Hedging Long-Term Exposures with Multiple Short-Term Futures Contracts

Anthony Neuberger
London Business School

This article analyzes the problem facing an agent who has a long-term commodity supply commitment and who wishes to hedge that commitment using short-maturity commodity futures contracts. As time evolves, the agent has to roll the hedge as old futures contracts mature and new futures contracts are listed. This gives rise to hedge errors. The optimal hedging strategy is characterized in a world where contracts of several different maturities coexist. The strategy is independent both of the agent's risk aversion and, under certain conditions, of beliefs about expected returns from holding futures contracts. The methodology is compared with approaches based on dynamic models of the term structure. It is tested on data from the oil futures market.

How well can a long-term exposure to commodity prices be hedged using traded commodity futures? Futures contracts provide an excellent tool for risk management. They are highly leveraged, they have low transaction costs, and counterparty risk is small. They are traded on markets which are generally transparent and closely regulated. But most contracts, and in particular the most liquid contracts, tend to be of short maturity. The investor who wishes to hedge a long-term exposure has to keep rolling the hedge into longer-dated contracts. If the prices of the contracts at the time of the rollover do not conform exactly to the particular model being used, errors will arise when positions in one contract are closed out and positions in another are opened up. Rollover risk becomes an important source of hedging error. A model that is used for the design of long-term hedging strategies needs to capture the cross-sectional features of futures prices.

The model presented in this article starts out from an assumption about the relationship between the price of the longest-dated contract in the market and the contemporaneous prices of shorter-dated contracts. The parameters of the model are estimated from cross-sectional rather than time-series regressions. In this way the model directly reflects the risks attendant on rolling over contracts. It turns out that for the particular problem addressed

I am grateful for valuable help and comments from Wayne Ferson (the editor), Adam Apter, Dick Brealey, Michael Brennan, Mark Britten-Jones, Mark Grinblatt, Stephen Hall, Julian Franks, Michael Rockinger, an anonymous referee, and also from participants at the 1995 European Finance Association meeting, the 1996 Western Finance Association, and the 1996 Corporate Risk Management Conference at UCLA. I am also grateful to Rajiv Guha for substantial research assistance. Address correspondence to Anthony Neuberger, London Business School, Sussex Place, Regents Park, London NW1 4SA, UK, or e-mail: aneuberger@lbs.ac.uk.

The Review of Financial Studies Fall 1999 Vol. 12, No. 3, pp. 429–459

here, it is then unnecessary to make any explicit assumptions about the dynamics of prices.

The basic setup posits an agent who has written a long-term fixed price supply contract for a commodity. The agent is assumed to have access to a futures market where contracts of different maturities are traded, but where the longest maturity matures before the commitment falls due. The desire for hedging comes because the agent is risk averse and wishes to maximize expected utility. The model shows how an optimal hedge can be constructed.

The equating of optimal hedging with utility maximization may seem rather restrictive, particularly in light of the growing literature on why firms hedge [see Smith and Stulz (1985), Bessembinder (1991), and Froot, Scharfstein and and Stein (1993); see also Mello and Parsons (1995) in the specific context of commodity price hedging]. Other motives for corporate hedging — such as minimizing the variance of cash flows, taxable income or accounting earnings, or avoidance of financial distress — may all be important, and will in general lead to somewhat different hedging strategies. Yet even in these cases the characteristics of an economic value hedge, such as is considered here, are important. To the extent that an asset or liability can be hedged, it can also be valued. The hedging strategy creates a synthetic long-term forward price; knowledge of that price is useful in many different contexts.

Previous research on the use of futures for hedging [notably Ederington (1979), Baesel and Grant (1982), McCabe and Franckle (1983), and Grant (1984)] restricts itself to a world in which just one futures contract trades at a time. But in most futures markets, many different maturities trade simultaneously. It seems plausible that hedging could be greatly improved by using several maturities simultaneously. We show that this is indeed the case.

More recent work in this area has built on Gibson and Schwartz (1990) and Brennan (1991). In particular, Gibson and Schwartz model the term structure of commodity prices by assuming that the spot price and the convenience yield from holding the physical commodity follow a joint diffusion process of specified form. By imposing no arbitrage conditions they obtain a partial differential equation which all contingent claims must satisfy. Provided that the set of hedging instruments is at least as large as the number of state variables (two in this case), any other contingent claim can be valued exactly and hedged perfectly.

This approach is powerful because it embeds the hedging problem in the much wider problem of creating a consistent framework for pricing and hedging all contingent claims. But it does have certain drawbacks. First, the model assumes that all futures are fairly priced relative to each other, and hence implicitly assumes away rollover risk. Second, since any two futures provide a perfect hedge, the model cannot suggest which particular maturities should be used if there are more than two.

Ross (1995) has a model of commodity prices that admits many factors. He recognizes that in principle it would be possible to hedge all the factors, and hence hedge any long-term commitment perfectly, if there were enough contracts with different maturities and if all the parameters of the model were known. He points out the difficulty in estimating the model with sufficient precision to design perfect hedging strategies, and argues that in practice it is not possible to hedge away all risk. He therefore models the problem in an incomplete market setting by restricting the number of maturities traded to be less than the number of factors; in particular he concentrates on strategies that use just one maturity at a time.

