −5835.24 | −94.40 | 294.45 | −835.89 | −28.64 | 90.58 | −271.65 |

For each market and maturity, the first line of results corresponds to using the realized volatility over the option's remaining lifetime at all times, thus eliminating model risk due to the volatility forecast. The next line shows the results for the best volatility predictor for that horizon, based on out of sample performance (i.e., the shaded cells from Table I). For exponentially weighted forecasts, the mean lag in trading days corresponding to the average decay factor for that market is given in parentheses. The third line presents results based on suboptimal volatility forecasts: five years of daily data for one-month options, and three months of daily data for two-year contracts. Thus for the short-term options, this line uses a historical sample that is substantially longer than indicated, and for the longer term contracts, we use "too little" data, in the form of frequent observations drawn from a very short sample period.

This table allow easy comparison across several important dimensions. Comparing calls versus puts, we see that, though results are obviously different there is no clear distinction in the character of the outcomes. Whether call writing or put writing is more profitable, or riskier, varies from case to case. It is worth noting in passing that, with delta hedging, the profitability of the trade can be expected to depend on the behavior of volatility as much or more than on market direction. If volatility has been underestimated, hedged writing of both calls and puts should lose money, as illustrated by the case of trading two-year S&P 500 options using volatilities computed from the last three months of daily data.

Comparing across strikes for a given market shows that even with hedging there is considerable risk in writing out of the money options, especially those that are deep out of the money. Standard deviations for these contracts are nearly all more than 50 percent of the initial premium, and much greater in some cases. However, some of the truly huge numbers, such as those for DM puts, are somewhat artificial. The strike levels for these options are set in terms of standard deviations based on the "best" volatility forecasts. In cases where the "bad" estimator predicts much lower volatility, these options appear to be very far out of the money. Their model values are then extremely low and very large numbers must be written to produce $100 of premium income. Losses are enormous if even a few cases end up very far in the money.[16]

Model risk due to inaccurate volatility estimates is a significant contributor to overall risk, as evidenced by the differences between the first and second lines in Table IV. For the interest rate contracts, even knowing the true volatility that would occur does not hold the risk down—standard deviation exceeds 100 percent in the majority of cases. Risk is relatively more manageable for the equity and currency options. Comparing the results for the optimal and the suboptimal volatility predictors, one sees that, for the most part, using too much past data for the one-month options does not hurt performance as much as using too little data for the two-year contracts.

[16] It is worth noting that this type of problem is likely to be more significant in theory than in practice because any bank actively writing options is aware of how its competitors are pricing similar contracts. The bank would notice if its prices are well below the market because of an unusual volatility estimate.

Finally, it is interesting to note that the per trade returns do not seem to be much riskier for the long-term contracts than for the shorter ones. This is odd, but it is consistent with the equally counterintuitive forecasting results, shown above and in greater detail in Figlewski (1997), that volatility forecasts are more accurate for longer horizons than for shorter ones.

## V. Writing Options with a Volatility Markup

The results so far indicate that a regular option writing program would entail a large amount of risk even with delta hedging of the market exposure. Moreover, the pricing models are based on the principle that a perfectly hedged position involving an option and its underlying asset should earn the riskless interest rate. But, of course, no financial institution would enter into this business if it is expected to earn only the T-bill rate. The normal procedure when writing an over the counter option contract is to obtain the best available volatility forecast, then increase it by some suitable amount in pricing the option. The higher volatility produces a higher option price, which is expected to compensate the writer for bearing the various kinds of risks and to provide an expected profit above the riskless rate.

Table V examines the volatility markup strategy for call options in our markets. The table shows results for two maturities, daily writing of three-month options and monthly writing of two-year contracts. The premiums are computed by increasing the volatility input from the optimal historical model by 10 percent, 25 percent, or 50 percent (i.e., from $\sigma$ to 1.1 $\sigma$, 1.25 $\sigma$, or 1.5 $\sigma$). Hedging, however, is always done using the unadjusted forecasts. We report the mean and standard deviation of the returns, the percentage of positions that lose money, and the average size of the loss for a losing position.

