

Testing the volatility term structure using option hedging criteria

March 1998

Robert F. Engle
Department of Economics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
(619) 534-1054

Joshua V. Rosenberg
Department of Finance
NYU - Stern School of Business
44 West 4th Street, Suite 9-190
New York, New York 10012-1126
(212) 998-0311

Abstract

The volatility term structure (VTS) reflects market expectations of asset volatility over different horizons. These expectations change over time, giving dynamic structure to the VTS. This paper evaluates volatility models on the basis of their performance in hedging option price changes due to shifts in the VTS. An innovative feature of the hedging approach is its increased sensitivity to several important forms of model misspecification relative to previous testing methods.

Volatility hedge parameters are derived for several volatility models incorporating different predicted VTS dynamics and information variables. Hedging tests using S&P500 index options indicate that the GARCH components with leverage VTS estimate is most accurate. Evidence is obtained for mean-reversion in volatility and correlation between VTS shifts and S&P500 returns. While a convexity hedge dominates the volatility hedges for the observed sample, this result appears to be due to sample selection bias.

This paper has benefited from the suggestions of Alex Kane, Bruce Lehmann, Allan Timmermann, Halbert White, Steve Figlewski, Peter Carr, Jose Lopez, Torben Andersen, Joel Hasbrouck, Kaushik Amin, participants at the 1998 American Finance Association meetings, and the 1997 Time Series Analysis of High Frequency Data Conference. This research was supported by the Project in Econometric Analysis.

The volatility term structure (VTS) reflects market expectations of asset return volatility over different time horizons. In a stochastic volatility setting, relative changes in delta-neutral medium and short-term straddle prices should be primarily attributable to shifts in the VTS. By deriving and implementing volatility hedges consistent with several volatility models, we are able to rank volatility models based on performance in hedging shifts in the VTS.

Heynen, Kemna, and Vorst (HKV, 1994) rank volatility models based on the closeness of the forecast VTS and the implied volatility term structure. Several other papers including Stein (1989), Diz and Finucane (1993), and Xu and Taylor (1994) directly estimate the volatility term structure using option implied volatilities.

Hedging tests provide several advantages over comparing a forecast VTS to a realized VTS. First, hedging tests focus on the relative levels of forecast variance over different horizons and are less sensitive to unconditional variance forecasts. So, hedging tests may be able to distinguish among alternative models with different term structure shapes but similar levels of unconditional variance. HKV (1994) find that differences in unconditional volatility forecasts drive their test results.

Second, hedging tests evaluate predicted term structure dynamics rather than the closeness of fit to the levels. In tests of VTS levels, it may be difficult to distinguish among models with similar rates of mean reversion but different explanatory variables. In other words, two volatility models with different information variables may perform similarly in pricing options, but quite differently in hedging options. Hedging tests may be superior at identifying omitted variables or interrelationships in the volatility model, because hedging performance depends on eliminating sensitivity to all of sources of volatility.

Hedging performance tests of the stochastic volatility models indicate that the GARCH components with leverage VTS estimate is most accurate. The hedging performance of this GARCH model is significantly better than the performance of all of the other stochastic volatility models, including two simpler GARCH models. Overall, the volatility hedges in which the VTS is sensitive to the magnitude of underlying asset returns outperform models in which the VTS is uncorrelated with returns. The least effective hedging performance is provided by a model with the greatest volatility persistence and no correlation between return magnitudes and volatility. This is the familiar delta-vega hedge.

It is also found that the constant volatility delta-gamma (convexity) hedge dominates the stochastic volatility hedges in the observed sample. This result is surprising, since it indicates that S&P500 volatility is either constant or extremely insensitive to news in the observed data sample. This is inconsistent with evidence for stochastic volatility based on the underlying price process and option implied volatilities. For example, HKV (1994) find that the VTS levels are best explained by an EGARCH model, and Amin and Ng (1994) find that GARCH models with asymmetric effects provide the best fit to individual equity option prices.

However, it is shown that the dominance of the convexity hedge over the volatility hedges may be due to sample selection bias and that the results might be reversed if all option price changes over the sample period could be observed. This bias does not appear to affect the relative performance among the volatility

hedges. It is suggested that the strongest results from the hedging tests are the relative rankings of the stochastic volatility hedges.

Previous papers concerned with option hedging performance are limited by their focus on a single type of stochastic volatility model and reliance on interpolated prices or option values at expiration for empirical hedging results. For instance, Hull and White (1987a) present simulation results that show that under a continuous-time integrated variance process, constant volatility (CV) delta-gamma hedging is more effective when implied volatility is relatively stable, and CV delta-vega hedging is more effective otherwise. The HW empirical hedging results depend on interpolating OTC option prices using Black-Scholes evaluated at the market implied volatility.

Melino and Turnbull (1995) present an option hedging technique under stochastic volatility using deltas and vegas obtained from the pricing partial differential equation. They find evidence that a constant volatility model performs poorly in pricing and hedging long-term currency options under stochastic volatility. Model performance is measured using the replicating error of a hedge portfolio held until the target option expiration.

This paper provides an empirical comparison of the efficacy of techniques for hedging medium-term S&P500 index at-the-money straddles with short-term at-the-money straddles. All hedge portfolio price changes are based on market prices rather than interpolated prices. The hedge ratios derived in this paper are only applicable to at-the-money options and may not be directly used to hedge an options book which has options with a variety of moneynesses. Engle and Rosenberg (1995) provide an alternative GARCH option hedging methodology, based on Monte-Carlo simulation, which is applicable to options of any moneyness.

This paper is structured as follows. Section 1 presents an approximate option pricing formula applicable to at-the-money options under stochastic volatility. Section 2 details the derivation of stochastic volatility hedging parameters. Section 3 describes the estimation of the S&P500 volatility models and volatility hedge ratios. Section 4 presents the accuracy of the approximate BSP hedge ratios. In Section 5, five volatility models are tested based on their ability to hedge the effects of volatility changes on medium-term S&P500 index straddles. Section 6 investigates sample selection bias and the effect on volatility versus convexity hedges. Section 7 concludes the study.

1. Approximate option pricing in a stochastic volatility environment

The Black-Scholes (1973) option pricing model assumes that the underlying asset variance is constant over the life of the option. Substantial empirical evidence has been presented, see e.g. Bollerslev, Chou, and Kroner (1992), that stochastic variance characterizes many financial return time-series. A number of option pricing models, such as Hull and White (1987b), Melino and Turnbull (1990, 1995), Amin and Ng (1993, 1994), and Duan (1995) have been developed that allow for stochastic volatility. However, each model provides a different specification for the variance process. The method for selecting the appropriate variance process and thus the correct model for option pricing and hedging is left open.

Since pricing options in a generic stochastic volatility environment is not a solved problem, we utilize an approximate option pricing formula for *at-the-money* options that may be applied to a variety of volatility models. The use of a single approximate option pricing formula for different volatility models facilitates the derivation of the option hedge parameters using a consistent methodology.

This paper uses the following approximate pricing formula, which will be referred to as the Black-Scholes-plug-in formula or BSP.

$$(1) \quad P_t \cong BSP(E_t[\bar{\sigma}_{t,T}(S_t)], S_t, T) \quad C_t \cong BSP(E_t[\bar{\sigma}_{t,T}(S_t)], S_t, T)$$

In equation (1), BSP is the Black-Scholes pricing formula for a put or call. P_t and C_t are the call and put premia, S_t is the current underlying asset price, and T is the number of days until option maturity. Dependence on the riskless rate and strike price are suppressed. In this approximate pricing formula, expected average volatility $\bar{\sigma}_{t,T}(S_t)$ is “plugged into” the Black-Scholes formula to obtain the stochastic volatility option price.

The average expected volatility over the next $T-t$ days, $\bar{\sigma}_{t,T}(S_t)$, is a single point on the volatility term structure. The entire volatility term structure (VTS) at date t is defined by the relationship between average expected volatility and the forecast horizon. The VTS depends on state variables describing current volatility, unconditional volatility, and the rate of volatility persistence. Average expected volatility may also depend on the most recent return magnitude which is a function of both S_t and S_{t-1} . For simplicity, dependence on S_{t-1} is suppressed in equation (1).

The accuracy of the BSP formula depends on several factors. First, the options to be priced should be at-the-money, since the BSP method relies on the approximate linearity of the Black-Scholes formula in the volatility parameter for at-the-money options. Second, the effect of volatility risk premia must be small, since average expected volatility rather than average volatility under the risk-neutral measure is used. Third, approximate conditional log-normality under the risk-neutral measure is necessary. The accuracy of hedge ratios derived under BSP will be verified using simulations under the appropriate risk-neutral measure in Section 4.

