NCCC-134
APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

Comparing the Performances of the Partial Equilibrium and Time-Series Approaches to Hedging

by

Henry L. Bryant and Michael S. Haigh

Suggested citation format:

Bryant, H. L., and M. S. Haigh. 2003. "Comparing the Performances of the Partial Equilibrium and Time-Series Approaches to Hedging." Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.uiuc.edu/nccc134].

Comparing the Performances of the Partial Equilibrium and Time-Series Approaches to Hedging

Henry L. Bryant

and

Michael S. Haigh*

Paper presented at the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, St. Louis Missouri, April 21-22, 2003

* Bryant (**h-bryant@tamu.edu**) is a doctoral candidate in the Department of Agricultural Economics at Texas A&M University and Haigh (**mhaigh@arec.umd.edu**) is an Assistant Professor in the Department of Agricultural and Resource Economics at the University of Maryland.

Comparing the Performances of the Partial Equilibrium and Time Series Approaches to Hedging

Practitioner's Abstract

This research compares partial equilibrium and statistical time-series approaches to hedging. The finance literature stresses the former approach, while the applied economics literature has focused on the latter. We compare the out-of-sample hedging effectiveness of the two approaches when hedging commodity price risk using a simple derivative with a linear payoff function (a futures contract). For various methods of parameter estimation and inference, we find that the partial equilibrium models cannot out-perform the time series model. The partial equilibrium models' unpalatable assumption of deterministically evolving futures volatility seems to impede their hedging effectiveness, even when potentially foresighted option-implied volatility term structures are employed.

Keywords: Commodity contingent claims, multivariate GARCH, hedging.

I. Introduction

Two broad strategies for optimally hedging risky market commitments have emerged in the academic literature and in practice. The applied economics literature has focused on the use of statistical models of the observed time series of cash and futures prices in hedging. Early development of this type of optimal hedging is found in Johnson (1960), Peck (1975), and Kahl (1983), among others. Typically, this type of hedging considers an agent with a non-tradable position in a cash commodity, who plans to buy or sell some number of commodity futures contracts that will maximize her utility. This traditionally involved making static estimates of the variances of changes in the cash and futures prices and the covariance between those changes, and then choosing a level of hedging that would minimize the variance of changes in the hedger's portfolio value. More recently, Cecchetti, Cumby and Figlewski (1988), Myers (1991), and Baillie and Myers (1991) have adopted the use of models of time-varying conditional variance for optimal hedging. Noting that the use of differenced data will loose information about the long-run relationship between two time series, Kroner and Sultan (1993) incorporate the co-integrating relationship between cash and futures prices into their model. Gagnon, Lypny and McCurdy (1998) and Haigh and Holt (2000) extend these models to include multiple risks.

The finance literature on the other hand has typically stressed the use of partial equilibrium derivative pricing models for hedging. This began when Black and Scholes (1973) and Merton (1973) noted that the seller of a derivative could form a risk-free portfolio by holding just the right quantity of the underlying security. This quantity is determined by the rate a which the price of the derivative will change as the price of the underlying changes, referred to as the "delta." This type of hedging is therefore often referred to as "delta hedging."

In application, different types of hedgers have tended to make use of the two strategies. Holders of large derivative portfolios generally have employed partial equilibrium hedging. This is the

realm of financial institutions and "financial engineers" that sell derivatives to their customers at a markup to the value of a portfolio with price dynamics that replicate, as closely as possible, those of the derivative. Commodity producers and consumers, on the other hand, have more often used the time series approach. A typical picture is that of the agricultural producer with a crop in the ground, who wishes to minimize the risk that the price of the output will fall before the harvest. Despite their differences, these two types of hedgers face exactly the same problem: they each hold a position in one market (either underlying or derivative), and wish to take a position in the other market that will result in maximum benefit. Either hedger might use either of the two approaches to hedging, despite the traditional divide.

Each approach has it own merits and drawbacks. The time series approach does not require the imposition of theory *a priori*, thereby avoiding potential misspecification. Also, available time series models can very effectively represent time-varying covariability among price series, a commonly observed market phenomenon that is central to the hedging problem. This approach does not, however, make use of all available information. For example, time series hedging models consider neither the arbitrage activity that constrains the price of a derivative relative to its underlying security, nor theories regarding derivatives' price variability (e.g. the Samuelson (1965) hypothesis that a futures contract's volatility should increase as expiration approaches).

By contrast, the partial equilibrium approach directly incorporates the arbitrage relationship(s) between the derivative and underlying instrument(s). An additional benefit of this approach is the ability to use observed market prices to infer the expectations of market participants. For example, option prices can be used to infer the future levels of volatility that knowledgeable industry participants are anticipating in an associated underlying market. The adoption of the partial equilibrium approach comes at the price however of requiring various simplifying assumptions, which have varying degrees of implausibility. Crucially, most partial equilibrium models do not incorporate the stochastically time-varying volatility that is widely acknowledged to exist in most financial and commodity markets.

Given the above stated benefits and drawbacks of each of the two approaches to hedging, it is not immediately clear that one approach should be preferred in any given situation. No previous research has directly compared the effectiveness of these two hedging strategies, and we thus undertake such an evaluation here. We directly compare the in-sample and out-of-sample hedging performance of the two approaches for a trader that is long physical crude oil, and uses a simple derivative with a linear payoff function (a futures contract) to hedge the associated price risk. We assume that the hedger maximizes a mean-variance utility function, and hedging effectiveness is measured by the increases in utility that the hypothetical trader realizes by implementing each strategy (relative to not hedging at all). Two partial equilibrium models that have been developed in the commodity contingent claims pricing literature are considered – the Schwartz (1997) one-factor model, and the two-factor model of Gibson and Schwartz (1990). Various strategies for estimating and inferring these models' parameters are employed. The competing time series model is a vector error-correction model, with a generalized autoregressive conditional heteroskedastic error structure.

The rest of this paper is organized as follows: In section II, we describe the hedging problem and the time series model. In section III, we describe the partial equilibrium models, and

describe how they can adapted for optimal hedging by an agent with mean-variance type utility. This continuous time mean-variance hedging can be considered a generalization of delta hedging. We also show these models can be extended to allow for spatial and form differences between the commodity to which a hedger is committed and the commodity underlying the futures contract. Section IV discusses the estimation of all models, and section V reports their hedging effectiveness. Section VI concludes.

II. Hedging Commodity Price Risk Using Time Series Models

We consider a hedger that is long a physical commodity, and wishes to optimally select of a quantity of futures contracts to sell. The hedge ratio, *b*, is the ratio of the size of the futures market position to the size of the cash market position. The change in the hedger's portfolio value over the discrete interval from time *t-1* to time *t* is given by

$$P_t - P_{t-1} = (L_t - L_{t-1}) - b_{t-1}(F_t - F_{t-1}) \tag{1}$$

where P_t, L_t, and F_t represent portfolio value, the local cash price of the commodity held by the hedger, and the futures price, respectively, in period *t*. Note that the commodity held by hedger does not necessarily correspond exactly to the commodity underlying the futures contract. The hedger may be holding a different grade of the commodity than is called for by the futures contract (or a different commodity altogether), and she may not be able to deliver her commodity against the futures contract at par value locally. We therefore distinguish between a local cash price of an arbitrary commodity, and the price at the specified futures delivery location of the specified commodity. We refer to the former as a local cash price L_t as above, and to the latter as the spot price S_t.

We assume that the hedger maximizes mean-variance type utility. The hedger's problem for each period is then is formulated as follows:

$$\underset{b_{t-1}}{Max}\left[E(\Delta P_t \mid \Omega_{t-1}) - \frac{\lambda_U}{2} \operatorname{var}(\Delta P_t \mid \Omega_{t-1}) \right] \tag{2}$$

where *E* is the conditional expectation operator, ΔP_t is the change in portfolio value from *t*-1 to *t*, Ω_{t-1} is the information available as of *t*-1, λ_U is the coefficient of absolute risk aversion, and var() is the conditional variance operator. Note that the risk-minimizing objective is a special case of (2) where $\lambda_U = \infty$. Note that the conditional variance term in (2) can be expanded, using (1), as

$$\operatorname{var}(\Delta L_t \mid \Omega_{t-1}) + b_{t-1}^2 \operatorname{var}(\Delta F_t \mid \Omega_{t-1}) - 2b_{t-1} \operatorname{cov}(\Delta L_t, \Delta F_t \mid \Omega_{t-1}) \tag{3}$$

where cov() is the conditional variance operator. The utility-maximizing hedge ratio is then given by

$$b_{t-1} = \frac{-\lambda_U^{-1} E(\Delta F_t \mid \Omega_{t-1}) + \mathrm{cov}(\Delta L_t, \Delta F_t \mid \Omega_{t-1})}{\mathrm{var}(\Delta F_t \mid \Omega_{t-1})}. \tag{4}$$

The second-order condition for this problem is the negative of the risk aversion coefficient multiplied by the conditional variance of changes in the futures price, and we are thus guaranteed a global maximum for a risk-averse hedger. If we have $\lambda_U = \infty$, the first term in the numerator is zero and we have the minimum-variance hedge ratio. For $0 > \lambda_U > \infty$, the optimal hedge ratio contains the minimum-variance component, and a speculative component. If our hedger anticipates a decrease in the futures price, he will reduce the level of hedging to below the minimum variance level to avoid losses in the futures market. Likewise, an anticipated increase in the futures price will compel our hedger to increase the size of the futures position.