Boosting the volatility clearly helps to increase the mean returns to option writing. Standard deviations are only slightly affected, but the fraction of trades that lose money is decreased rather sharply in most cases. For example, writing three-month, at the money S&P 500 index calls loses money 46.8 percent of the time when options are priced at their estimated Black–Scholes values, but increasing the volatility input by one-quarter lowers the loss percentage to 16.0. Interestingly, although each formerly losing trade should lose less after the volatility boost, in many cases the average loss becomes worse—that is, only a much smaller fraction of the trades end up losing, but the losses that remain are big ones.

Thus, boosting the volatility input turns a very risky trading strategy that exposes the writer to very large potential losses into one that may allow quite substantial profits on average and relatively few losses. However, the losing trades still entail large losses. For example, for deep out of the money S&P 500 calls at the two-year horizon, boosting volatility by 25 percent cuts the fraction of losing trades from 37.3 percent to 15.4 percent, but those that do lose money lose on average almost 140 percent of the initial premium. This pattern is more striking for writing put options. We find that put writing is profitable in most cases, and raising the volatility input could produce

## Table V

## Return and Risk in Writing and Delta Hedging Call Options with a Volatility Markup

The table reports the performance of a strategy of writing and hedging call options each day for three-month options, or each month for two-year options. Option prices are set by valuing the option with the appropriate model but with a volatility input that is "marked up" by 0 percent, 10 percent, 25 percent, or 50 percent (i.e., the best volatility forecast is multiplied by 1.0, 1.1, 1.25, or 1.5.) Hedging and hedge rebalancing are done with the unadjusted volatility. The table shows the mean return per $100 of option premium, the standard deviation of the return, the percentage of positions that lose money, and the average loss for losing positions. See also the notes to Table IV.