2. Hedging options in a stochastic volatility environment

Hedge parameters measure the sensitivity of an option price to changes in the state variables. When a day passes, the option price will change in part because the underlying asset price changes and in part because the volatility forecast changes. Changes due solely to the passage of time or changes in interest rates will be ignored. Thus, the approximate hedge parameters developed in this section are appropriate for hedging over a short time period, such as one day, and may not perform well for hedging over a longer horizon.

The change in the option value due to changes in the state variables may be approximated using a Taylor series expansion. In this case, it is natural to think of expanding the end-of-day option price as a function of the end-of-day state variables. Just as Black-Scholes delta and gamma are derived by taking the first-derivative of the Black-Scholes formula with respect to the end-of-day underlying price (S_t), the BSP delta and gamma are obtained by taking derivatives of BSP with respect to S_t . If volatility (σ) is considered to be a separate stochastic state variable, a volatility hedge parameter may be obtained using a partial derivative of BSP with respect to volatility.

Evaluating the derivatives of BSP under constant volatility (CV) at current values of the state variables gives the familiar Black-Scholes delta, gamma, and vega hedge parameters. Typically, these hedge parameters are used to hedge option price changes in response to the first and second-order effect of changes in the underlying asset price and the first-order effects of changes in the underlying asset variance.

$$(2) \quad \Delta_{CV} = \frac{\partial P_t}{\partial S_t} \quad \Gamma_{CV} = \frac{\partial^2 P_t}{\partial S_t^2} \quad \Lambda_{CV} = \frac{\partial P_t}{\partial \sigma}$$

The chain rule may be used to develop BSP hedge parameters in a stochastic volatility environment. The potential dependence of expected average volatility on the underlying price, return, and the parameters of the volatility process suggests that the BSP hedge parameters will incorporate additional terms. In fact, the stochastic volatility hedge parameters will be combinations of CV delta, CV gamma, CV vega, and derivatives of the volatility term structure defined by the volatility process parameters.

For example, BSP delta is obtained by differentiating equation (1) with respect to the underlying price. BSP delta measures the option price change due to a small (first-order) change in the current underlying price at the current level of volatility.

$$(3) \quad \Delta = \Delta_{CV} + \Lambda_{CV} * VM \quad VM = \frac{\partial \bar{\sigma}_{t,T}}{\partial S_t} \quad \bar{\sigma}_{t,T}^2 = \frac{1}{T} E_t \left[\sum_{i=1}^T \sigma_{t+i}^2 \right]$$

The delta formula indicates that a change in the underlying price affects the option price directly through CV delta and indirectly through CV vega and a shift in the volatility term structure. The vega multiplier (VM) in equation (3) measures the change in the average expected volatility, $\bar{\sigma}_{t,T}$, due to a first-order change in the current underlying price.

The BSP deltas are fairly similar for the volatility models considered in this paper, since their vega multipliers are all zero. The vega multiplier is zero when a first-order change in the underlying price does not affect average expected volatility. In the CV and ARIV models, the underlying price change conveys no volatility news. In GARCH models, volatility changes due to underlying price changes including the leverage effect are captured by second and higher order terms. Setting the vega multiplier equal to zero simplifies equation (3) to CV delta evaluated at the forecast average volatility.

While deltas will only vary based on the average volatilities predicted by different models, the volatility hedge parameters are potentially quite different across stochastic volatility environments. This means that hedging tests that use the volatility hedge parameters should be able to distinguish the relative accuracy of the volatility models. So, delta-vega or delta-gamma hedging tests are preferred to delta hedging tests for ranking volatility models.

2.1 CV model volatility hedge parameters

While the CV model is not a stochastic volatility model, one might consider the effects on the option price of continual updating of volatility estimates based on an investor's expanding information set. If volatility is constant but is estimated with error, then updating of volatility estimates will result in changes in the option price.

The type of updating information used motivates two distinct types of CV volatility hedges. In the first case, volatility news might be viewed as independent of the current underlying price. Then, the volatility hedge parameter is CV vega as given in equation (2), which measures the option price change due to a change in σ . CV vega is increasing with time-to-maturity so that longer maturity options are more sensitive to volatility changes than short maturity options. Volatility news causes a parallel shift in the term structure, since an unconditional volatility estimate revision affects all average variance forecasts equally.

If the unconditional volatility is estimated using historical returns, then the estimated volatility will be a function of the current underlying price and return. Using a long historical time-series, the volatility effect of a change in the current underlying price will typically be negligible. Thus, the volatility hedge parameter is approximately equal to CV gamma as given in equation (2). CV gamma measures the second-order effect of a change in the underlying price on the option price due to the convexity of the option pricing formula, which dominates the volatility effect in this context. CV gamma is decreasing with time-to-maturity.

CV vega and CV gamma provide opposite forecasts of the relative changes in short and medium-term options due to volatility news. These contrasting results are based on the hypothesized relationship between return magnitudes and volatility. In either case, the CV volatility term structure is flat, since average volatilities over all horizons are equal to the estimated unconditional variance.

2.2 ARIV model volatility hedge parameters

In this paper, the autoregressive implied volatility model (ARIV) is based on the AR(1) specification given by Heynen, Kemna, and Vorst (1994). The ARIV model may be viewed as a reduced form of a stochastic autoregressive volatility model (SARV) model in which the factors that drive the volatility process are unobservable, but volatility is observable. For further discussion of SARV models, see Andersen (1994) or Taylor (1994).

For the ARIV model, the volatility hedge parameter is ARIV vega. This is obtained by differentiating the BSP formula with respect to the volatility news using the chain rule. The first term is the derivative of

BSP with respect to volatility, which is CV vega. This term is multiplied by the derivative of average volatility with respect to the volatility news, which is the vega multiplier (VM). ARIV vega is evaluated using current information including the average expected volatility given by the ARIV model.

$$(4) \quad \Lambda_{ARIV} = \Lambda_{CV} * VM \quad VM = \frac{\sqrt{\bar{s}_{t,T}}}{\sqrt{s_{t+1}^2}}$$

Consider a first-order autoregressive volatility model in which σ_t is the volatility on day t, and σ_{t-1} is the one-day lag. The ARIV model may be written in variances as:

$$(5) \quad s_t^2 = w + r s_{t-1}^2 + e_t$$

Using equation (5), the ARIV model volatility term structure (in variances) and vega multiplier are defined by:

$$(6) \quad \bar{s}_{t,T}^2 = \bar{s}^2 + \frac{1}{T} \left[\frac{1-r^T}{1-r} \right] (s_{t+1}^2 - \bar{s}^2), \quad \bar{s}^2 = \frac{w}{1-r}$$

$$VM = \frac{\sqrt{\bar{s}_{t,T}}}{\sqrt{s_{t+1}^2}} = \frac{1}{2\bar{s}_{t,T}T} \left[\frac{1-r^T}{1-r} \right]$$

A mean-reverting ARIV model has the property that volatility news has the greatest effect on the one-day ahead volatility, and the effect on future volatilities decays with time. If the volatility process were integrated, the volatility news would affect all future volatilities equally. While the VTS at any given time might be upward or downward sloping, volatility news will always decay at a rate defined by the autoregressive parameter of the ARIV process. This parameter will be less than one if the process is mean-reverting, and it determines the VTS curvature.

2.3 GARCH model volatility hedge parameters

Consider hedging changes in volatility when volatility forecasts depend on the magnitude of the current underlying asset return such as in the GARCH models. The volatility hedge parameter for GARCH models is based on the second derivative of BSP with respect to the underlying price, which captures the impact of volatility news.

As with the GARCH delta formula, the GARCH gamma formula is derived by eliminating the terms in the vega multiplier that are zero when evaluated in the default case of no volatility news. GARCH gamma is then estimated using current information including the realized return for date t, the current underlying price, and the GARCH average expected volatility.

$$(7) \quad \Gamma_G = \Gamma_{CV} + \Lambda_{CV} * VM \quad VM = \frac{\mathbb{1} \bar{\mathbf{s}}_{t,T}}{\mathbb{1} \mathbf{s}_t^2}$$

The GARCH gamma in equation (7) incorporates both a volatility hedge and a hedge against non-linear price response, since it includes both CV vega and CV gamma. The vega multiplier (VM) measures the second-order shift in the volatility term structure due to a change in the underlying price. The VM incorporates all the parameters of the GARCH process including leverage terms.

The VM determines the weights on CV gamma and CV vega in GARCH gamma. If variances are insensitive to second-order changes in the underlying price, then the VM will be close to zero and the weights will give more emphasis to the CV gamma. In the extreme case that volatility is constant, the vega multiplier will be zero and GARCH gamma will simplify to CV gamma. If the volatility process is integrated, then more weight will be given to BS vega. Notice that GARCH gamma is always at least as large as CV gamma, since GARCH gamma adds the volatility effect of a price shock to the non-linear effect of a price shock.