Calculating the optimal hedge ratio in equation (4) requires the time-series modeler to provide two types of information – the conditional expected futures price change and conditional variance-covariance forecasts. Recent academic hedging research advocates obtaining the first piece of information using a vector error correction (VEC) model. This an the appropriate modeling technique in the event that each of the two price series is found to follow a unit root process, but a linear combination of the two is found to be stationary (Engle and Granger, 1987). This linear combination is interpreted as representing a long-run equilibrium between the two levels series. The VEC model is essentially a vector auto-regression model in which a deviation from the long-run equilibrium (the "error") in one time period is subject to some degree of correction in the following time period. A basic representation of a VEC for 2 variables is as follows:

$$\Delta y_t = \pi_0 + \sum_{i=1}^{r} \pi_i \Delta y_{t-i} + \alpha\beta y_{t-1} + \varepsilon_t \tag{5}$$

where y_t is the 2×1 vector of observations at time t, π_0 is a 2×1 parameter vector, π_i is a 2×2 coefficient matrix, β is the co-integrating vector characterizing the long-run equilibrium, α is a 2×1 coefficient vector, and ε_t is a vector of innovations. The inner product βy_{t-1} is the deviation from the long-run equilibrium, and α characterizes the rate at which each of the two variables responds to this deviation. Forming y using cash and futures prices, Equation (5) can then be used to generate forecasts of futures price changes – one of the components of the optimal hedge ratio above.

The other pieces of information that are required to calculate the optimal hedge ratio in (4) are the conditional variances and covariance. These can be forecast using multivariate versions of the auto-regressive conditional heteroskedasticity (ARCH) model of Engle (1982) or the generalized ARCH (GARCH) model of Bollerslev (1986). A GARCH error structure implies that the conditional second moment of the innovation vector of a model follows an autoregressive, moving average process – it is a function of past innovation vectors and past second moments. Here we employ a GARCH(1,1) model with the diagonal vech parameterization of Bollerslev, Engle, and Wooldridge (1988). The conditional distribution of the error vector from (5) is then given by

(6) $$\varepsilon_t \mid \Omega_{t-1} \sim N(0, H_t)$$

(7) $$\text{vech}(H_t) = W + A\ \text{vech}(\varepsilon_{t-1}\varepsilon_{t-1}^T) + B\ \text{vech}(H_{t-1}).$$

Here, vech() is the column stacking operator that stacks the lower triangular portion of a symmetric matrix, W is a 3×1 vector of constants, and A and B are a diagonal 3×3 coefficient matrices. Equation (7) can be used to form one-period ahead forecasts of the variance of futures price changes and the covariance between futures and cash price changes. The VEC-GARCH model given by equations (5) through (7) thus provides a means by which a hedger can select the optimal level of hedging.

III. Hedging Commodity Price Risk Using Partial Equilibrium Models

Early models for pricing contingent claims included only a single stochastic factor, the price of the underlying asset. These models assumed that a risk-free portfolio consisting of a short position in the derivative contract and a long position in the underlying asset could be formed, and that this portfolio should earn the risk-free rate of return. Ross (1978) noted that this assumption is inappropriate in the event that there are benefits to holding an actual asset, rather than merely holding a contract calling for future delivery. When the asset is a commodity, the flow of these benefits is referred to as a convenience yield. Kaldor (1939) describes this phenomenon, and it features prominently in the theory of storage developed in Working (1949) and Brennan (1959). Consideration of the convenience yield motivated the development of the Brennan and Schwartz (1985) model for pricing commodity contingent claims, which assumed that a commodity's convenience yield was a constant proportion of the spot price. This assumption that the convenience yield could be specified as a deterministic function of a commodity's spot price was investigated empirically in Brennan (1991), and Gibson and Schwartz (1991). Both studies decisively concluded that such an assumption was inappropriate, and that the convenience yield should be specified as a second stochastic factor.

Gibson and Schwartz (1990) thus developed a model for pricing commodity contingent claims with two stochastic factors, the first being the spot price of the commodity and the second being the instantaneous net (of storage costs) convenience yield of the commodity. In this model, the holder of a commodity derivative faces not only the risk that the spot price of the commodity will change, but also the risk associated with changes in the convenience yield. As it is not possible to hedge the latter risk, the hedger will not be able to form a completely risk-free portfolio, and the Gibson-Schwartz (GS) model is one of incomplete markets.

The GS model assumes that the spot price of a commodity S and associated instantaneous net convenience yield δ follow the joint diffusion process given by

(8) $$dS / S = \mu dt + \sigma_1 dz_1$$

(9) $$d\delta = k(\alpha - \delta)dt + \sigma_2 dz_2$$

where μ is the drift of spot price returns, σ_1^2 and σ_2^2 are the instantaneous variances of spot price returns and the convenience yield respectively, dz_1 and dz_2 are increments to correlated Brownian motions, with the multiplication rule $dz_1 dz_2 = \rho_{12} dt$, and ρ_{12} being the correlation coefficient. The convenience yield is assumed to revert at rate k to a long-run mean level α. By Ito's Lemma, the price $G(S,\delta,\tau)$ of a commodity contingent claim that is a function of time, and a twice continuously differentiable function of S and δ then follows the diffusion

$$(10) \qquad dG = \left[-G_\tau - \frac{1}{2} G_{SS} \sigma_1^2 S^2 + G_{S\delta} S \rho_{12} \sigma_1 \sigma_2 + \frac{1}{2} G_{\delta\delta} \sigma_2^2 + G_S \mu S \right] dt + \left[\sigma_1 S G_S \right] dz_1 + \left[\sigma_2 G_\delta \right] dz_2$$

where $\tau = T - t$ is the length of time from the present (t) until expiration of the derivative (T), and G_X represents the partial derivative of G with respect to X. Gibson and Schwartz present a no-arbitrage argument that leads to following partial differential equation that must be satisfied by the price $F(S,\delta,\tau)$ of a futures contract:

$$(11) \qquad \frac{1}{2} F_{SS} S^2 \sigma_1^2 + \frac{1}{2} F_{\delta\delta} \sigma_2^2 + F_{S\delta} S \rho_{12} \sigma_1 \sigma_2 + F_\delta S (r - \delta) + F_\delta \left[k(\alpha - \delta) - \lambda \sigma_2 \right] - F_\tau = 0$$

where r is the risk-free rate of return, and λ is the market price of convenience yield risk. The solution to (11), as reported in Hilliard and Reis (1998)[1] is

$$(12) \qquad F(S,\delta,\tau) = S \exp\left\{ \left[\frac{-\delta(1-e^{-k\tau})}{k} \right] + \tau \left[r - \alpha + \frac{\lambda\sigma_2}{k} + \frac{\sigma_2^2}{2k^2} - \frac{\sigma_1 \sigma_2 \rho_{12}}{k} \right] + \left(\frac{1-e^{-k\tau}}{k} \right) \left[\alpha - \frac{\lambda\sigma_2}{k} - \frac{\sigma_2^2}{k^2} + \frac{\sigma_1 \sigma_2 \rho_{12}}{k} \right] + \left[\frac{\sigma_2^2 \left(1 - e^{-2k\tau}\right)}{4k^3} \right] \right\}.$$

We now turn to the task of adapting the GS model for use in hedging. Using (12) to find the appropriate partial derivatives to substitute into (10), we find the diffusion followed by a futures contract to be

$$(13) \qquad dF = \left[F\left(\mu - (r - \delta) - \frac{\lambda\sigma_2}{k}\left(1 - e^{-k\tau}\right) \right) \right] dt + \left[F\sigma_1 \right] dz_1 + \left[F\left(\frac{-\sigma_2 (1 - e^{-k\tau})}{k} \right) \right] dz_2 .$$

For a hedger whose local cash price corresponds to the spot price, changes in portfolio value are given by $dP = dS - b\,dF$. Using this, applying Ito's Lemma to equation (8), and using equation (13), we find that the short hedger's portfolio dynamics are described by the diffusion

$$(14) \qquad dP = \left[S\left(\mu + \frac{\sigma_1^2}{2} \right) - bF\left(\mu - (r - \delta) - \frac{\lambda\sigma_2}{k}\left(1 - e^{-k\tau}\right) \right) \right] dt$$

$$+\left[(S-bF)\sigma_1\right]dz_1+\left[bF\left(\frac{\sigma_2\left(1-e^{-k\tau}\right)}{k}\right)\right]dz_2$$

Defining another standard Brownian motion z and a parameter σ_P such that

(15) $$\sigma_P dz=\left[(S-bF)\sigma_1\right]dz_1+\left[bF\left(\frac{\sigma_2\left(1-e^{-k\tau}\right)}{k}\right)\right]dz_2,$$

we can simplify equation (14) to

(16) $$dP=\mu_P dt+\sigma_P dz$$

with drift $\mu_P=\left[S\left(\mu+\frac{\sigma_1^2}{2}\right)-bF\left(\mu-(r-\delta)-\frac{\lambda\sigma_2}{k}\left(1-e^{-k\tau}\right)\right)\right]$ and instantaneous variance