| | | Deep Out of the Money | | | | Out of the Money | | | | At the Money | | | | In the Money | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Markup | Mean | Std. Dev. | Losses | Avg. Loss | Mean | Std. Dev. | Losses | Avg. Loss | Mean | Std. Dev. | Losses | Avg. Loss | Mean | Std. Dev. | Losses | Avg. Loss |
| S&P 500 | | | | | | | | | | | | | | | | | |
| 3 month | 0% | 3.7 | 87.6 | 41.1% | −71.5 | 3.9 | 41.8 | 45.2% | −31.9 | 1.0 | 24.7 | 46.8% | −18.9 | −1.8 | 14.3 | 50.9% | −11.6 |
| | 10% | 30.5 | 88.2 | 27.7% | −73.7 | 17.9 | 42.2 | 31.4% | −29.3 | 9.5 | 25.1 | 32.2% | −17.5 | 3.1 | 14.5 | 33.7% | −11.6 |
| | 25% | 73.6 | 89.4 | 15.4% | −77.7 | 38.9 | 43.1 | 16.4% | −27.3 | 22.3 | 25.6 | 16.0% | −17.5 | 10.6 | 14.8 | 17.8% | −12.0 |
| | 50% | 150.8 | 92.6 | 5.4% | −84.7 | 74.2 | 44.7 | 4.8% | −25.7 | 43.6 | 26.6 | 5.7% | −15.6 | 23.5 | 15.4 | 6.5% | −11.6 |
| 2 year | 0% | 7.5 | 121.0 | 37.3% | −100.7 | 7.8 | 47.3 | 46.7% | −28.4 | −0.4 | 22.3 | 59.9% | −14.1 | −3.2 | 10.8 | 73.1% | −8.4 |
| | 10% | 32.4 | 122.2 | 27.8% | −107.6 | 20.3 | 48.7 | 34.0% | −25.6 | 7.1 | 23.4 | 39.2% | −13.7 | 0.9 | 11.5 | 45.8% | −8.6 |
| | 25% | 71.4 | 124.6 | 16.0% | −139.4 | 39.1 | 50.9 | 17.5% | −24.6 | 18.3 | 24.9 | 23.6% | −10.3 | 7.4 | 12.7 | 30.2% | −6.5 |
| | 50% | 139.3 | 130.7 | 9.9% | −129.9 | 70.5 | 55.0 | 6.1% | −22.0 | 37.3 | 27.7 | 5.7% | −6.5 | 18.8 | 14.5 | 5.2% | −2.8 |
| US$ LIBOR | | | | | | | | | | | | | | | | | |
| 3 month | 0% | −70.3 | 308.9 | 47.1% | −244.5 | −40.2 | 161.9 | 52.6% | −138.3 | −32.1 | 115.1 | 57.3% | −100.1 | −27.9 | 86.5 | 60.5% | −79.0 |
| | 10% | −39.4 | 308.9 | 35.2% | −291.5 | −23.8 | 161.9 | 45.7% | −141.5 | −21.9 | 115.1 | 51.5% | −100.7 | −21.9 | 86.5 | 57.3% | −77.3 |
| | 25% | 11.4 | 308.9 | 23.8% | −370.0 | 1.1 | 161.9 | 35.3% | −155.1 | −6.6 | 115.1 | 43.7% | −102.0 | −12.8 | 86.5 | 52.5% | −74.9 |
| | 50% | 104.6 | 308.8 | 16.3% | −430.4 | 43.3 | 161.9 | 24.1% | −176.5 | 18.7 | 115.1 | 32.7% | −107.0 | 2.6 | 86.5 | 44.1% | −72.1 |
| 2 year | 0% | −233.7 | 543.1 | 67.4% | −367.4 | −142.5 | 253.0 | 70.3% | −212.5 | −113.5 | 166.5 | 80.4% | −145.8 | −96.0 | 117.9 | 87.7% | −112.7 |
| | 10% | −198.7 | 543.8 | 54.3% | −417.2 | −124.0 | 253.4 | 61.6% | −223.0 | −102.1 | 166.8 | 71.0% | −153.2 | −89.3 | 118.0 | 83.3% | −111.7 |
| | 25% | −141.7 | 544.8 | 35.5% | −568.8 | −96.1 | 254.0 | 51.4% | −236.6 | −85.0 | 167.1 | 60.9% | −160.7 | −79.2 | 118.2 | 77.5% | −109.5 |
| | 50% | −37.9 | 546.4 | 23.2% | −744.5 | −49.3 | 254.8 | 36.2% | −281.5 | −56.9 | 167.6 | 49.3% | −167.5 | −62.2 | 118.5 | 63.8% | −114.4 |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10-year Treasury yield | | | | | | | | | | | | | | | | | |
| 3 month | 0% | −100.9 | 265.8 | 60.0% | −199.1 | −49.2 | 111.8 | 63.3% | −95.7 | −34.5 | 71.1 | 67.7% | −63.0 | −25.9 | 49.2 | 69.5% | −46.0 |
| | 10% | −69.8 | 265.8 | 41.1% | −253.4 | −32.7 | 111.8 | 49.1% | −104.8 | −24.3 | 71.1 | 55.7% | −65.3 | −19.9 | 49.2 | 62.3% | −45.0 |
| | 25% | −18.8 | 265.8 | 28.9% | −301.9 | −7.7 | 111.9 | 34.8% | −118.5 | −9.0 | 71.2 | 42.3% | −68.3 | −10.7 | 49.2 | 52.2% | −43.6 |
| | 50% | 74.9 | 265.8 | 18.2% | −361.2 | 34.7 | 111.9 | 23.4% | −124.8 | 16.5 | 71.2 | 27.4% | −74.0 | 4.8 | 49.2 | 35.1% | −46.2 |
| 2 year | 0% | −161.0 | 296.6 | 65.8% | −259.1 | −83.6 | 123.4 | 67.4% | −134.8 | −60.7 | 82.6 | 68.9% | −97.0 | −46.7 | 59.8 | 74.1% | −69.3 |
| | 10% | −124.6 | 295.8 | 50.3% | −296.8 | −64.4 | 123.3 | 57.0% | −138.2 | −48.8 | 82.7 | 62.7% | −94.2 | −39.6 | 59.9 | 67.4% | −68.9 |
| | 25% | −64.8 | 294.4 | 37.3% | −329.8 | −35.1 | 123.1 | 46.1% | −137.8 | −30.8 | 82.8 | 52.9% | −92.5 | −29.0 | 60.1 | 59.1% | −67.2 |
| | 50% | 44.6 | 291.9 | 26.4% | −340.0 | 14.4 | 122.9 | 32.1% | −137.8 | −1.1 | 83.0 | 40.4% | −86.3 | −10.9 | 60.4 | 50.8% | −58.9 |
| Deutsche mark | | | | | | | | | | | | | | | | | |
| 3 month | 0% | 2.4 | 95.6 | 39.0% | −79.6 | 6.2 | 45.5 | 38.8% | −35.8 | 2.7 | 27.2 | 41.3% | −21.5 | 0.4 | 16.8 | 45.0% | −13.0 |
| | 10% | 30.6 | 96.0 | 24.5% | −91.7 | 21.0 | 45.7 | 24.7% | −38.0 | 11.8 | 27.3 | 27.2% | −21.9 | 5.6 | 17.0 | 29.4% | −13.7 |
| | 25% | 76.2 | 97.3 | 14.3% | −99.4 | 43.3 | 46.2 | 13.4% | −39.0 | 25.4 | 27.7 | 14.0% | −23.7 | 13.7 | 17.3 | 15.6% | −15.1 |
| | 50% | 159.1 | 101.8 | 7.6% | −76.8 | 81.0 | 48.0 | 5.3% | −35.0 | 48.1 | 28.9 | 5.5% | −22.9 | 27.4 | 18.2 | 5.8% | −17.7 |
| 2 year | 0% | −2.6 | 67.9 | 59.0% | −40.8 | 0.6 | 36.8 | 57.5% | −22.1 | −2.7 | 25.4 | 56.1% | −18.0 | −4.7 | 15.7 | 64.6% | −12.4 |
| | 10% | 26.8 | 73.0 | 34.4% | −35.0 | 15.8 | 40.8 | 35.8% | −19.8 | 6.6 | 27.0 | 42.5% | −15.2 | 0.7 | 16.3 | 50.5% | −10.6 |
| | 25% | 75.1 | 86.6 | 13.7% | −31.5 | 39.2 | 48.6 | 17.5% | −15.3 | 20.8 | 30.6 | 18.9% | −15.5 | 9.1 | 18.1 | 28.3% | −9.8 |
| | 50% | 164.0 | 125.8 | 1.4% | −7.4 | 79.3 | 65.6 | 2.4% | −13.7 | 44.8 | 38.8 | 5.7% | −11.5 | 23.5 | 22.5 | 6.6% | −16.3 |