In contrast to the ARIV model, the GARCH models use historical data from the underlying asset to estimate the volatility process and the volatility term structure. Bollerslev's (1986) GARCH(p,q) model was developed as an extension to the Engle's (1982) ARCH(p) formulation to incorporate volatility persistence in a more parsimonious manner. In high frequency data, the GARCH-in-mean model developed by Engle, Lilien, and Robins (1987) which allows for a time varying risk-premium is frequently rejected. We find this to be the case for daily S&P500 returns, so a constant risk premium is used in all of the GARCH models.

The GARCH(1,1) model with a constant risk premium may be written as:

$$(8) \quad \ln(S_t / S_{t-1}) - r = \mu - \frac{1}{2} \mathbf{s}_t^2 + \mathbf{e}_t, \quad \mathbf{e}_t \sim N(0, \mathbf{s}_t^2)$$

$$(9) \quad \mathbf{s}_t^2 = \mathbf{w} + \mathbf{a} \mathbf{e}_{t-1}^2 + \mathbf{b} \mathbf{s}_{t-1}^2$$

where $\ln(S_t/S_{t-1})-r$ is the excess log-return, μ is a constant risk premium, \mathbf{s}_t^2 is the one-day conditional variance, and r is the risk-free rate. The parameters \mathbf{a} and \mathbf{b} determine the relative importance lagged squared return and lagged conditional variance on current conditional variance.

The GARCH(1,1) volatility term structure (in variances) and vega multiplier are given by:

$$(10) \quad \bar{\mathbf{s}}_{t,T}^2 = \bar{\mathbf{s}}^2 + \frac{1}{T} \left(\frac{1 - (\mathbf{a} + \mathbf{b})^T}{1 - (\mathbf{a} + \mathbf{b})} \right) (\mathbf{s}_{t+1}^2 - \bar{\mathbf{s}}^2), \quad \bar{\mathbf{s}}^2 = \frac{\mathbf{w}}{1 - (\mathbf{a} + \mathbf{b})}$$

$$(11) \quad VM(T) = \frac{\mathbf{a}}{T \bar{\mathbf{s}}_{t,T} \mathbf{s}_{t-1}^2} \left(\frac{1 - (\mathbf{a} + \mathbf{b})^T}{1 - (\mathbf{a} + \mathbf{b})} \right)$$

The GJR model (Glosten, Jagannathan, and Runkle, 1993) or TARARCH model (Zakoian, 1994) generalizes the GARCH(1,1) model to allow negative return shocks to disproportionately increase volatility. This asymmetric effect is frequently called leverage, reflecting the increase in the debt-equity ratio that follows a reduction in a firm's market capitalization. In this paper, the GJR model with a constant risk premium is defined by:

$$(12) \quad \ln(S_t / S_{t-1}) - r = \mathbf{m} - \frac{1}{2} \mathbf{s}_t^2 + \mathbf{e}_t, \quad \mathbf{e}_t \sim N(0, \mathbf{s}_t^2)$$

$$(13) \quad \mathbf{s}_t^2 = \mathbf{w} + \mathbf{a}\mathbf{e}_{t-1}^2 + \mathbf{g}[\text{Max}[0, -\mathbf{e}_{t-1}]^2] + \mathbf{b}\mathbf{s}_{t-1}^2$$

The leverage effect is measured by γ which determines the effect of a negative return shock on conditional variance. The α and β terms have the same interpretations as in the GARCH (1,1) model.

In this model specification, the second derivative of volatility with respect to the current underlying price is nonexistent. From the left, the second derivative is zero, and from the right it is γ , so $.5\gamma$ is used as an approximation. The GJR volatility term structure (in variances) and vega multiplier are given by:

$$(14) \quad \bar{\mathbf{s}}_{i,T}^2 = \bar{\mathbf{s}}^2 + \frac{1}{T} \left(\frac{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})^T}{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})} \right) (\mathbf{s}_{i+1}^2 - \bar{\mathbf{s}}^2), \quad \bar{\mathbf{s}}^2 = \frac{\mathbf{w}}{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})}$$

$$(15) \quad VM(T) = \frac{\mathbf{a} + 5\mathbf{g}}{T \bar{\mathbf{s}}_{i,T}^2 \mathbf{s}_{i-1}^2} \left(\frac{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})^T}{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})} \right)$$

The GARCH components with leverage model (GCOMP) developed by Engle and Lee (1993) allows for richer volatility dynamics and a leverage effect. In this model, volatility shocks have different effects on a long-run and short-run volatility component, and each volatility component has a different rate of mean reversion. Consider the following GCOMP formulation:

$$(16) \quad \ln(S_t / S_{t-1}) - r = \mathbf{m} + \mathbf{e}_t, \quad \mathbf{e}_t \sim N(0, \mathbf{s}_t^2)$$

$$(17) \quad \mathbf{s}_t^2 = \mathbf{q}_t^2 + \mathbf{a}(\mathbf{e}_{t-1}^2 - \mathbf{q}_{t-1}^2) + \mathbf{g}[\text{Max}[0, -\mathbf{e}_{t-1}]^2 - 5\mathbf{q}_{t-1}^2] + \mathbf{b}(\mathbf{s}_{t-1}^2 - \mathbf{q}_{t-1}^2)$$

$$(18) \quad \mathbf{q}_t^2 = \mathbf{w} + \mathbf{r}\mathbf{q}_{t-1}^2 + \mathbf{f}(\mathbf{e}_{t-1}^2 - \mathbf{s}_{t-1}^2)$$

Of the parameters, α reflects the effect of a shock on the short-run component of volatility, γ reflects the short-run asymmetric effect of "bad news" on volatility, β reflects the influence of the lagged conditional volatility, ρ reflects the persistence of the long-run component, and ϕ reflects the effect of a volatility shock on the long-run component.

The GCOMP volatility term structure (in variances) and vega multiplier are given by:

$$(19) \quad \bar{s}_{i,T}^2 = \bar{s}^2 + \frac{1}{T} \left(\frac{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})^T}{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})} \right) (s_{i+1}^2 - q_{i+1}^2) + \frac{1}{T} \left(\frac{1 - \mathbf{r}^T}{1 - \mathbf{r}} \right) (q_{i+1}^2 - \bar{s}^2)$$

$$\bar{s}^2 = \frac{w}{1 - \mathbf{r}}$$

$$(20) \quad VM(T) = \frac{1}{T \bar{s}_{i,T} s_{i-1}^2} \left((\mathbf{a} + 5\mathbf{g}) \frac{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})^T}{1 - (\mathbf{a} + \mathbf{b} + 5\mathbf{g})} + \mathbf{f} \frac{1 - \mathbf{r}^T}{1 - \mathbf{r}} \right)$$

The GARCH(1,1) and GJR models, like the ARIV model, generate a monotonic upward or downward sloping VTS with the curvature determined by the level of volatility persistence. The GCOMP model is compatible with a non-monotonic VTS. In the GCOMP model, the VTS curvature is dominated by the sum of α , β , and $.5\gamma$ over the short-run and by ρ over the long-run. Notice that S_{t-1} appears in the denominator of all of the GARCH vega multipliers as a result of taking the second derivative with respect to S_t of the current squared return in the volatility equation, which is a function of S_{t-1} .

Figure 1 illustrates the differences in several volatility hedge parameters evaluated for at-the-money straddles with 1 to 100 days until maturity using estimated parameters for S&P500 return volatility models. Average expected volatilities are set to one-percent per day, the riskless rate and dividend yield are set to zero, and the strike and underlying price are set to 100. CV vega is increasing in time-to-maturity, while CV gamma and GCOMP gamma are both declining in time-to-maturity. ARIV vega is increasing initially, but then decreasing in time-to-maturity.

3. Estimating volatility hedge ratios for S&P500 index options

In order to implement hedging tests, it is necessary to estimate the volatility hedge ratios which predict the relative price change for a medium-term option position compared to a short-term option position due to volatility news. The volatility hedge ratios depend on the estimated S&P500 return volatility model parameters as well as the characteristics of the option position. Using the methods described in this section, the GARCH, ARIV or CV hedge ratios could be estimated for at-the-money European options traded on other indices, commodities, or individual stocks.

First, consider the estimated CV model presented in Table 2. The annualized sample standard deviation of log returns (σ) is 15.47%. The sample consists of S&P500 daily log-returns from June 20, 1982 to June 30, 1996, which includes the 1000 days prior to the first day on which hedging tests begin and ends on the last test day. The CV model sample standard deviation is also estimated for out-of-sample hedging tests using a daily expanding window of trailing returns.