(17) $$\sigma_P^2=S^2\sigma_1^2+b^2F^2\left[\sigma_1^2-2\sigma_1\sigma_2\rho_{12}\left(\frac{\left(1-e^{-k\tau}\right)}{k}\right)+\sigma_2^2\left(\frac{\left(1-e^{-k\tau}\right)}{k}\right)^2\right]$$
$$-2bSF\left[\sigma_1^2-\sigma_1\left(\frac{\sigma_2\left(1-e^{-k\tau}\right)}{k}\right)\rho_{12}\right]$$

This expression for the instantaneous variance of changes in portfolio value is analogous to equation (3) – the first term is the instantaneous variance of spot price changes, the second term is b^2 multiplied by the instantaneous variance of futures price changes, and the third term is $-2b$ multiplied by the instantaneous covariance between spot and futures price changes. Armed with the above specification for the controlled stochastic process followed by the hedger's portfolio, we are in a position to solve the continuous time version of the hedging problem given by equation (2). In the context of the GS model, we find the following expression for the optimal hedge ratio:

(18) $$b_{GS}=\frac{-\lambda_U^{-1}\mu_F+SF\left[\sigma_1^2-\sigma_1\left(\frac{\sigma_2\left(1-e^{-k\tau}\right)}{k}\right)\rho_{12}\right]}{F^2\left[\sigma_1^2-2\sigma_1\sigma_2\rho_{12}\left(\frac{\left(1-e^{-k\tau}\right)}{k}\right)+\sigma_2^2\left(\frac{\left(1-e^{-k\tau}\right)}{k}\right)^2\right]}$$

where

(19) $$\mu_F=F\left(\mu-(r-\delta)-\frac{\lambda\sigma_2}{k}\left(1-e^{-k\tau}\right)\right).$$

Note that the above expression for the optimal hedge ratio has been developed for a hedger whose cash market commitment exactly corresponds to the commodity underlying the futures contract (i.e. $L = S$). This result is of limited usefulness, as many hedgers' cash market commitments vary from the specifications of the futures contract. The GS model can be augmented, however, to derive a more general formulation. We define the difference between the hedger's cash price and the spot price as

$$B \equiv L - S \tag{20}$$

and we propose the following stochastic process for B:

$$dB = \gamma(\beta - B)dt + \sigma_3 dz_3 \tag{21}$$

where σ_3^2 is the instantaneous variance of changes in B, dz_3 is a third Brownian motion, and we add the multiplication rules $dz_1 dz_3 = \rho_{13} dt$ and $dz_2 dz_3 = \rho_{23} dt$. We assume that B reverts to level β at rate γ. The mean-reverting nature of B is justified in the event that a stable long-run relationship between the cash and spot prices exists. In the event that no such relationship existed, the futures contract would make an inappropriate hedging vehicle for the cash price concerned. Changes in the hedger's portfolio are then given by $dP_A = dB + dS - b dF$, and we can follow a succession of steps similar to those above to arrive at the following diffusion:

$$dP_A = \mu_{AP} dt + \sigma_{AP} dz \tag{22}$$

with drift $\mu_{AP} = \left[\gamma(\beta - B) + S\left(\mu + \frac{\sigma_1^2}{2}\right) - bF\left(\mu - (r - \delta) - \frac{\lambda\sigma_2}{k}\left(1 - e^{-k\tau}\right)\right)\right]$ and instantaneous variance

$$\begin{aligned}\sigma_{AP}^2 = {} & \left[S^2\sigma_1^2 + 2S\sigma_1\sigma_3\rho_{13} + \sigma_3^2\right] \\ & + b^2F^2\left[\sigma_1^2 - 2\sigma_1\sigma_2\rho_{12}\left(\frac{\left(1 - e^{-k\tau}\right)}{k}\right) + \sigma_2^2\left(\frac{\left(1 - e^{-k\tau}\right)}{k}\right)^2\right] \\ & - 2b\left[SF\sigma_1^2 - SF\sigma_1\left(\frac{\sigma_2\left(1 - e^{-k\tau}\right)}{k}\right)\rho_{12} - F\sigma_3\left(\frac{\sigma_2\left(1 - e^{-k\tau}\right)}{k}\right)\rho_{23} + F\sigma_1\sigma_3\rho_{13}\right].\end{aligned} \tag{23}$$

The differences between expressions (17) and (23) are in the terms that represents the instantaneous variance of cash price changes and the covariance between cash and futures price changes. The variance of cash price changes now reflects the interaction between the spot price and its difference with the local cash price. The covariance term now contains portions that reflect the covariation of B with the other stochastic factors in the model. This results in an expression for the optimal hedge ratio, analogous to (18), of

$$(24)\qquad b_{AGS} = \frac{-\lambda_U^{-1}\mu_F + \left[SF\sigma_1^2 - SF\sigma_1\left(\frac{\sigma_2\left(1-e^{-k\tau}\right)}{k}\right)\rho_{12} - F\sigma_3\left(\frac{\sigma_2\left(1-e^{-k\tau}\right)}{k}\right)\rho_{23} + F\sigma_1\sigma_3\rho_{13}\right]}{F^2\left[\sigma_1^2 - 2\sigma_1\sigma_2\rho\left(\frac{\left(1-e^{-k\tau}\right)}{k}\right)_{12} + \sigma_2^2\left(\frac{\left(1-e^{-k\tau}\right)}{k}\right)^2\right]}$$

This is a more general optimal hedge ratio that could be used by a hedger who does not plan to make delivery at the delivery location specified by the futures contract, or who is implementing a cross hedge.

Schwartz (1997) presents a one-factor model for pricing commodity contingent claims, hereafter referred to as the S97 model. Rather than arguing that a risk-free portfolio of a derivative and the underlying commodity can be formed, however, this model is developed by attaching a market price of (spot price) risk to the derivative. The S97 model does not therefore follow in the spirit of Kaldor, Working and Brennan's theory of storage as the GS model did, but instead follows Keynes (1930) and Hicks (1939) in emphasizing the role of risk and return in determining the value of contingent claims. In the S97 model, the spot price is assumed to follow the process

$$(25)\qquad dS = k(\mu - \ln S)Sdt + \sigma_1 Sdz_1.$$

where as before σ_1^2 is the instantaneous variance of changes in the natural logarithm of the spot price, and the log of the spot price reverts to level μ at rate *k*. The price of a futures contract must satisfy, as discussed by Schwartz, the partial differential equation

$$(26)\qquad \frac{1}{2}\sigma_1^2 S^2 F_{SS} + k(\mu - \lambda - \ln S)SF_S - F\tau$$

where λ is the market price of risk. Schwartz gives the solution as

$$(27)\qquad F(S,\tau) = \exp\left[e^{-k\tau}\ln S + \left(1-e^{-k\tau}\right)\left(\mu - \frac{\sigma_1^2}{2k} - \lambda\right) + \frac{\sigma_1^2}{4k}\left(1-e^{-2k\tau}\right)\right],$$

Following the discussion of hedging using the GS model, when *C* = *S*, we have the following process for the short hedger's portfolio under the S97 model

$$(28)\qquad dP = \mu_P dt + \sigma_P dz$$

where the drift is

$$(29)\qquad \mu_P = k(\mu - \ln S)S - bFe^{-k\tau}k\lambda$$

and instantaneous variance is

(30) $$\sigma_P^2 = S^2\sigma_1^2 + b^2F^2e^{-2k\tau}\sigma_1^2 - 2bSFe^{-k\tau}\sigma_1^2.$$

The optimal hedge ratio for the short hedger when $L = S$ is then

(31) $$b_{S97} = \frac{-\lambda_U^{-1}k\lambda + S\sigma_1^2}{Fe^{-k\tau}\sigma_1^2}.$$

Note that if we ignore the speculative component, the variance-minimizing hedge ratio is $(S/F)\exp(k\tau)$. Using equation (27), it is easy to see that this is identical to F_S. This demonstrates that the adaptation of contingent claims models for mean-variance hedging that we outline here can be considered a generalization of delta hedging. Augmenting the S97 for the case where $L \neq S$, again using equation (20) and specifying

(32) $$dB = \gamma(\beta - B)dt + \sigma_2 dz_2,$$

similar to before we find the diffusion followed by the hedger's portfolio is

(33) $$dP_A = \mu_{AP}dt + \sigma_{AP}dz$$

with drift

(34) $$\mu_{AP} = \gamma(\beta - B) + k(\mu - \ln S)S - bFe^{-k\tau}k\lambda$$

and instantaneous variance

(35) $$\sigma_{AP}^2 = \left[S^2\sigma_1^2 + \sigma_2^2 + 2S\sigma_1\sigma_2\rho_{12}\right] + b^2\left[F^2e^{-2k\tau}\sigma_1^2\right] - 2h\left[SFe^{-k\tau}\sigma_1^2 + Fe^{-k\tau}\sigma_1\sigma_2\rho_{12}\right].$$

The optimal hedge ratio for the short hedger when $L \neq S$ is then given by

(36) $$b_{AS97} = \frac{-\lambda_U^{-1}k\lambda + S\sigma_1^2 + \sigma_1\sigma_2\rho_{12}}{Fe^{-k\tau}\sigma_1^2}.$$

IV. Data, Parameter Estimation, and Parameter Inference

The data we use are week ending observations of the New York Mercantile Exchange (NYMEX) crude oil futures contracts, options on those futures, and the associated spot price. The futures and spot price data are observed over the period January 6, 1984 through June 21, 2002. We use option prices observed January 3, 1992 through June 21, 2002. Option prices were available before 1992, but trading volumes were not sufficient for the purposes outlined below. All data were provided by Commodity Research Bureau. We divide the data into three periods. The first time period, January 6, 1984 through December 27, 1991 (417 observations), is used strictly for parameter estimation. The second time period, January 3, 1992 through December 27, 1996 (261 observations), is used for both parameter estimation and the evaluation of in-sample hedging

effectiveness.[2] Out-of-sample hedging effectiveness is evaluated over the final time period, January 3, 1997 through June 21, 2002 (286 observations).