handsome profits. However, the standard deviations are generally larger than for calls and the average losses for those positions that do lose money could be quite severe.

## VI. Conclusion

The results we have obtained indicate that model risk in trading and hedging derivatives is quite large. Of the three risk management strategies discussed at the outset, our analysis makes clear that simply writing and holding options as if they are ordinary risk assets entails a very large risk exposure, with standard deviations of per trade returns well over 100 percent of premium income in many cases. Cash flow matching, if it can be done precisely, is obviously a sound risk management strategy because it removes the impact of model risk entirely, and market risk, as well. The entire risk in a cash-flow matched position is essentially credit risk—the risk that one of the counterparties in the matched position defaults on a losing trade and leaves the hedger exposed. Cash flow matching, however, is not a strategy that can be followed generally, as we have argued, because the nature of the derivatives business is that the public wants to be long options, so that the dealer community must be short on balance. Dealers as a group must hold exposed, unmatched, option positions.

This leaves delta hedging based on valuation models implemented with forecasted volatilities as the only viable trading and risk management strategy for most institutional options writers. Our results show that this can be expected to involve considerable risk exposure. Two important components of an overall risk management strategy for a derivatives book, therefore, should be use of the best pricing models and volatility estimators, and also diversification of (delta hedged) risk exposures across a variety of derivatives markets and instruments, with the hope that a "worst year" for one asset class may be mitigated by good years for others. A third component is simply to charge higher prices than the model indicates for the options that are written. Our results suggest that increasing the volatility input by one-quarter to one-half would substantially increase mean returns and reduce the fraction of losing trades. However, the worst losses to the strategy would remain very large—there would just be fewer of them.