The CV gamma and CV vega hedge ratios are:

$$(21) \quad \frac{\Gamma_{CV}(\mathbf{s}, S_t, T_m)}{\Gamma_{CV}(\mathbf{s}, S_t, T_s)} \quad \frac{\Lambda_{CV}(\mathbf{s}, S_t, T_m)}{\Lambda_{CV}(\mathbf{s}, S_t, T_s)}$$

Table 1 describes additional characteristics of the S&P500 log-return time-series. The negative return skewness and positive excess kurtosis suggest that the unconditional distribution is poorly approximated by a log-normal distribution. Persistence in the squared log-returns, as evidenced by the significant Ljung-Box statistic of 234.83, suggests time-varying volatility.

The estimated ARIV model, presented in Table 2, is mean-reverting with an estimated autoregressive parameter of .91. The ARIV process is estimated following the methodology of HKV (1994) and Stein (1989). In these papers, the short-term implied variance is taken as a proxy for the one-day implied variance, and the volatility process is estimated by regression of the short-term implied variance on its first lag. For out-of-sample hedging tests, an ARIV model is estimated daily using an expanding window of short-term implied variances.

Implied S&P500 return variances are obtained by inverting the Black-Scholes formula using the closest-to-the-money, nearest maturity S&P500 call and put options with at least five trading days left until maturity. The 1296 days over the period from January 1986 to June 1996 on which both short and medium-term straddle price changes are available are used to extract the implied variances and their first lags. The Black-Scholes formula includes a dividend adjustment in which the current index price is discounted by the present value of dividends paid over the life of the option. Realized daily index dividends are estimated using the total and capital appreciation returns reported in the CSP stock file index database. The risk-free rate used is the 13-week Treasury Bill yield.

The ARIV volatility hedge ratio is:

$$(22) \quad \frac{\Lambda_m(\bar{S}_{t,T_m}, S_t, T_m)VM(T_m)}{\Lambda_s(\bar{S}_{t,T_s}, S_t, T_s)VM(T_s)}$$

The medium and short-term average volatilities used in equation (22) are the implied standard deviations for the medium and short-term S&P500 index options. $VM(T_m)$ and $VM(T_s)$ are the medium and short-term vega multipliers, which are derivatives of volatility term structure with respect to volatility news.

Consistent with a mean-reverting ARIV model, changes in short-term implied standard deviations are more volatile than changes in medium-term implied standard deviations as shown in Table 1. Interestingly, the mean and standard deviation of medium-term implied variances are slightly higher than the mean and standard deviation of short-term implied variances. The large excess kurtosis of the changes in short and medium-term volatility indicate the presence of large but infrequent implied volatility jumps.

The three GARCH models are presented in Table 2. These models are estimated by maximum likelihood using S&P500 daily log-excess returns from June 20, 1982 to June 30, 1996. Volatility persistence in the GARCH models is greater than in the ARIV model. For example, the GARCH(1,1) model has volatility persistence as measured by the sum of α and β of about .99 compared to .91 for the ARIV model. The GJR volatility persistence is about .98. The GCOMP model has higher long-run, but

lower short-run volatility persistence than the other models. The leverage effect is significant in the GJR and GCOMP models using standard t-statistics, but insignificant using the robust t-statistic of Bollerslev and Wooldridge (1992). All of the models are successful in eliminating persistence in the standardized squared returns as measured by the Ljung-Box statistic.

For out-of-sample hedging tests, the GARCH models are re-estimated daily using an expanding window of trailing returns. In this case, the first estimation begins with return data from June 20, 1982 through January 1, 1986 for estimation on January 1, 1986. For estimation, excess returns are calculated using the 13-week Treasury Bill yield as a proxy for the riskless rate of interest, and returns during the week of the October 1987 crash are down-weighted by twenty percent in the log-likelihood function.

The GARCH volatility hedge ratio is:

$$(23) \quad \frac{\Gamma_m(\bar{S}_{t,T_m}, S_t, T_m) + \Lambda_m(\bar{S}_{t,T_m}, S_t, T_m)VM(T_m)}{\Gamma_s(\bar{S}_{t,T_s}, S_t, T_s) + \Lambda_s(\bar{S}_{t,T_s}, S_t, T_s)VM(T_s)}$$

To better understand the volatility hedge ratios in equations (21), (22), and (23), consider hedging the volatility sensitivity of a medium-term at-the-money straddle with 30 days to maturity with a short-term at-the-money straddle with 10 days to maturity. Straddles are a natural position to consider for volatility hedging, since straddles are especially sensitive to changes in volatility and insensitive to the direction of underlying price movements.

Table 3 indicates that a CV vega hedge requires the purchase of 1.73 short-term straddles for each medium-term straddle written. A hedge ratio greater than one reflects the increase in vega with maturity. This corresponds to an experiment where volatility is changed once and for all, and therefore has a larger impact on longer-lived options. Figure 1 confirms that CV vega is increasing as the option time-to-maturity increases. In contrast, the CV gamma hedge requires the purchase of .58 short-term straddles per medium-term straddle written.

In the ARIV model, the ARIV volatility hedge ratio incorporates mean reversion in volatility, which counteracts the increase in vega with time-to-maturity. For this example and the estimated S&P500 ARIV model, an ARIV volatility hedge requires the purchase of .91 short-term straddles for every medium-term straddle written. Figure 1 shows that ARIV vega declines with time-to-maturity for straddles with greater than two weeks until maturity resulting in a hedge ratio less than one.

In this example, the GARCH(1,1) volatility hedge ratio is .95, the GJR hedge ratio is .92, and the GCOMP hedge ratio is .76. The volatility hedge ratios are largest for GARCH models with greatest volatility persistence, since increased volatility persistence increases sensitivity of medium-term options to volatility news.

Using the estimated volatility models, it is straightforward to calculate BSP volatility hedge ratios over the sample period. This gives a another picture of the model differences. For hedging at-the-money medium-term S&P500 straddles with short-term straddles, all average hedge ratios are less than one,

except for the CV vega hedge ratio. As shown in Table 3, the CV gamma hedge and the GCOMP gamma hedge have the lowest average hedge ratios.

4. Evaluating the accuracy of the approximate BSP hedge ratios

The final step before utilizing the BSP hedge ratios in hedging tests is an evaluation of the accuracy of the BSP hedge ratios as approximations of the true hedge ratios. Hedging error due to use of approximate hedge ratios will be proportional to the difference between the true and approximated hedge ratios. Thus, accuracy of the approximate hedge ratios is properly evaluated on an absolute rather than relative basis.

Using Monte-Carlo simulation and the appropriate risk-neutralized processes for the ARIV and GARCH models, the “true” hedge ratios may be calculated under the risk-neutral measure and compared with the BSP approximations. Centered finite difference approximations of option pricing formula derivatives are used to obtain the hedge parameters. The centered finite difference approximations for GARCH delta and GARCH gamma are described in Engle and Rosenberg (1995). For this study, simulation under Amin and Ng’s (1994) risk-neutralization for GARCH processes is used with the estimated S&P500 volatility process parameters, and derivatives are evaluated at a one-tenth standard deviation shock centered around the expected asset return.

ARIV delta and ARIV vega are calculated using simulation under the Hull and White (1987b) risk-neutralization using the estimated ARIV parameters for S&P500 return volatility. The centered finite difference approximation for ARIV vega is calculated by first taking the difference between a simulated straddle price evaluated at the initial volatility plus and minus a one-tenth standard deviation shock. The simulated option price difference is divided by the difference in the initial volatilities, giving an estimate of ARIV vega. ARIV delta is calculated in an analogous manner using one-tenth standard deviation changes in the initial underlying price.

All true hedge parameters for at-the-money calls and puts are calculated with 50,000 simulation replications for maturities from 5 to 90 days and moneynesses of -.010, .005, 0.000, .005, and .010 representing the range of option maturities and moneynesses in the empirical hedging tests. Simulated hedge parameters are smoothed according to the methodology in Engle and Rosenberg (1995). In each simulation experiment, the level of volatility is set to the unconditional volatility for the given process. The gamma and vega hedge ratios are based on hedging an at-the-money medium-term straddle with 25 to 90 days left with a short-term straddle with 20 fewer days until maturity. The delta hedge ratios are generated for straddles with 5 to 90 days until maturity.

The “true” hedge ratios are compared with the BSP hedge ratios evaluated at the unconditional level of volatility for the original process using equations (3), (21), (22), and (23). The difference between the BSP hedge ratio and the fully-simulated alternative is defined to be the approximation error. The hedge ratio errors are aggregated over maturities and moneynesses to evaluate accuracy and bias for the entire sample.

We expect that the particular method for delta hedge ratio calculation will have little effect on the hedging tests for two reasons. First, Table 4 shows that average at-the-money simulated straddle deltas

across maturities are close to zero, ranging from .0255 to .0647. This indicates that changes in the underlying asset price have a relatively small contribution to straddle price variance. Empirically, a delta hedge provides a negligible amount of hedging benefit for a straddle because of the insensitivity of this position to the direction of underlying price changes. For example, a CV delta hedge for a medium-term straddle results in a reduction of only 1.1% in portfolio variance as shown in Table 6.