There is one NYMEX crude oil futures delivery per month. The price data for individual futures contracts were used to construct a rolling nearby futures series (*NEAR*) that is used in the parameter estimation and evaluation of hedging effectiveness. Where price changes were required, as in the unit root testing and VEC model estimation, care was taken to take changes of the individual futures series before selecting those changes that were nearby. That is to say, we use nearby futures changes (*NEARD*) rather than a changes in the nearby futures series (*DNEAR*). The latter series would result in roughly one out of every four observations being the composition of a change in a futures price and the spread between the expiring and new nearby futures prices (due to monthly contract expiration and the weekly observation frequency). Such a series has no natural interpretation in the context of hedging, and an uncertain (at best) interpretation in the context of time series econometrics. The *NEARD* series, however, contains no observations that are corrupted by futures spreads and is consistent with the futures price changes that an actual trader would realize. The differenced spot price series (*DS*) contains the usual first differences of the spot prices (*S*).

Following Gibson and Schwartz (1990), we employ the annualized one-month forward convenience yield when estimating the stochastic processes underlying the GS model. This is estimated using the price F^1 of a nearby futures contract and the price F^2 of the subsequent contract expiring using the following relation

$$\delta = r^1 - 12\ln\left(\frac{F^1}{F^2}\right) \tag{37}$$

where r^1 is the one-month forward riskless interest rate.

We first discuss the in-sample time series analysis. Augmented Dickey-Fuller (ADF) Tests for unit roots were carried out on all series over the in-sample estimation period (January 6, 1984 through December 27, 1996), with results presented in the first four rows of Table 1. Test results suggest non-stationary behavior, and differenced spot and nearby futures changes series are thus used for the remainder of the time-series estimation. We test for the presence of cointegration between *S* and *NEAR* using the Engle-Granger (1987) methodology.[3] Regressing *S* on *NEAR* and a constant results in the following potential cointegrating relation

$$ECT = S + 0.014 - 1.001NEAR\,. \tag{38}$$

An ADF test statistic on the recovered *ECT* series, presented in the last row of Table 1, strongly rejects the null hypothesis of a unit root, and we conclude that *S* and *NEAR* are indeed cointegrated.

Preliminary univariate analysis of the *DS* and *NEARD* series suggested the presence of GARCH effects as expected. Bollerslev's GARCH(1,1) process was then fitted to each series under the assumption of normality, with the results found in Table 2. Consistent with Baillie and Myers, no autoregressive terms in the mean equations were necessary to render the standardized

residuals free of autocorrelation, as evidenced by the reported Ljung-Box tests on the standardized residuals for up to 12^{th}-order autocorrelation. The sample skewness and kurtosis of the standardized residuals from each model suggest no significant deviation from normality. Asymptotic standard errors for the conditional variance equation parameter estimates confirm the presence of GARCH behavior in the series, and the Ljung-Box test on the squared standardized residuals indicates that the GARCH(1,1) specification adequately represents this behavior.

Based on the results of the univariate time series analyses, the multivariate VEC-GARCH(1,1) model given by equations (5) through (7) was fitted to the *DS* and *NEARD* series under the assumption or normality. The mean equations for each variable include the *ECT* recovered using equation (38). Schwarz (1978) information criterion was employed in the specification of the mean equations otherwise, and it was determined that neither constants nor autoregressive terms were desirable. Results are presented in Table 3. Residual diagnostics suggest no serious misspecification. All parameter estimates are significant at the 1% level. The speed of adjustment coefficients on the *ECT* suggest that deviations from the long-run equilibrium are subject to rapid correction, as expected given the frequency of futures deliveries used to construct the *NEARD* series. The parameters estimates associated with the conditional variance dynamics (A_{ii}, B_{ii}, W_i ; $i = 1,3$) are similar to those obtained in the univariate estimation. The parameter estimates associated with the conditional covariance dynamics (A_{22}, B_{22}, W_2) indicate substantial interaction between the two series.

The GS model parameters were estimated using an iterated seemingly unrelated regressions (SUR) procedure on the linear discrete approximations to equations (8) and (9). The resulting annualized parameter estimates are $\mu = -0.017$, $\alpha = 0.177$, $k = 9.183$, $\sigma_1 = 0.349$, $\sigma_2 = 1.157$, and $\rho_{12} = 0.431$. The large estimate of k suggests a high degree of mean-reversion in the convenience yield, and the large estimate of σ_2 suggests that it is highly volatile as well. We refer to this method of parameter estimation as estimating the stochastic differential equations (SDEs).

In order to implement the optimal hedging scheme outlined in section III, an estimate of the market price of convenience yield risk in the GS model is also needed. To accomplish this task, we follow Gibson and Schwartz (1990) by finding the least-squares fit of the futures pricing formula in the GS model to the market data. Specifically, for each available week-ending futures price observation for each delivery in the data set over the in-sample period, we collect the 5-tuple (F, S, δ, τ, r). We then use all such observations to find the value of λ that minimizes the sum of squared pricing errors implied by equation (12), using the estimates of parameters other than λ found by estimating the SDEs. The value of λ that we find is –0.132. As discussed in Gibson and Schwartz (1990), finding a negative price of convenience yield risk is consistent with the fact that the partial derivative of the futures price with respect to the convenience yield is negative.

In addition to estimating the SDEs, it is also possible to directly estimate the parameters of the term structure of volatility (TSV) in the GS model, using market data observed during the recent past. This provides a means by which the restrictive assumption of a constant TSV can be somewhat relaxed. The TSV for the GS model is given by

(39) $$\sigma_F(\tau;\sigma_1,\sigma_2,\rho_{12},k)=\sqrt{\sigma_1^2-2\sigma_1\sigma_2\rho_{12}\left(\frac{\left(1-e^{-k\tau}\right)}{k}\right)+\sigma_2^2\left(\frac{\left(1-e^{-k\tau}\right)}{k}\right)^2}$$

Computing the annualized sample standard deviations of observed futures log price changes for the most recent 2 months of daily observations for the *n*th nearby futures series provides us with a pair $(\hat{\sigma}_F,\tau)$ where τ is the average length of time until expiration. Collecting these pairs for the 12 nearest nearby futures price series provides 12 observations with which we find the values of $\sigma_1,\sigma_2,\rho_{12}$ and k that result in the best fit, in the least squares sense, of equation (39) to the market data. This exercise can be carried out at any point in time to arrive at a TSV that reflects more recent market activity, rather than a very long run average TSV found by estimating the SDEs. The estimated TSV might be thought of as the generalization of what is commonly referred to as "historical volatility." Rather than estimating the annualized volatility of only the spot price using a moving window of observations, however, the entire TSV is estimated. This provides a second means that a hedger might use arrive at the GS parameters needed to calculate his optimal hedge ratio. As an example, Figure 1 presents the GS term structure of volatility found by estimating the SDEs, and the TSV found by direct estimation on June 21, 1996 (a date chosen to illustrate an example of a high level of volatility in nearby futures). In both cases, the TSV is a decreasing function of time until maturity, as predicted by the Samuelson hypothesis. The functional form for the TSV in the GS model does not require this, however; gentle increases at longer times until maturity are permitted and are observed over some intervals in the data set.

In addition to the two parameter estimation methods discussed above, it is also theoretically possible to infer the TSV from observed futures option prices if a closed-form solution for those prices is available for a given model. In the case of the GS model, the value *C* at time *t,* of a European call option with strike price *X,* expiring at time T_1, on a futures contract with price *F,* expiring at time *T,* is given in Hilliard and Reis (1998) as

(40) $$C(F,X,t,T_1,T)=e^{-r(T_1-t)}\left[FN(d_1)-XN(d_1-v)\right]$$

where

(41) $$d_1=\frac{\ln(F/X)+0.5v^2}{v}$$

and

(42) $$v^2(t,T_1,T)=\sigma_1^2(T_1-t)-\frac{2\sigma_1\sigma_2\rho_{12}}{k}\left[(T_1-T)-\frac{e^{-k(T-T_1)}-e^{-k(T-t)}}{k}\right]$$
$$+\frac{\sigma_2^2}{k^2}\left[(T_1-t)-\frac{2}{k}\left(e^{-k(T-T_1)}-e^{-k(T-t)}\right)+\frac{1}{2k}\left(e^{-2k(T-T_1)}-e^{-2k(T-t)}\right)\right].$$

$N(d_1)$ represents the standard normal distribution function evaluated at d_1. To infer the TSV on a given date, the price for one approximately at-the-money option on each futures contract was collected (where available). All such available option prices and the corresponding values of F, X, r, T, and T_1, were then used in attempts to find the values of $\sigma_1, \sigma_2, \rho_{12}$ and k that provided the best least-squares fit of (the highly non-linear) equation (40). Just as the direct estimation of the TSV can be thought of as a generalization of "historical volatility", the option-implied TSV can be thought of as a generalization of "implied volatility". Unfortunately, in many cases as few as 5 observations were available for this task, and the inferred parameter values were often unreasonable. Given that this task could not be performed reliably with the available data, we do not use option-implied term structures of volatility for hedging in the context of the GS model.