## Appendix A. Data Sources

This appendix describes the sources and handling of the data series used in the study.

*Standard and Poors Stock Index:* The data for January 1971 to December 1996 come from CRSP. The daily dividend payout on the index portfolio is constructed as the difference between the dividend inclusive return series and the price change series. Monthly data are constructed from the daily series.

*90-day U.S. Dollar LIBOR:* For the period January 1971 to January 1975 the data are from the Harris Bank, and are weekly. From January 1975 to December 1996 the data are daily, from Datastream (series name: EURO-CURRENCY (LDN) US$ 3 MONTHS—MIDDLE RATE ECUSD3M).

*10-year Treasury Yield:* The source is the Constant Maturity 10-Year Yield series from the Federal Reserve H.15 report. Data come directly from the Federal Reserve for January 1971 to February 1996. The H.15 data are obtained from Bloomberg for February 1996 to December 1996.

*Deutsche mark/Dollar Exchange Rate:* Data for January 1971 to December 1995 are from the Federal Reserve H.10 report. From December 1995 to December 1996 the source is Datastream (series name: DMARKE$).

*Futures Prices:* Daily closing prices for the 90-day Eurodollar futures contract at the Chicago Mercantile Exchange and the 20-year U.S. Treasury Bond futures contract at the Chicago Board of Trade are obtained from the Futures Industry Association.

*U.S. Riskless Interest Rate:* The 90-day Eurodollar interest rate is used, after conversion to the equivalent continuously compounded rate.

*German Riskless Interest Rate:* The EuroDeutschemark three-month rate is used, after conversion to the equivalent, continuously compounded rate. For January 1971 to December 1980 the source is the Harris Bank. From January 1981 to December 1996 the data are from Datastream (series name: EURO-CURRENCY (LDN) D-MARK 3 MONTHS—MIDDLE RATE ECWGM3M).

## Appendix B. Calculation of Volatility Estimates

To explain the details of the basic OUCH estimation procedure used here, we focus on the single example of forecasting the volatility of three-month LIBOR over a 24-month horizon using 24 months of historical data, shown in Table I, Panel B. The data sample begins in January 1971. Since we need up to five years of past data to compute some of the estimates and we want to be able to compare the performance of alternative methods over the same historical time period, the first forecast date is January 1976. Continuously compounded changes for the observed interest rates from the last 24 months are computed by taking the first differences of their logarithms. Rather than calculating the standard deviation around the sample mean, which is generally a very inaccurate estimate of the true mean, we impose zero as the mean change in rates. The variance is then just the average of the squared log changes. This is annualized by multiplying the monthly variance by 12; then taking the square root gives the volatility. This estimate is used as the volatility forecast going forward from January 1976. The historical volatility over the past 24 months just calculated becomes the first forecast for both the 24-month horizon and the 60-month horizon.

Next we compute the realized volatility over the 24-month period from January 1976 to January 1978, also constraining the mean to zero. The difference between the forecast and the realized volatility becomes the forecast error for January 1976. The forecast date is then advanced one month and the procedure is repeated. This continues until December 1991, the last date for which our data sample allows calculation of the realized volatility over a 60-month forecast horizon. The series of volatility forecast errors are then squared and averaged, and the square root is taken to produce the root mean squared forecast error of 13.3 percent. This is the measure of forecast accuracy for the strategy of using 24 months of past data to predict future LIBOR monthly volatility over a 24-month horizon. The same procedure is used to fill in RMSE values for all combinations of market, historical sample, and forecast horizon.

We also consider an "All available" estimator that uses all past data back to the beginning of the sample as of each date. Unlike the other methods, the amount of historical data in this volatility estimate grows over time, from five years on the first date to 20 years or more on the last.