Second, Table 4 also shows that the BSP at-the-money straddle deltas provide a reasonable approximation to the simulated deltas. For example, the average errors for the GARCH deltas range from .0081 to .0372, while the average ARIV delta error is -.0013. The presence of a leverage effect does impart a slight downward bias in the GJR and GCOMP deltas evaluated using BSP.

The ARIV and GARCH volatility hedge ratio approximation errors are also acceptably small, as shown in Table 4. The average GARCH volatility hedge ratio error ranges from -.0090 to -.0128 while the average ARIV error is -.0179. These results suggest that the contribution of approximation error to the hedging results is not significant.

5. Hedging tests

In this section, volatility hedging tests are developed for a \$100 at-the-money medium-term S&P500 index option straddle positions using short-term S&P500 index straddles. Each straddle position consists of an equal number of calls and puts with identical strike prices and times to maturity. The positions are also delta hedged to minimize the influence of correlation with the underlying asset on the tests. The hedging tests are conducted using daily closing option price data from the Chicago Board Options Exchange over the period January 1986 through June 1996.

Table 5 summarizes the data used in the option hedging tests. There are 1296 days over the 10.5 year sample period for which price changes for all four options are available out of a total of 2654 trading days. The short-term straddle is selected as the nearest-maturity closest-to-the-money position with at least five trading days until maturity with a moneyness (index price/strike price - 1) less than .01. The medium-term straddle selected is the next maturity straddle of the same moneyness as the short-term straddle with available closing price change data. The average moneyness of the straddles is close to zero indicating the positions are on average almost exactly at-the-money. The average time-to-maturity of the short-term and medium-term straddles are about 14 and 38 trading days respectively.

Hedge ratios are estimated using the delta hedge ratio in equation (3) and the volatility hedge ratios in equations (21), (22), and (23). Hedge parameters for each model are evaluated using the BSP formulas given in section (2). In the BSP formula, the underlying price used is the closing S&P500 index price, and the risk-free rate used is the 13-week Treasury Bill yield. The BSP formula is adjusted for realized index dividends by discounting the index price by the present value of dividends paid over the life of the option. The S&P500 daily dividend series is derived from capital appreciation and total returns reported in the CRSP stock index file database.

Two sets of hedge ratios are estimated. The first set of “in-sample” hedge ratios are based on the estimated volatility models reported in Table 2. These hedge ratios correspond to a test of the model that best characterizes the volatility term structure over the sample period. A second set of “out-of-sample” hedge ratios are based on volatility models estimated using an expanding window of data. In this case, only data that would have been available on each hedging date is used to construct the hedge portfolio.

The hedging tests are implemented as follows. Each trading day, a medium-term at-the-money straddle position worth \$100 is written. The volatility and delta hedge ratios corresponding to the volatility models are then calculated. The number of short-term straddles to purchase is given by the volatility hedge ratio times the number of medium-term straddles written. The number of shares of the underlying to purchase or sell per medium-term straddle is given by the delta hedge ratio. These transactions are made, creating the volatility hedge portfolio. This hedge portfolio is held for one day and then sold. Each day, hedge portfolio price changes are calculated, and new positions established.

Hedging effectiveness is analyzed in Table 6. In continuous-time hedging, the optimal hedge portfolio would have zero variance, but in discrete-time hedging the optimal hedge portfolio will exhibit volatility. Thus, several hedging effectiveness criteria are reported to ensure that multiple dimensions of risk that arise in discrete-time hedging are considered. We also examined hedging performance using unexpected price changes, that is changes net of theta, and the outcomes were quite similar. We do not expect that sensitivity to interest rate changes will have a substantial impact on test results.

Since the hedging performance results based on multiple criteria are quite similar, this discussion will focus on two particular measures: variance reduction and absolute hedge error. The “variance reduction” reported is defined as the percent reduction in variance of the hedged portfolio relative to the unhedged portfolio. The absolute hedge error and associated t-statistic are used to compare the relative effectiveness of two selected hedging methods. For example, the “absolute hedge error of model A versus model B” is the average daily difference in the centered absolute hedge portfolio price changes for models A and B. A positive value for this measure implies that model A has larger average absolute hedging errors, i.e. inferior performance relative to model B.

The statistical significance of hedging performance differences for two models is evaluated using White’s (1980) heteroskedasticity consistent t-statistic for the absolute hedge error. In this case, this t-statistic is the same as the conventional t-statistic. A positive t-statistic greater than two indicates that the model A is inferior to model B, while a negative t-statistic less than two indicates the reverse.

The results in Table 6 indicate that the CV delta-gamma (convexity) hedge offers the greatest reduction in hedge portfolio variance (27.7%). The second greatest reduction in hedge portfolio variance is generated by the GCOMP delta-gamma hedge (24.5%). These rankings are unchanged when the entire sample is divided into two sub-periods: 1986-1991 and 1992-1996.

The absolute hedge error t-statistic is used to test whether the CV delta-gamma hedge (or the GCOMP delta-gamma hedge) is significantly better than the alternatives. The fifth column of the second panel compares absolute hedge errors of alternative models (model A) with the CV delta-gamma hedge (model B). For all alternative models, the average absolute hedge error differences are positive indicating these

models are inferior to the CV delta-gamma hedge. For example, the average daily absolute hedge error for the ARIV delta-vega hedge is \$.25 larger than the average absolute CV delta-gamma hedge error over the sample period. The statistical significance of these differences is confirmed by robust t-statistics greater than two, as reported in the sixth column of the second panel.

The same test is applied to compare each hedge to the GCOMP delta-gamma hedge. The seventh and eighth column of the second panel show that all of the alternative hedges, except the convexity hedge, are significantly inferior to the GCOMP hedge. This is indicated by positive t-statistics greater than two in the eighth column. In terms of variance reduction, the GCOMP model improves on the alternative GARCH models with a value of 24.5% compared to 19.0% for the GJR model and 17.1% for the GARCH(1,1) model.

It is also notable that the CV delta-vega hedge increases portfolio variance above that of an unhedged portfolio. Clearly, treating a volatility shock as a one-and-for-all change that affects all parts of the volatility term structure equally is unrealistic based on this result. The fact that all of the GARCH models are superior to the ARIV model indicates that the addition of mean-reversion to the volatility process, without a realistic specification of the relationship between return magnitudes and future volatility, is inadequate for modeling volatility term-structure dynamics and hedging changes in volatility. Thus, the hedging performance tests highlight a particular form of model misspecification: an omitted variable in the volatility equation.

The third panel of Table 6 provides an out-of-sample evaluation of hedging performance. For the out-of-sample tests, all volatility models and volatility hedge ratios are estimated using only data available at the time of hedge portfolio construction. These results are similar to those in the first part of Table 6 in that the CV delta-gamma hedge is most effective followed by the GCOMP delta-gamma hedge. The t-statistics in the out-of-sample hedging tests are lower than for the in-sample tests, as is expected due to the daily re-estimation of the volatility models, and are not significant in all cases. However, the out-of-sample results are consistent with the in-sample results in terms of variance reduction, and no stochastic volatility hedge is a statistically significant improvement over the GCOMP delta-gamma hedge.

6. Convexity versus volatility hedges: the role of sample selection

The hedging results of section 5 indicate that if volatility is to be considered as a risk factor, it should be expressed as a function of returns. However, the superior performance of the CV delta-gamma hedge relative to the GARCH models suggests that volatility has a zero or negligible reaction to return magnitudes. Relative S&P500 straddle price changes seem to be most accurately predicted by their sensitivity to second-order changes in the underlying price. It is puzzling that stochastic volatility models appear to perform well in predicting option prices, but poorly in predicting relative option price changes.

A possible source of the apparent superior performance of the convexity hedge relative to the volatility hedges is sample selection bias. Of the 2654 trading days in the sample period, there are 1296 days for which one-day option price changes are available for short and medium term straddles. Since at-the-money

options are the most liquid contracts, it is most likely that an at-the-money straddle will be traded for two days in a row when the underlying price does not change significantly. In other words, straddle price changes tend to be available on low volatility days.

The first panel of Table 7 analyzes the movements in the S&P500 index for days with and without option price change data. The difference in return volatility for sample and non-sample days is striking: the annualized S&P500 return standard deviation on sample days is 11.7% compared to 19.2% for excluded days. This implies that days without option price change data are high volatility days.

If relative hedging performance is different on low and high volatility days, the observed test results may be biased. The second panel of Table 7 measures hedging performance on high volatility days by constructing a subsample of the observed option price change sample into the 100 days with largest S&P500 return magnitudes. In this subsample, all of the GARCH volatility hedges outperform the convexity hedge. For example, the variance reduction for the GCOMP hedge is 53.5% compared to 48.9% for the CV delta-gamma hedge. If the relative performance of the models for the unobserved days follows this pattern, the GARCH hedges would dominate the convexity hedge.