We now turn to the estimation of the parameters of the S97 model. The linear discrete approximation of equation (25) was estimated over the in-sample estimation period using ordinary least squares, resulting in the following annualized parameter estimates: $\mu = 3.038$, $\alpha = 2.993$, $k = 1.334$, and $\sigma_1 = 0.347$. The market price of risk in the S97 model was estimated using a procedure analogous to that used to estimate the market price of convenience yield risk in the GS model. The resulting in-sample estimate of the market price of risk λ is 0.025. In addition to estimating the SDE of the S97 model, it is again possible to directly estimate the TSV. The TSV for the S97 model is given by

$$\sigma_F(\tau;\sigma_1,k) = e^{-k\tau}\sigma_1 . \tag{43}$$

Again pairs $(\hat{\sigma}_F, \tau)$ were collected for the 12 nearest nearby futures series, and the natural logarithm of $\hat{\sigma}_F$ was regressed on τ to arrive at least squares estimates for k and σ_1. In the case of the S97 model, we find that it is possible to reliably infer the TSV using observed futures option prices. The solution for European options on futures in the S97 model is given in Clewlow and Strickland (1999). The solution is equations (40) and (41) again, but equation (42) is replaced with

$$v^2(t,T_1,T) = \frac{\sigma_1^2}{2k}\left[e^{-2k(T-T_1)} - e^{-2k(T-t)}\right]. \tag{44}$$

The term structures of volatility estimated/inferred using the three methods outlined above for the S97 model on June 21, 1996 are presented in Figure 2. Note first that in all cases the TSV is a strictly decreasing function of time until maturity as dictated by its exponential decay functional form. The directly estimated TSV indicates a higher level of volatility at all times until maturity than the option-implied TSV. As it happened, the option-implied TSV indicated much higher levels of volatility one or two months earlier. This highlights the lagged effect that an increase in the general level of volatility will have on the TSV that is directly estimated using a moving window of historical data. The option-implied TSV, on the other hand, is calculated using data observed on a single day and can therefore adjust instantly to changes in market conditions.

Careful examination of the dynamics of the implied TSV, however, reveals a more subtle problem. We found the S97 option-implied TSV displayed a teetering behavior – an increase in

implied spot price volatility was generally accompanied by a decrease in the implied volatility of futures far from maturity and vice versa. Evidence of this is presented in Figure 3. Over a six week period, the implied spot price volatility increased roughly 8%, while the implied volatility of futures one year from expiration decreased about 4%. This phenomenon seems difficult to justify economically, and more likely result from the assumption of a constant TSV. In actual practice, option traders anticipate mean reversion in volatility levels - an increase in spot price volatility is likely to die out as time passes. As discussed in Hull and White (1987), the prices of options in a stochastic volatility environment should be a function of the expected levels of volatility over the life of the option. A short-term increase in spot price volatility has a large impact on the average level of volatility over the life of option that is nearing expiration, but a relatively small impact on the average level of volatility expected over the life of an option far from expiration. A significant increase in the premiums for options on nearby futures may therefore be accompanied by only a modest increase in the prices of options on distant futures. A significant increase in nearby option prices necessarily results in an increase in the value of σ_1 in the fitted TSV, but the rate of decay of volatility k must also increase if the distant option prices have not risen by much.

V. Hedging Effectiveness

We consider the problem of a hypothetical crude oil trader with mean-variance utility that wishes to take an optimal position in crude oil futures using equation (4). We assume that the cash position is 100,000 barrels, and that this position is hedged using the nearby futures contract. We further assume that the hedger's cash position corresponds to the specifications of the futures contract (i.e. $L = S$). Optimal hedge ratios in the time series hedging scheme are formed in each period by using the appropriate elements of the conditional variance-covariance matrix H_t. When employing partial equilibrium hedging, hedge ratios are formed using either equation (18) or equation (31) after any appropriate parameter estimation or inference. Two methods of parameter estimation are devised above for the GS model: 1) simply estimating the SDEs and 2) directly estimating the TSV each time a new hedge ratio is formed. These two methods of parameter estimation are also available when using the S97 model, and we additionally are able to infer the TSV from futures option prices. We thus have five competing partial equilibrium hedging schemes.

The time paths of hedge ratios generated by the VEC-GARCH and GS models during the last 18 months of the in-sample period are presented in Figure 4. As expected, the hedge ratios generated by the GS model with SDE parameters estimates are fairly stable relative to those generated by the VEC-GARCH and GS model employing a freshly estimated TSV each period. Nonetheless, the paths of the VEC-GARCH and GS with SDE parameter estimates are similar – steady in late 1995, then dipping in the spring and summer of 1996 and then increasing in late 1996. Over the portion of the in-sample period for which we evaluate hedging effectiveness, the correlation coefficient between these two models' hedge ratios is 0.53, while the correlation between the hedge ratios from the VEC-GARCH and the GS model with an estimated TSV is –0.10. The time paths of the optimal hedge ratios generated by the three S97 models over the same time period are presented in Figure 5. All three consistently follow a saw tooth pattern due to the functional form of the TSV in the S97 model. Ignoring the speculative component of equation (31), and assuming the ratio of the spot price to the futures price is approximately one,

the optimal hedge ratio is then approximately $\exp(k\tau)$. This is greater than one before futures expiration, and decays to one at the time of expiration. As one might expect after examining Figures 4 and 5, the S97 hedge ratios are highly correlated with one another, but not with the GS or VEC-GARCH hedge ratios.

To evaluate in-sample hedging effectiveness, the realized levels of utility, based on the realized price changes and conditional variances and covariances from the VEC-GARCH model, were evaluated for each week over the period January 3, 1992 through December 27, 1996. The average level of utility was then calculated for each of the six hedging models. Table 4 presents these averages, along with the incremental utility increase (IUI) that are realized for each model versus holding only the unhedged cash position. Utility increases are large in all cases, demonstrating the excellent hedging performance of NYMEX crude oil contract in the present context. The VEC-GARCH model delivers the greatest utility increase, at 85.16%. Among the partial equilibrium models, there is no clear-cut pattern. Neither the GS nor S97 models' performance dominates the other. Also, neither of the two available methods of parameter estimation is clearly superior. Hedging using the GS model with estimated SDEs results in hedging performance that is very similar to hedging using the S97 model with estimated term structures of volatility. The S97 model with option-implied terms structures of volatility provides the second worst hedging performance, despite the attempt to glean insight into the future volatility conditions expected by option traders.

Previous optimal hedging literature considers not only in-sample hedging effectiveness, but stresses the need to evaluate out-of-sample hedging effectiveness as well. This provides a fair test of how an optimal hedging scheme is likely to perform in real-world conditions. To evaluate out-of-sample hedging effectiveness, we re-estimated each model each period using all available data at that point in time for each of the models, and then used each to make one period ahead forecasts of the components of the hedger's optimal hedge ratio. The resulting utilities in each period were assessed using the ensuing actual price changes in the following week and the conditional variances and covariances recovered from a final VEC-GARCH model estimated using the entire data set. Again the utilities in each period were averaged and incremental increases in utility over that of the unhedged portfolio were calculated for each hedging model. Results are presented in Table 5. These results are very similar to those found in the in-sample period. The VEC-GARCH model results in the largest utility increase, but delivers a utility increase almost 5% greater than the second place model versus only about 2% in-sample. Again the S97 model with estimated TSV and the GS model with estimated SDEs deliver similar performance, roughly tying for second place. The remaining three models again share the dishonor of being the three worst performing.

To determine if the superior hedging effectiveness of the VEC-GARCH model was attributable to superior futures price forecasting (associated with the speculative component of the hedge ratio) or the superior variance and covariance forecasting, the out-of-sample forecasts of nearby futures one-week price changes were evaluated. All models delivered very similar root mean squared forecast errors (RMSEs), however the VEC-GARCH model provided the worst forecasts. The RMSE of the VEC-GARCH forecasts was \$1.161 per barrel, while the partial equilibrium models' RMSEs were tightly distributed around an average of \$1.154 per barrel. It

therefore appears that the superior hedging performance of the VEC-GARCH model is due entirely to superior modeling of conditional variance and covariance dynamics.

Overall, the VEC-GARCH hedging model, which allows time-varying variances and covariance, provides the best hedging performance, despite producing the most variable hedge ratios (as measured by sample standard deviation). The partial equilibrium models' hedge ratios are less variable, but perform worse. This suggests that the hedge ratios generated by the partial equilibrium models are not sufficiently reflecting changes in volatility conditions. The cause of the inferior performances of the partial equilibrium hedging models thus appears to be the unrealistic assumption of a constant TSV. Attempts to compensate for this shortcoming by frequently estimating or inferring the TSV do not result in consistently improved hedging effectiveness, and in no case is the performance of the VEC-GARCH model matched. Estimating the TSV suffers from the problem of employing a moving window of historical data, and any change in volatility conditions is reflected with somewhat of a lag. Inferring the TSV from futures options prices (only practical for the S97 model) is still done in a constant TSV context, and suffers from the teetering effect described earlier. All methods of updating the parameters of the term structures of volatility in the partial equilibrium models also come at the expense of a significant increase in computational complexity.