The third historical volatility estimator we examine is an exponentially weighted forecast, shown in equation (7). We want to estimate the decay rate $w$ for each date using only historical data that would have been available at the time, with the optimality criterion being out of sample RMSE. For each date, this requires dividing the available historical data into an estimation sample and a forecast sample. The procedure we have adopted is somewhat arbitrary, but represents one plausible way for such a computation to be done.

As an example, consider the 24-month forecast horizon using monthly observations. To allow multiple past forecast horizons, on date $T$ we compute realized volatility over 12 overlapping 24-month periods, $\{T - 25$ to $T - 1$, $T - 26$ to $T - 2, \ldots, T - 36$ to $T - 12\}$. Given a trial value for $w$, for each of those periods we compute a volatility forecast using equation (7) with all available data prior to the beginning of that period. This produces 12 forecasts and forecast errors (overlapping, unfortunately) from which the RMSE is calculated. We then search over values for $w$ to obtain the value that would have minimized RMSE. This $w$ is used on date $T$ to forecast volatility over $T + 1$ to $T + 24$. We then advance a month and repeat the process to obtain a weighted volatility forecast for the period $T + 1$ to $T + 25$ based on a reestimated $w$ value, and so on. In each case, we impose the constraint $w \leq 1.0$.

The difficulty arises at the beginning of the sample. To the extent possible, we would like to use an estimation sample of at least five years and a forecast sample allowing multiple periods of length equal to the forecast horizon we are trying to optimize for, so that a meaningful RMSE can be computed. That is not feasible in every case. To give a larger historical sample for the first forecast dates, we set $w = 1.0$ for the first year with the daily data and for the 24-month horizon, and for the first four years with the 60-month horizon. This still can leave only three years of sample data to compute the volatility for the first forecasts that go into the RMSE we are testing to find the optimal $w$. This problem only affects the beginning of the sample, but it is inherent in the use of an exponentially declining weight forecast when the weight must be obtained by examining past data.

## REFERENCES


Basle Committee on Banking Supervision, 1994, *Risk Management Guidelines for Derivatives* (Bank for International Settlements, Basle).

Basle Committee on Banking Supervision, 1996, *Amendment to the Capital Accord to Incorporate Market Risks* (Bank for International Settlements, Basle).

Black, Fischer, 1976, The pricing of commodity contracts, *Journal of Financial Economics* 3, 167–179.

Black, Fischer, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–659.

Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics* 31, 307–327.

Bollerslev, Tim, Ray T. Chou, and Kenneth Kroner, 1992, ARCH modeling in finance, *Journal of Econometrics* 52, 5–59.

Canina, Linda, and Stephen Figlewski, 1993, The informational content of implied volatility, *Review of Financial Studies* 6, 659–681.

Figlewski, Stephen, 1997, Forecasting volatility, *Financial Markets, Institutions & Instruments* 6, 2–87.

Figlewski, Stephen, 1998, Derivatives risks, old and new, *Brookings-Wharton Papers on Financial Services* 1, 159–217.

Galai, Dan, 1977, Tests of market efficiency of the Chicago Board Options Exchange, *Journal of Business* 50, 167–197.

Garman, Mark B, and Steven W. Kohlhagen, 1983, Foreign currency option values, *Journal of International Money and Finance* 2, 231–237.

Group of Thirty Global Derivatives Study Group, 1993, *Derivatives: Practices and Principles* (Group of Thirty, Washington).

Merton, Robert C., 1973, Theory of rational option pricing, *Bell Journal of Economics and Management Science* 4, 141–183.

Merton, Robert C., Myron S. Scholes, and Mathew L. Gladstein, 1978, The returns and risk of alternative call option portfolio investment strategies, *Journal of Business* 51, 183–242.

Merton, Robert C., Myron S. Scholes, and Mathew L. Gladstein, 1982, The returns and risks of alternative put-option portfolio investment strategies, *Journal of Business* 55, 1–56.

Risk Standards Working Group, 1996, *Risk Standards for Institutional Investment Managers and Institutional Investors*.

*RiskMetrics*, 1995, 3rd ed. (J.P. Morgan).

U.S. Government Accounting Office, 1994, *Financial Derivatives: Actions Needed to Protect the Financial System* (U.S. Printing Office, Washington).