Notice that the relative performance among the volatility hedges does not change on the high volatility days. In other words, in this subsample the GCOMP hedge is superior to the alternative GARCH hedges as well as the ARIV and CV delta-vega hedges. This reinforces the evidence for mean reversion in volatility and correlation between return magnitudes and volatility. It also suggests that the strongest results from the hedging tests are the relative rankings of the stochastic volatility hedges.

7. Conclusions

This paper provides a methodology for testing volatility models based on their ability to construct option portfolios hedged against shifts in the volatility term structure. Approximate at-the-money option hedge parameters are developed for several volatility model specifications, and volatility hedging tests are implemented using S&P500 index options.

Of the stochastic volatility hedges, the GARCH components with leverage delta-gamma hedge is most effective, indicating that it provides the most accurate estimates of VTS dynamics. The poorer hedging performance of the CV delta-vega and ARIV models suggests that volatility changes must be linked to the magnitude of underlying asset returns.

In the observed sample, a convexity hedge outperforms the volatility hedges. The effect of missing observations on high volatility days (i.e. sample selection bias) indicates that volatility hedges might outperform the convexity hedge if option price changes on all days could be observed.

Bibliography

- Amin, K. I. and V. K. Ng, 1993, "Option Valuation with Systematic Stochastic Volatility," *Journal of Finance*, 48, 881-910.
- Amin, K. I. and V. K. Ng, 1994, "A Comparison of Predictable Volatility Models using Option Data," Manuscript.
- Andersen, T. G., 1994, "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling," *Mathematical Finance*, 4, 75-102.
- Black, F. and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637-654.
- Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T., R. Y. Chou and K. F. Kroner, 1992, "ARCH Modeling in Finance - A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5-59.
- Bollerslev, T. and J. M. Wooldridge, 1992, "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances," *Economics Letters*, 12, 143-172.
- Diz, F. and T. J. Finucane, 1993, "Do the Options Markets Really Overreact?," *Journal of Futures Markets*, 13, 299-312.
- Duan, J. C., 1995, "The GARCH Option Pricing Model," *Mathematical Finance*, 5, 13-32.
- Engle, R. F., 1982, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.
- Engle, R. F. and G. Lee, 1993, "A Permanent and Transitory Component Model of Stock Return Volatility," University of California - San Diego, UCSD discussion paper, 92-44R.
- Engle, R. F., D. M. Lilien and R. P. Robins, 1987, "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M model," *Econometrica*, 55, 391-407.
- Engle, R. F. and J. V. Rosenberg, 1995, "GARCH Gamma," *Journal of Derivatives*, 2, 47-59.
- Glosten, L. R., R. Jagannathan and D. E. Runkle, 1993, "On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779-1801.
- Heynen, R., A. Kemna and T. C. F. Vorst, 1994, "Analysis of the Term Structure of Implied Volatilities," *Journal of Financial and Quantitative Analysis*, 29, 31-56.
- Hull, J. and A. White, 1987a, "Hedging the Risks from Writing Foreign Currency Options," *Journal of International Money and Finance*, 6, 131-152.

- Hull, J. and A. White, 1987b, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42, 281-301.
- Melino, A. and S. M. Turnbull, 1990, "Pricing Foreign Currency Options with Stochastic Volatility," *Journal of Econometrics*, 45, 239-265.
- Melino, A. and S. M. Turnbull, 1995, "Misspecification and the Pricing and Hedging of Long Term Foreign Currency Options," *Journal of International Money and Finance*, 14, 373-393.
- Stein, J., 1989, "Overreactions in the Options Market," *Journal of Finance*, 44, 1011-1023.
- Taylor, S. J., 1994, "Modeling Stochastic Volatility: A Review and Comparative Study," *Mathematical Finance*, 4, 183-204.
- White, H., 1980, "A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.
- Xu, X. and S. J. Taylor, 1994, "The Term Structure of Volatility Implied by Foreign Exchange Options," *Journal of Financial and Quantitative Analysis*, 29, 57-74.
- Zakoian, J.-M., 1994, "Threshold Heteroskedastic Models," *Journal of Economic Dynamics and Control*, 18, 931-955.

Table 1

Descriptive statistics for data used in estimation of the volatility models

Descriptive statistics for S&P500 return data used in CV and GARCH model estimation

	Number obs.	Mean	Standard deviation	Skewness	Excess kurtosis	Ljung-Box (15) on log returns	Ljung-Box (15) on squared log returns
S&P500 index daily log return	3655	0.1286	0.1547	-3.72	88.85	32.19	234.83

Descriptive statistics for implied variance data (short term implied variance is used in ARIV model estimation)

	Number obs.	Mean	Std. dev.	Skewness	Excess kurtosis	First order autocorrelation
Short term implied std. dev.	1296	0.1403	0.0384	1.08	0.96	0.93
Medium term implied std. dev.	1296	0.1475	0.0397	1.01	0.67	0.95
Short term implied std., 1st diff.	1295	0.0000	0.0143	0.65	7.92	-0.23
Medium term implied std., 1st diff.	1295	0.0000	0.0123	0.65	5.46	-0.14

This table presents descriptive statistics for data used in estimation of the volatility models. S&P500 daily log returns are calculated using the log-daily capital appreciation for the S&P500 index from June 20, 1982 through June 30, 1996. The mean (annualized), standard deviation (annualized), skewness, and excess kurtosis are sample statistics. Ljung-Box refers to the Ljung-Box Q-statistic (Box, 1970) which measures autocorrelation and is calculated using 15 lags. The critical value is approximately 25. Summary statistics for implied standard deviations and their first differences are reported for the 1296 days over the period January 1986 through June 1996 for which short and medium term at-the-money S&P500 index straddle price changes are available. Implied standard deviations (annualized) are estimated by inverting the Black-Scholes (1973) formula for short and medium term at-the-money S&P500 puts and calls using closing option prices. The short and medium term implied standard deviations are the square root of the average of the at-the-money put and call implied variances for the respective maturity. The Black-Scholes formula is adjusted for dividends by discounting the current index level by the present value of realized dividends over the life of the option. Realized dividends are extracted using total and capital appreciation returns reported in the CRSP stock file index database. The riskless interest rate used is the 13 week U.S. Treasury Bill yield. The number of days until expiration is calculated using trading days.

Table 2

Estimated volatility models used in volatility term structure tests

CV model	
Parameter	Sample standard deviation
σ	0.1547

ARIV model			
Parameter	Estimated coefficient	t-statistic	RMSE
ω	7.18E-06	6.54	2.01E-05
ρ	0.9140	81.22	
Adjusted r-squared	0.84		
First-order autocorr.	-0.17		

GARCH(1,1) model			
Parameter	Estimated coefficient	t-statistic	Robust t-statistic
μ	0.0004	3.17	3.06
ω	1.10E-06	6.83	2.48
α	0.0643	37.18	1.41
β	0.9245	247.47	22.54
γ			
ϕ			
ρ			
Maximized log-likelihood func. value		12339.27	
Ljung-Box (15) on std. sqr. ret.		8.45	

GJR model		
Parameter	Estimated coefficient	Robust t-statistic
μ	0.0003	2.18
ω	1.36E-06	2.57
α	0.0387	2.82
β	0.9213	29.44
γ	0.0489	1.03
Maximized log-likelihood func. value		12349.92
Ljung-Box (15) on std. sqr. ret.		5.86

GCOMP model		
Parameter	Estimated coefficient	Robust t-statistic
μ	0.0003	2.40
ω	3.76E-07	2.93
α	0.0000	0.00
β	0.7044	5.33
γ	0.1477	1.41
ϕ	0.0256	2.76
ρ	0.9954	369.74
Maximized log-likelihood func. value		12387.11
Ljung-Box (15) on std. sqr. ret.		5.71

This table presents the estimated volatility models which are used to forecast the volatility term structure and to calculate the volatility hedge ratios. The CV model refers to the constant volatility model, which is estimated using the annualized sample standard deviation of daily log-returns for the S&P500 index over the period June 20, 1982 to June 30, 1996. The ARIV model refers to the autoregressive implied volatility model which is estimated as an AR(1) model with short term implied variance proxying the one-day implied variance, following the methodology of Heynen, Kemna, and Vorst (1994). The characteristics of the 1296 short term implied variances used in ARIV estimation are detailed in Table 1. The ARIV model is defined in equation (5), and RMSE refers to the root mean squared error from the regression. The three GARCH models are estimated by maximization of the log-likelihood function using daily log-returns excess returns for the S&P500 index over the period June 20, 1982 to June 30, 1996. Returns during the week of the October 1987 crash are down-weighted by twenty percent in the log-likelihood function. Robust t-statistics are calculated according to the method of Bollerslev and Wooldridge (1992). The GARCH(1,1) model is defined in equations (8) and (9), the GJR model is defined in equations (12) and (13), and the GCOMP model is defined in equations (16)-(18). The Ljung-Box statistic measures persistence in squared standardized returns using 15 lags.