In a sense, the hedging problem formulated here was the easiest possible for the partial equilibrium models. The assumption was made that the hedger's cash position corresponded with the futures contract specifications (i.e. $L = S$). We thus employed the optimal hedge ratios in equations (18) and (31) rather than those from the augmented models in equations (24) and (36). For many hedgers this will not be the case, and the use of the augmented models would be necessary. This would likely result in hedging performance that fell further short of that of the VEC-GARCH model, for the following reason. The use of an augmented partial equilibrium models would add another layer of constant variance-covariance assumptions – likely exacerbating the problem that led to the poor performance when $L = S$. On the other hand, the case where $L \neq S$ presents no special problem for the time series model, as one would simply employ the appropriate local cash price series rather than the spot price series, and proceed as usual with a model that fully incorporates conditional variance and covariance dynamics.

VI. Conclusion

This paper compares the performances of time series and partial equilibrium based optimal hedging models for trader that is long in a cash commodity market, and maximizes mean-variance utility using futures contracts. We find that the time series approach delivers superior hedging performance to that of each of the other models considered. This appears to be due to the partial equilibrium models' unpalatable assumption of a constant volatility term structure. The constant volatility term structure framework hampers even the seemingly promising technique of inferring option market participants' expectations regarding future volatility conditions.

This research has considered only a single type of derivative, however. These results suggest that the attractiveness of employing a simple partial equilibrium model (i.e. one that does not incorporate stochastic volatility) when hedging a commodity market cash position using futures

contracts (or *vice versa*) is questionable. Few would doubt the usefulness of partial equilibrium models in hedging a position in a derivative with a non-linear payoff function (e.g. an option), however. The conclusion then is that different types of hedging models are suited to different tasks, and the best approach in still other situations (e.g. hedging with swaps) is uncertain. Furthermore, this research has considered only a single hedging objective. When commodity producers or consumers purchase options they generally think of them as being similar to insurance contracts. This suggests that they are maximizing utility of a form other than that employed here (and in much of the optimal hedging literature). These issues illuminate the necessity of further research.

References

Ballie, R. T.; and Myers, R. J. (1991): "Bivariate GARCH Estimation of the optimal Commodity Futures Hedge," *Journal of Applied Econometrics,* 6:109-124.

Black, Fischer (1976): "The Pricing of Commodity Contracts," *Journal of Financial Economics*, 3(1/2):167-179.

Black, Fischer and Scholes, M. (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.; Engle, R. F.; and Wooldridge, J. M. (1988): "A Capital Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy*, 96:116-131.

Brennan, Michael J. (1958): "An Approach to the Valuation of Uncertain Income Streams," *American Economic Review*, 48:50-72.

Brennan, Michael J. (1991): "The Price of Convenience and the Valuation of Commodity Contingent Claims," in *Stochastic Models and Option Values,* D. Lund and B. Oskendal (ed.), North-Holland, Amsterdam.

Brennan, Michael J. and Schwartz, Eduardo S. (1985): "Evaluating Natural Resource Investments," *The Journal of Business*, 58(2):135-157.

Cecchetti, S. G.; Crumby, R. E.; and Figlewski, S. (1988): "Estimation of the Optimal Futures Hedge," *Review of Economics and Statistics*, 70:623-630.

Clewlow, Les; and Strickland, Chris (1999): "Valuing Energy Options in a One Factor Model Fitted to Forward Prices," Working paper, University of Technology, Sydney, Australia.

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

Engle, R. F. and Granger, Clive W. J. (1987): "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55:251-276.

Gagnon, Louis; Lypny, Gregory J. and McCurdy, Thomas H. (1998): "Hedging Fogeign Currency Portfolios," *Journal of Empirical Finance,* 5:197-220.

Gibson, Rajna and Schwartz, Eduardo S. (1990): "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims," *Journal of Finance,* 45(3):959-976.

Gibson, Rajna and Schwartz, Eduardo S. (1991): "Valuation of Long Term Oil-Linked Assets," in *Stochastic Models and Option Values,* D. Lund and B. Oskendal (ed.), North-Holland, Amsterdam .

Gibson, Rajna and Schwartz, Eduardo S. (1993): "The Pricing of Crude Oil Futures Options Contracts," *Advances in Futures and Options Research,* 6:291-311.

Haigh, Michael S. and Holt, Mathew T. (2000): "Hedging Multiple Price Uncertainty in International Grain Trade," *American Journal of Agricultural Economics*, 65:603-605.

Hicks, J. R. (1939): *Value and Capital*, Oxford: Oxford University Press.

Hilliard, Jimmy E.; and Reis, Jorge (1998): "Valuation of Commodity Futures and Options Under Stochastic Convenience Yields, Interest Rates, and Jump Diffusions in the Spot," *Journal of Financial and Quantitative Analysis*, 33(1):61-86.

Hull, John C.; and White, A. (1987): "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42:281-300.

Johansen, S. (1988): "Statistical Analysis of Cointegrating Vectors," *Journal of Economics Dynamics and Control*, 52:169-210.

Johnson, Leland. L. (1960): "The Theory of Hedging and Speculation in Commodity Futures," *Review of Economic Studies,* 27:139-151.

Kahl, Kandice H. (1983): "Determination of the Recommended Hedging Ratio," *American Journal of Agricultural Economics,* 65:603-605.

Kaldor, Nicholas (1939): "Speculation and Economics Stability," *Review of Economic Studies*, 7:1-27.

Keynes, J. Maynard (1930): *A Treatise on Money,* London: Macmillan.

Kroner, K. F. and Sultan, J. (1993): "Time Varying Distributions and Dynamic Hedging with Foreign Currency Futures," *Journal of Financial and Quantitative Analysis,* 28:535-551.

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

Myers, R. J. (1991): "Estimating Time-Varying Optimal Hedge Ratios On Commodity Futures Markets," *The Journal of Futures Markets*, 11:39-53.

Peck, Anne E. (1975): "Hedging and Income Stability: Concepts, Implications, and an Example," *American Journal of Agricultural Economics,* 57:410-419.

Ross, Stephen A. (1978): "A Simple Approach to the Valuation of Risky Streams," *The Journal of Business,* 51(3):453-475.

Samuelson, Paul (1965): "Proof that Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review*, 6(2):41-49.

Schwartz, E. S. (1997): "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging," *Journal of Finance,* 52:923-974.

Schwarz, G. (1978): "Estimating the Dimensions of a Model," *Annals of Statistics*, 6:461-464.

Stock, J. H. (1987): "Asymptotic Properties of Least-Squares of Co-integrating Vectors," *Econometrica*, 55:1035-1059.

Working, Holbrook (1949): "The Theory of Price of Storage," *American Economic Review*, 39:1254-1262.

Table 1

Results from Augmented Dickey-Fuller (ADF) tests for the presence of unit roots (test is on the estimated coefficient θ_1 from the following prototype model: $\Delta X_t = \theta_0 + \theta_1 + \sum_{k=1}^{K} \beta_k \Delta X_t - k$).[a]

Series	K	θ_1
Spot Price	0	-2.907
Spot Price Changes	0	-28.819
Nearby Future Price	2	-2.973
Nearby Futures Price Changes	1	-16.474
ECT	3	-11.458

[a] The critical value –3.43 (1%) is given in Fuller (1976). The optimal lag length (*K*) was chosen using the Schwarz (1978) information criterion.

Table 2

Parameter estimates and residual diagnostics for the univariate GARCH(1,1) models:

$$x_t = \mu + \varepsilon_t$$
$$\varepsilon_t \mid \Omega_{t-1} \sim N(0, h_t^2)$$
$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}$$

	DSPOT		*NEARD*	
μ	-0.006	(0.027)	0.057	(0.028)
ω	0.051	(0.013)	0.089	(0.019)
α	0.184	(0.037)	0.200	(0.038)
β	0.764	(0.042)	0.679	(0.052)
Log-likelihood	-197.416		-182.392	
m^3	-0.330		-0.069	
m^4	2.180		2.504	
$Q(12)$	17.541	(0.130)	17.249	(0.140)
$Q^2(12)$	6.745	(0.874)	9.581	(0.653)

[b]The numbers in parenthesis beside the parameter estimates are asymptotic standard errors. m^3 and m^4 are the sample skewness and sample kurtosis, respectively, of the standardized residuals. $Q(12)$ and $Q^2(12)$ denote Ljung-Box test statistics for 12th-order autocorrelation in the standardized and squared standardized residuals, respectively, with the numbers in parenthesis being the associated p-values.

Table 3

Parameter estimates and residual diagnostics[c] for the multivariate GARCH(1,1) model:

$$\Delta y_t = \alpha ECT_{t-1} + \varepsilon_t; \quad \Delta y_t = \left(DSPOT_t, NEARD_t\right)^T$$

$$\varepsilon_t \mid \Omega_{t-1} \sim N(0, H_t)$$

$$\text{vech}(H_t) = W + A\ \text{vech}(\varepsilon_{t-1}\varepsilon_{t-1}^T) + B\ \text{vech}(H_{t-1})$$

α_1	-0.572	(0.102)
α_2	0.299	(0.106)
W_1	0.078	(0.006)
W_2	0.069	(0.001)
W_3	0.070	(0.005)
A_{11}	0.121	(0.011)
A_{22}	0.093	(0.009)
A_{22}	0.096	(0.011)
B_{11}	0.761	(0.014)
B_{22}	0.785	(0.011)
B_{33}	0.786	(0.018)
Log-likelihood	412.628	
DSPOT equation		
m^3	-0.264	
m^4	2.417	
$Q(12)$	17.157	(0.144)
$Q^2(12)$	11.972	(0.448)
NEARD equation		
m^3	-0.150	
m^4	2.795	
$Q(12)$	17.067	(0.147)
$Q^2(12)$	15.493	(0.216)

[c]The numbers in parenthesis beside the parameter estimates are asymptotic standard errors. m^3 and m^4 are the sample skewness and sample kurtosis, respectively, of the standardized residuals. $Q(12)$ and $Q^2(12)$ denote Ljung-Box test statistics for 12th-order autocorrelation in the standardized and squared standardized residuals, respectively, with the numbers in parenthesis being the associated p-values.

Table 4

In-sample hedging effectiveness: average levels of utility and incremental utility increases (IUI)[d]

	Average Utility	I.U.I
Unhedged	-5.648E+09	
S97 (estimated SDE)	-1.018E+09	81.98%
S97 (estimated TSV)	-9.537E+08	83.11%
S97 (inferred TSV)	-1.058E+09	81.26%
GS (estimated SDEs)	-9.498E+08	83.18%
GS (estimated TSV)	-1.143E+09	79.77%
VEC-GARCH	-8.380E+08	85.16%

[d]IUI is the percentage increase in average utility over that realized when not hedging.

Table 5

Out-of-sample hedging effectiveness: average levels of utility and incremental utility increases (IUI)[e]

	Average Utility	I.U.I
Unhedged	-9.370E+09	
S97 (estimated SDE)	-2.908E+09	68.97%
S97 (estimated TSV)	-2.723E+09	70.93%
S97 (inferred TSV)	-2.989E+09	68.10%
GS (estimated SDEs)	-2.773E+09	70.41%
GS (estimated TSV)	-2.954E+09	68.48%
VEC-GARCH	-2.281E+09	75.66%

[e]IUI is the percentage increase in average utility over that realized when not hedging.

Figure 1

The term structure of volatility (TSV) of crude oil in the Gibson-Schwartz (GS) model using different parameter estimation techniques.

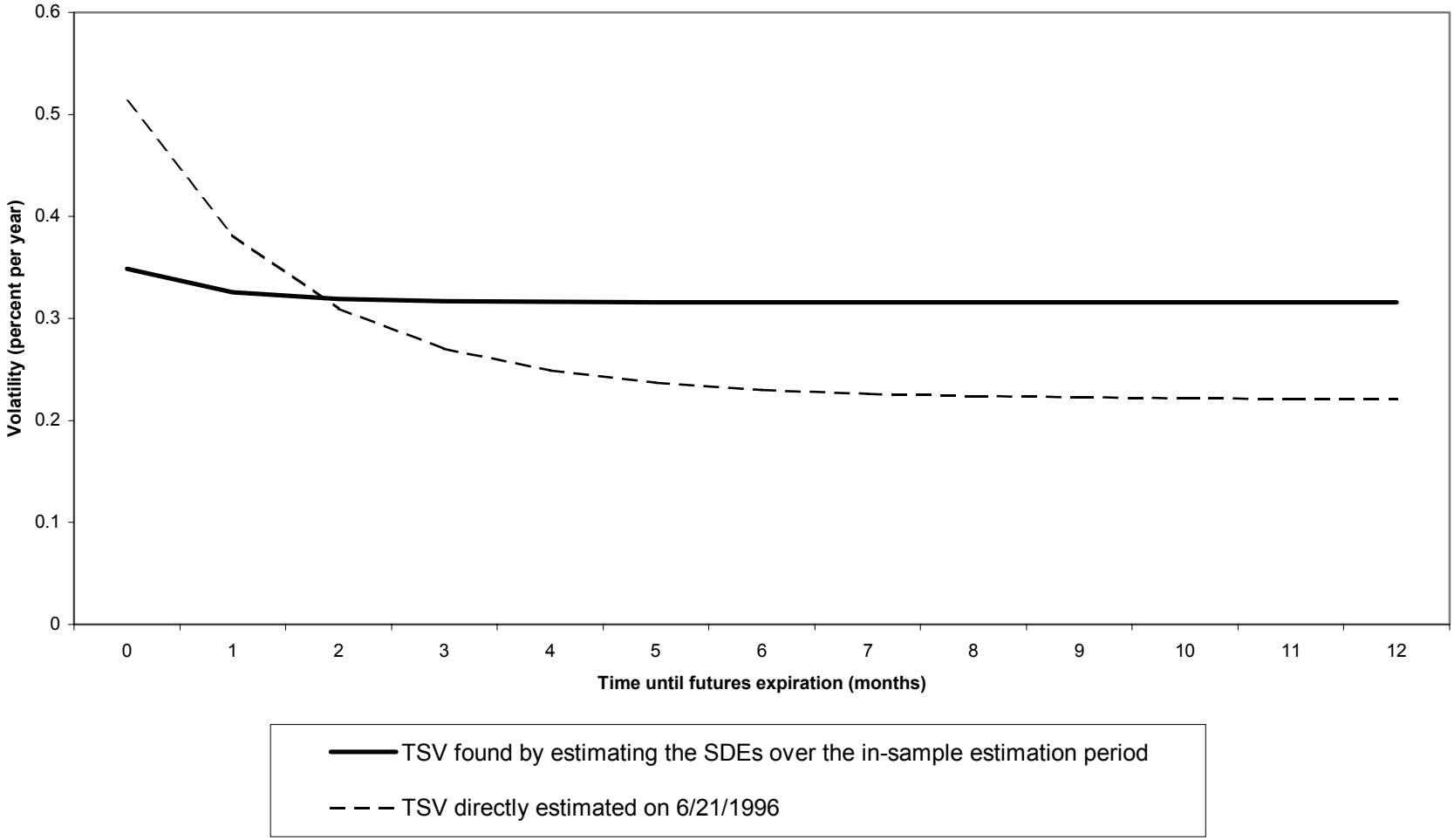


Figure 2

The term structure of volatility (TSV) of crude oil in the Schwartz 1997 (S97) model using different parameter estimation and inference techniques.

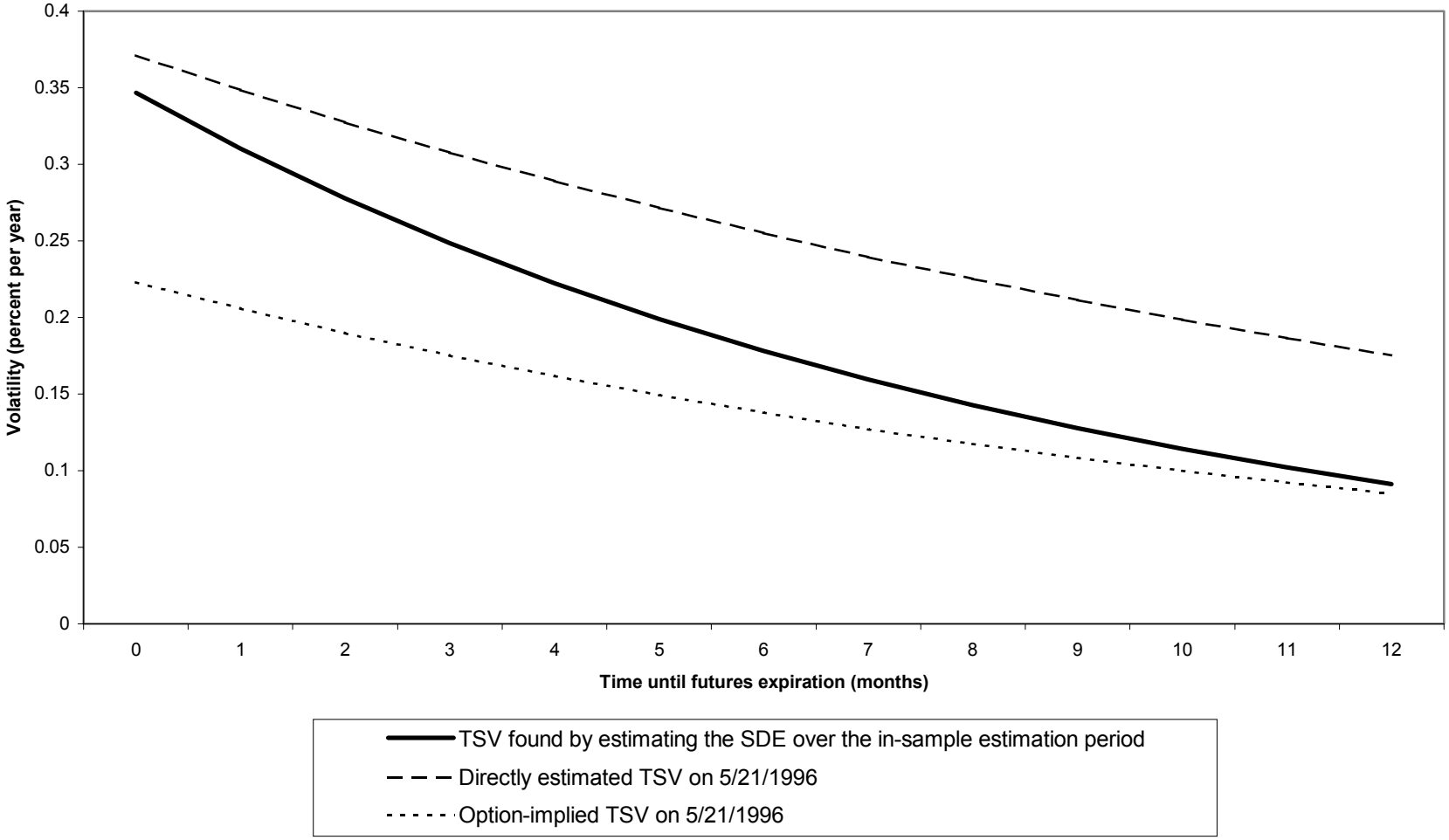


Figure 3

The option-implied term structure of volatility (TSV) of crude oil in the Schwartz 1997 (S97) model observed on two dates.