Table 3

Estimated volatility hedge ratios for at-the-money S&P500 index straddles

	Medium term days until maturity = 25	Medium term days until maturity = 30	Medium term days until maturity = 40		
	Short term days until maturity = 5	Short term days until maturity = 10	Short term days until maturity = 20	Average hedge ratio over sample	IQR of hedge ratios over sample
Volatility hedge ratio					
CV vega hedge ratio	2.24	1.73	1.41	1.72	0.45
ARIV vega hedge ratio	1.10	0.91	0.82	0.83	0.13
GARCH(1,1) gamma hedge ratio	0.82	0.95	1.01	0.93	0.08
GJR gamma hedge ratio	0.79	0.92	0.98	0.90	0.08
GCOMP gamma hedge ratio	0.63	0.76	0.88	0.79	0.13
CV gamma hedge ratio	0.45	0.58	0.71	0.60	0.16

This table presents the estimated volatility hedge ratios for a medium term at-the-money S&P500 index straddle with a short term straddle based on several volatility models. The volatility hedge ratio gives the number of short term straddles to be purchased for each medium straddle written to eliminate option sensitivity to volatility news. The particular medium and short term number of days until maturity are given in the column labels of the table. CV, ARIV, and GARCH volatility hedge ratios are given in equations (21), (22), and (23). Volatility multipliers are calculated using the estimated volatility model parameters in Table 2. Average annual expected volatility is set to 15.9%, the riskless rate is set to zero, the dividend yield is set to zero, and the index level and option strike price are set to 100. The average and interquartile range (IQR) of hedge ratios over the sample are calculated using the 1296 days of straddle price changes available from January 1986 to June 1996, the characteristics of the contracts traded, and the estimated volatility model parameters.

Table 4
Comparison of simulated and BSP approximate hedge ratios

Characteristics of delta hedge ratio approximation error

	GARCH(1,1)		GCOMP	
	delta	GJR delta	delta	ARIV delta
Mean error	0.0081	0.0372	0.0262	-0.0013
Std. error	0.0078	0.0077	0.0050	0.0039
Min. error	-0.0055	0.0165	0.0184	-0.0126
Max. error	0.0226	0.0495	0.0405	0.0093
Avg. sim. hedge ratio	0.0373	0.0647	0.0527	0.0255

Characteristics of gamma and vega hedge ratio approximation error

	GARCH(1,1)		GCOMP	
	gamma	GJR gamma	gamma	ARIV vega
Mean error	-0.0128	-0.0143	-0.0090	-0.0179
Std. error	0.0072	0.0068	0.0092	0.0174
Min. error	-0.0304	-0.0292	-0.0420	-0.0600
Max. error	0.0284	0.0210	0.0154	0.0061
Avg. sim. hedge ratio	0.9871	0.9558	0.9155	0.8478

This table presents the accuracy of the Black-Scholes plug-in (BSP) approximations to fully-simulated option hedge parameters under GARCH and autoregressive implied volatility (ARIV) models. Simulated option hedge parameters consistent with GARCH and ARIV models are obtained using Monte-Carlo simulation with 50,000 replications under the Amin-Ng (1994) and Hull and White (1987b) risk-neutral measures using empirical values for the volatility processes in Table 2 and centered finite difference approximations. Simulated hedge parameters are smoothed according to the methodology described in Rosenberg and Engle (1995). The current level of average expected volatility is set equal to the unconditional volatility for each model. BSP hedge ratios are estimated using the estimated volatility process parameters in Table 2 and equations (3), (4), and (7). The simulated hedge parameters are estimated for straddles with 5 to 90 days until maturity and for moneynesses (underlying price/strike price -1) of -.01, -.005, 0, .005, and .01 to reflect the moneyness range used in the empirical hedging tests. Delta hedge ratios are calculated for straddles with 5 to 90 days until maturity, while ARIV vega and GARCH gamma hedge ratios are calculated for pairs of equal moneyness straddles which have a difference of 20 days until maturity. The longer maturity straddle has from 25 to 90 days until maturity and the shorter maturity straddle has from 5 to 70 days until maturity. The approximation error is calculated by taking the difference of the simulated straddle delta, gamma, or vega hedge ratio and the BSP approximation. The sample properties of the approximation errors (mean, standard deviation, minimum, and maximum) over the range of moneyness and time to maturity for each model are listed in the table along with the average simulated hedge ratio.

Table 5

Summary of option data used in hedging performance tests

	Short term call options	Medium term call options	Short term put options	Medium term put options
Number of observations	1296	1296	1296	1296
Average price	5.38	9.60	5.04	8.45
Std. dev. price	1.98	2.90	1.77	2.29
Average price change	0.03	0.04	-0.14	-0.15
Std. dev. price change	1.53	1.56	1.56	1.52
Skewness of price change	0.87	0.42	1.32	1.09
Kurtosis of price change	2.14	1.66	6.11	6.68
Average daily return	0.0069	0.0081	-0.0268	-0.0291
Std. dev. daily return	0.3160	0.1670	0.3473	0.3295
Skewness of daily return	1.35	0.64	3.97	2.88
Kurtosis of daily return	5.05	1.68	52.62	37.28
Average time to maturity	14.31	38.11	14.31	38.11
Std. dev. time to maturity	6.44	11.79	6.44	11.79
Average moneyness	-0.0002	0.0038	-0.0002	0.0038
Std. dev. moneyness	0.0002	0.0038	0.0002	0.0038

This table presents summary statistics describing the S&P500 index option data used in the hedging performance tests of the volatility term structure. There are 1296 days for which short and medium term straddle price changes are available from January 1986 to June 1996. The yearly number of data points used from 1986 to 1996 are: 35, 97, 60, 95, 119, 132, 147, 181, 183, 163, 84. No data is used in October 1987 subsequent to October 15. Option returns are defined as the one-day net proportional change in the closing option price. Time to maturity is measured in days. Moneyness is defined as the ratio of the underlying price to the strike price minus one. Option prices are quoted in units (\$100 increments). The option data is from the Chicago Board Options Exchange.

Table 6

Tests of volatility term-structure forecasts using option hedging performance criteria

Hedge portfolios: (using in-sample forecasts)	<i>Standard deviation (annualized)</i>	<i>Interquartile range</i>	<i>5th percentile</i>	<i>Correlation with market factors (F probability)</i>			
No hedge	91.13	5.71	-9.52	0.0000			
Constant volatility delta hedge	90.65	5.62	-9.04	0.0000			
Constant volatility delta-vega hedge	136.98	7.51	-12.07	0.0000			
ARIV delta-vega hedge	85.27	5.60	-8.26	0.0000			
GARCH(1,1) delta-GARCH gamma hedge	82.96	5.53	-8.11	0.0000			
GJR delta-GARCH gamma	82.04	5.53	-8.02	0.0000			
GCOMP delta-GARCH gamma	79.18	5.35	-7.89	0.0014			
Constant volatility delta-gamma hedge	77.50	5.14	-7.88	0.0169			

Hedge portfolios: (using in-sample forecasts)	<i>Variance reduction (1986-1996)</i>	<i>Variance reduction (1986-1991)</i>	<i>Variance reduction (1992-1996)</i>	<i>Abs. hedge error vs CV delta- gamma</i>	<i>Robust t-statistic</i>	<i>Abs. hedge error vs GCOMP delta- GARCH gamma</i>	<i>Robust t-statistic</i>
No hedge	0.0%	0.0%	0.0%	0.54	7.02	0.46	4.88
Constant volatility delta hedge	1.1%	-0.6%	2.8%	0.45	5.73	0.37	3.82
Constant volatility delta-vega hedge	-125.9%	-145.6%	-105.2%	2.08	13.12	2.00	13.30
ARIV delta-vega hedge	12.4%	9.5%	15.5%	0.25	5.08	0.17	3.94
GARCH(1,1) delta-GARCH gamma hedge	17.1%	18.2%	15.7%	0.24	5.30	0.16	7.01
GJR delta-GARCH gamma	19.0%	20.1%	17.6%	0.20	4.79	0.12	6.34
GCOMP delta-GARCH gamma	24.5%	26.3%	22.4%	0.08	2.95	0.00	0.00
Constant volatility delta-gamma hedge	27.7%	29.4%	25.6%	0.00	0.00	-0.08	-2.95