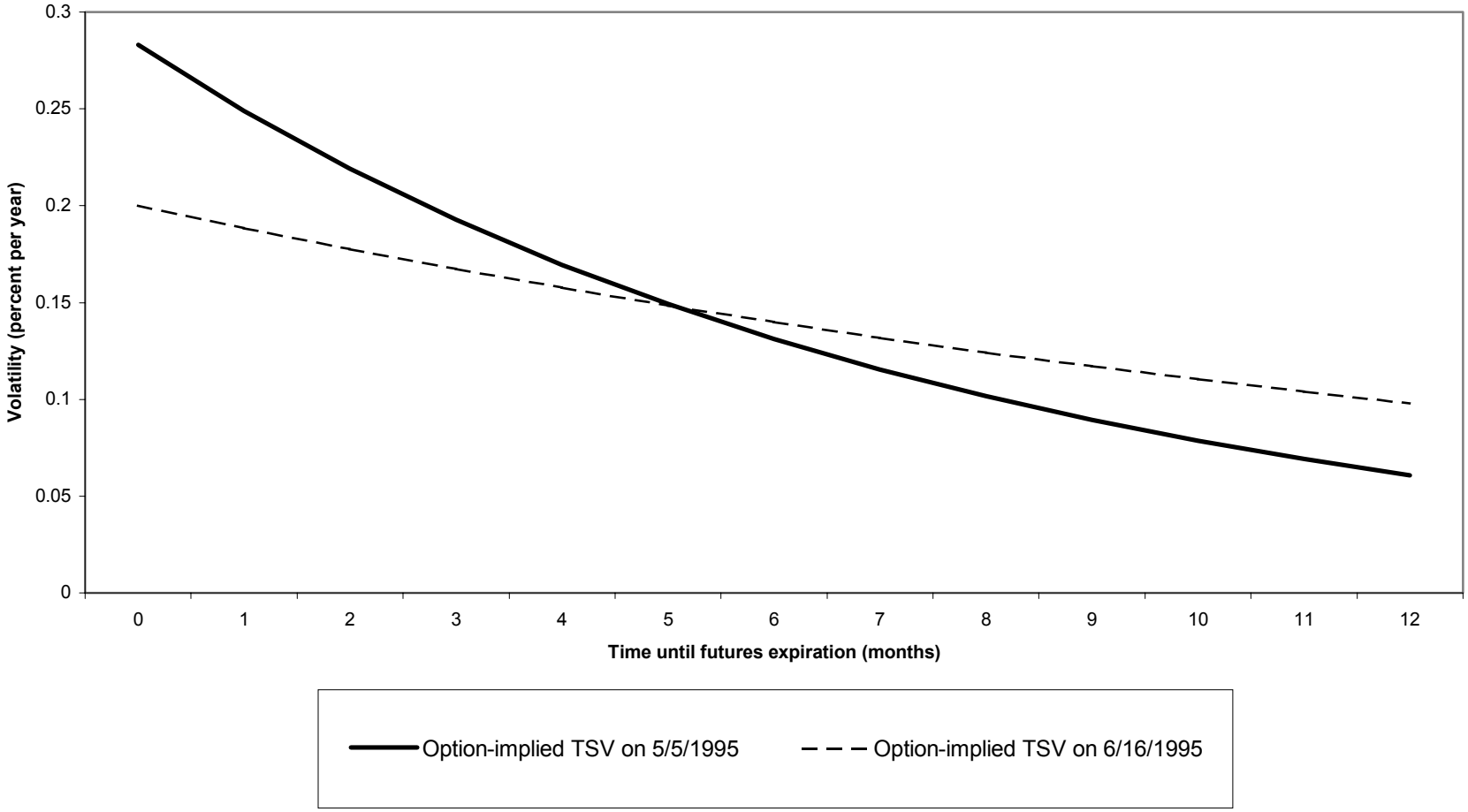
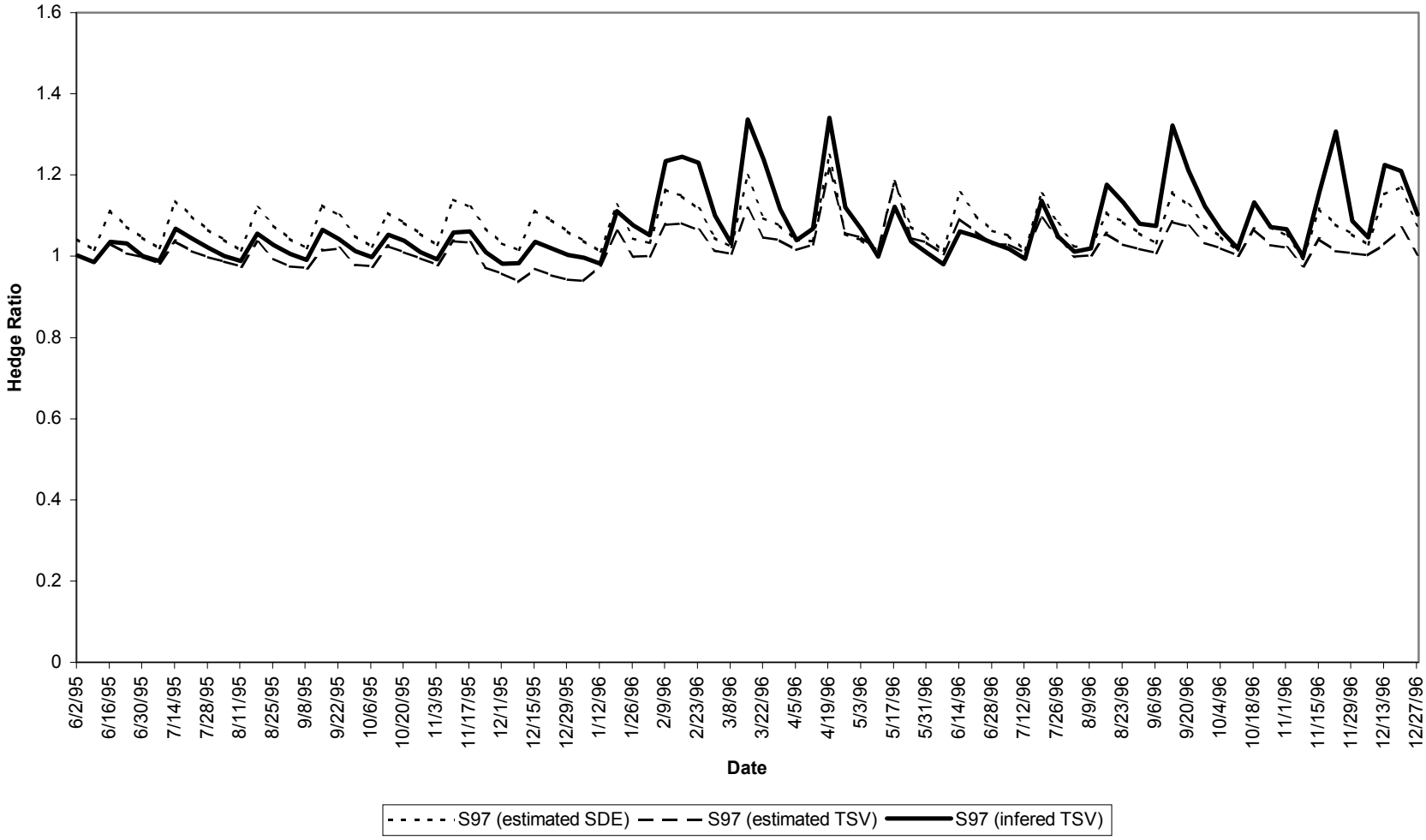


Figure 4

Partial time paths of the in-sample optimal hedge ratios generated by the GARCH and GS models

Figure 5

Partial time paths of the in-sample optimal hedge ratios generated by the S97 models



Endnotes

[1] Gibson and Schwartz (1993) and Schwartz (1997) also publish formulas for the price of a futures contract in the GS model, but these formulas appear to suffer from typographical errors as they do not seem to solve partial differential equation (11).

[2] In-sample hedging effectiveness is not evaluated over the entire in-sample estimation period because option trading volume was insufficient to carry out the inference of the term structure of volatility in the S97 model.

[3] Unfortunately, available implementations of Johansen's (1988) cointegration methodology perform data differencing automatically when forming the vector auto-regression. In the present context, given the series *NEAR,* an implementation of the Johansen methodology would then generate and subsequently employ the unacceptable differenced nearby series *DNEAR* described above. Hypothesis testing on the coefficients of the cointegrating vector within the Engle-Granger framework can be misleading (Stock 1987), however we carry out no such testing. The Engle-Granger methodology does provide a consistent estimate of a single cointegrating vector, however, which is all that we require here.