Hedge portfolios: (using out-of-sample forecasts)	<i>Variance reduction</i>	<i>Standard deviation (annualized)</i>	<i>Interquartile range</i>	<i>5th percentile</i>	<i>Correlation with market factors (F probability)</i>	<i>Abs. hedge error vs GCOMP delta- GARCH gamma</i>	<i>Robust t-statistic</i>
Constant volatility delta-vega hedge	-129.9%	138.17	7.50	-12.12	0.0000	2.00	13.28
ARIV delta-vega hedge	23.4%	79.75	5.37	-7.96	0.0010	0.04	1.42
GARCH(1,1) delta-GARCH gamma hedge	22.8%	80.06	5.38	-7.81	0.0000	0.04	3.36
GJR delta-GARCH gamma	23.7%	79.59	5.35	-7.81	0.0001	0.01	1.45
GCOMP delta-GARCH gamma	24.6%	79.12	5.35	-7.89	0.0006	0.00	0.00
Constant volatility delta-gamma hedge	27.7%	77.50	5.14	-7.90	0.0128	-0.08	-2.93

These panels describe the characteristics of one-day hedge portfolio price changes for hedge portfolios constructed according to several models of the S&P500 return volatility process. Hedge ratios are estimated for a \$100 written position of nearest maturity at-the-money S&P500 index option straddles. This position is hedged with the next nearest maturity at-the-money S&P500 index straddle available and the S&P500 index portfolio. The first two panels describe tests of the volatility term structure based on in-sample forecasts. The third panel presents tests of VTS forecasts using out-of-sample forecasts. The hedge portfolio price changes are determined using close-to-close price changes for the index options and the S&P500 index. Hedging tests are conducted using the 1296 days where data is available over the period January 1986 - June 1996. The standard deviation (annualized), interquartile range, and 5th percentile refer to sample statistics calculated using the one-day hedge portfolio price changes. The correlation with market factors (F probability) is the F probability from a regression of hedge portfolio price changes on the change in the S&P500 index level and the squared change. Variance reduction is defined as the ratio of the difference between the unhedged portfolio variance and the hedge portfolio variance with the unhedged portfolio variance expressed as a percentage. Absolute hedge error versus CV delta-gamma and GCOMP delta-GARCH gamma are defined as the average difference between the de-measured absolute daily hedge portfolio price change for an alternative hedge portfolio and the CV or GCOMP portfolio. The robust t-statistic is the heteroskedasticity consistent t-statistic of White (1980) for the null hypothesis that the hedge performance of the alternative model and the CV or GCOMP models are equal.

Table 7

Evidence of sample selection bias in option price change data

Characteristics of S&P500 daily log returns, January 1986 - June 1996

	<i>Number of observations</i>	<i>Return mean (annualized)</i>	<i>Return standard deviation (annualized)</i>
All trading days	2654	0.1167	0.1596
Days used in hedging tests	1296	0.1609	0.1170
Missing data days	1358	0.0762	0.1916

Hedge performance for 100 largest absolute S&P500 return dates

Hedge portfolios: (using in-sample forecasts)	<i>Variance reduction</i>
No hedge	0.0%
Constant volatility delta hedge	-7.0%
Constant volatility delta-vega hedge	-123.8%
ARIV delta-vega hedge	38.5%
GARCH(1,1) delta-GARCH gamma hedge	50.1%
GJR delta-GARCH gamma	50.9%
GCOMP delta-GARCH gamma	53.5%
Constant volatility delta-gamma hedge	48.9%

These panels present evidence of sample selection bias in the option price change data due to the fact that option price changes for initially at-the-money positions are typically available only when the underlying price does not change substantially. The first panel presents characteristics of daily S&P500 index log returns for all trading days from January 1986 to June 1996, for days where option price changes are available, and for days when data is not available. The second panel presents characteristics of one-day hedge portfolio price changes for hedge portfolios constructed according to several models of the S&P500 return volatility process. Only the 100 largest absolute S&P500 return dates are included in this analysis to approximate the performance of the hedges for the unobserved dates. Hedge ratios are estimated for a \$100 written position of nearest maturity at-the-money S&P500 index option straddles. This position is hedged with the next nearest maturity at-the-money S&P500 index straddle available and the S&P500 index portfolio. These results are based on in-sample forecasts. The hedge portfolio price changes are determined using close-to-close price changes for the index options and the S&P500 index. Variance reduction is defined as the ratio of the difference between the unhedged portfolio variance and the hedge portfolio variance with the unhedged portfolio variance expressed as a percentage.

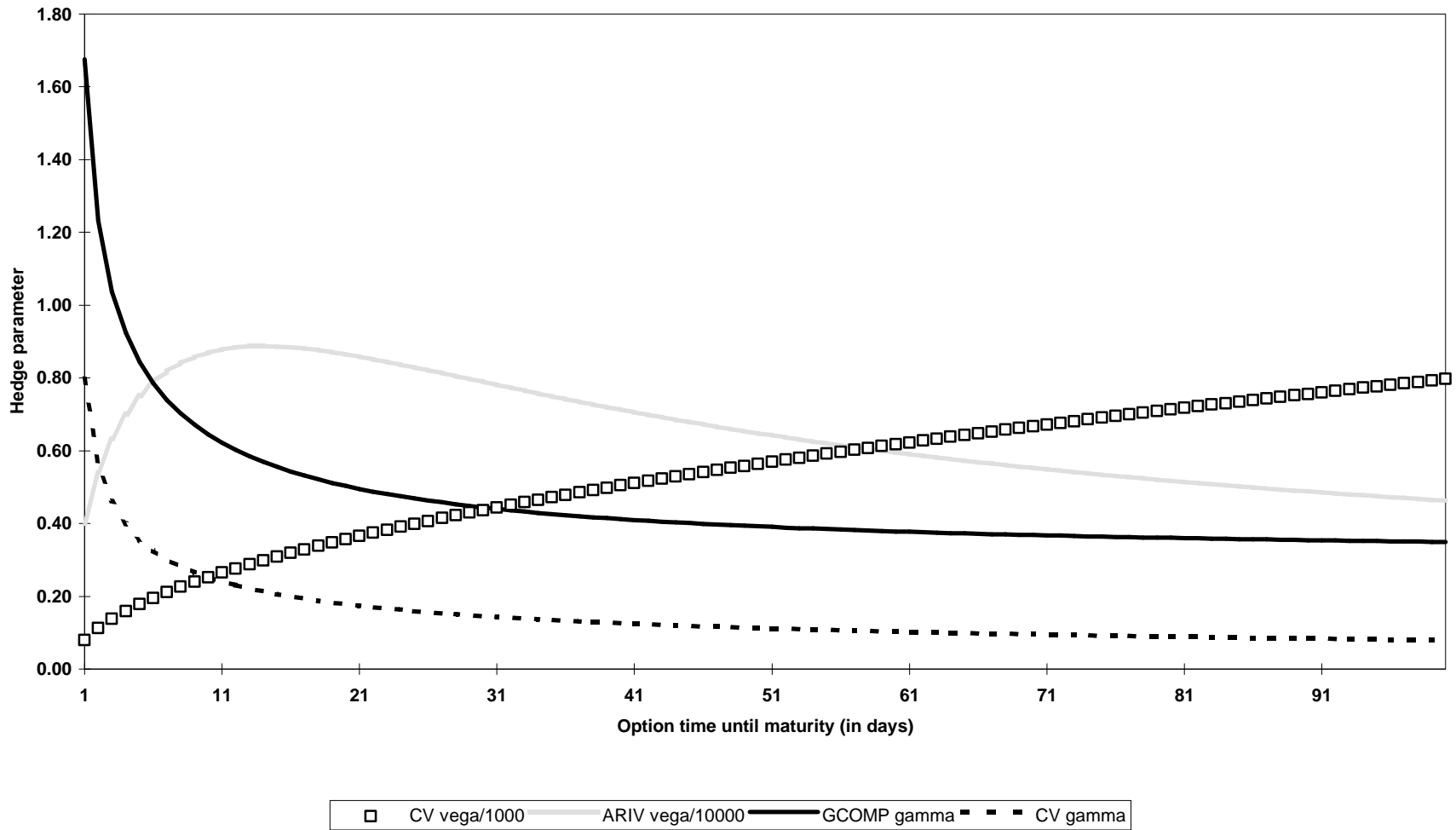


Figure 1
Comparison of volatility hedge parameters for an at-the-money S&P500 index straddle

CV vega and CV gamma are calculated using equation (2), and CV vega is scaled by dividing by 1000. ARIV vega is calculated using equations (4) and (6), the estimated ARIV parameters from Table 2, and is scaled by dividing by 10000. GCOMP gamma is calculated using equations (7), (19), and (20) with the estimated GCOMP parameters from Table 2. Average annual expected volatility is set to 15.9%, the riskless rate is set to 0, the dividend yield is set to zero, and the index level and option strike price are set to 100.