In our model, incompleteness is built in from the outset. The state variables are the prices of traded futures contracts. Futures contracts have a finite life. The price at which a contract first trades is a stochastic function of the prices of other contracts. This opening price provides new information that cannot be hedged using existing securities. The key assumption of the model is that the expected value of the opening price is a linear function of the prices of the other contracts. We impose no other restrictions on the dynamics of futures prices.

This article is laid out as follows: in Section 1 the basic model is set out and solved for the special case where futures prices follow a martingale. Section 2 relaxes the martingale assumption and asks whether agents' willingness to write long-term supply commitments is affected by their beliefs about the behavior of futures prices. In Section 3 the model is tested using oil futures data. Section 4 concludes.

1. The Basic Model

1.1 The setup

For convenience, the underlying commodity will be referred to as oil, but the approach applies to commodities generally. The problem to be examined is this: an agent contracts to sell one barrel of oil in T months time[1] at a price K and chooses to hedge her exposure in the traded oil futures market. What is her optimal trading strategy? How well can she hedge?

The model is set in continuous time. Time is measured in months. The supply contract is written at time 0 and matures at time T. At the beginning of each month a new futures contract starts trading. It matures exactly $N(<T)$ months later, so at any time during the month there are exactly N futures contracts trading; as one matures a new one is introduced.[2] $F(t, n)$ is the price at time t of the futures contract which matures at time n, where n is an integer. This price exists for all integers n in the range $[t, t+N]$. $\mathbf{F}_t$ denotes the $N \times 1$ vector of futures prices traded at time t, and F_i is the ith element

[1] The extension to flow contracts, where the agent contracts to deliver a certain flow of oil over a period, is done in Section 1.3.

[2] It is convenient to assume that the date in the month when new contracts are introduced is the same as the date on which contracts mature, but nothing in the model hangs on this.

of the vector where F_1 is the price of the longest maturity trading and F_N is the price of the near-month contract.

The sequence of events at the turn of the month is as follows. The new N-month contract opens. The futures prices at other maturities may adjust, and the agent may trade. Then all positions in the near-term contract are closed out at the spot price. Thus, at the very start of a month, when t is an integer, there are instantaneously $N + 1$ contracts trading — the maturing contract, the just-opened long-dated contract, and $N - 1$ continuing contracts. By convention the vector of futures prices $\mathbf{F}_t$ at the start of the month consists of the spot and the $N - 1$ continuing contracts.

The following assumptions are made:

Assumption 1. *There are no market frictions. There are no transaction costs, taxes, spreads, or market impacts from trades carried out by the agent. Contracts are cash settled immediately on a continuous basis; no other margins are required. Contracts are infinitely divisible.*

Assumption 2. *There is no risk of default. Both parties to the forward oil supply contract will certainly honor their commitments, as will the parties to futures market transactions.*

Assumption 3. *The agent can borrow or lend without limit at the riskless rate of interest.*

Since the focus of this article is on commodity price risk, interest rate risk is ignored. Without further loss of generality[3] we assume:

Assumption 4. *The riskless interest rate is zero.*

A futures position at time t is represented by the $N \times 1$ vector $\Delta = \{\delta_1, \ldots, \delta_N\}$. The agent's cash position at time t is X_t; X_0 is zero. At time T the agent buys oil on the spot market, delivers it to the customer, and receives K. The spot price equals the price of the futures contract which matures at that time, $F(T, T)$. The agent's terminal wealth W is $K + X_{\mathrm{T}} - F(T, T)$.

The effect of these assumptions is that the agent is free to choose a hedge strategy represented by a vector process Δ_t. The change in the agent's cash position at time t is the product of the futures position and the change in futures prices:

$$\begin{aligned} \mathrm{d}X &= \Delta' \cdot \mathrm{d}\mathbf{F} \\ \text{so } X_T &= \int_{t=0}^{T} \Delta_t' \cdot \mathrm{d}\mathbf{F}_t \end{aligned} \tag{1}$$

[3] The extension to nonzero interest rates is covered in Section 1.3.

If $N \geq T$ all risk can be removed by holding one contract of the T-maturing futures. This would lock in a cost of $F(0, T)$ and so ensure terminal wealth of $W = K - F(0, T)$. But in general $T > N$, and it is impossible to construct a perfect hedge.

The last two assumptions concern the behavior of futures prices:

Assumption 5. *Futures prices follow a martingale.*

Assumption 6. *At the beginning of a month, when t is an integer, a new N-month contract is listed. The price at which it is expected to open, given the prices of the other contracts, is*

$$
\begin{aligned}
E\left[F(t, t+N) \mid \mathbf{F}_t\right] &= \alpha + \sum_{i=1}^{N} \beta_i F(t, t+N-i) \\
&= \alpha + \beta' \cdot \mathbf{F}_t
\end{aligned} \tag{2}
$$

where β is a constant $N \times 1$ vector and α is a constant. The innovation is represented by ε_{t}, so $\varepsilon_{\mathrm{t}} \equiv F(t, t+N) - \mathrm{E}\left[F(t, t+N) \mid \mathbf{F}_t\right]$.

The martingale assumption (Assumption 5) will be relaxed in Section 2. Assumption 6 is critical to the model. It says that the price at which a new contract trades, given the prices of all other traded contracts, is a linear function of the prices of those traded contracts plus a noise term. The implications of this assumption, its consistency with other models of commodity prices, and its empirical validity will be tested against oil futures prices in Section 3. We will refer to Equation (2) as the *forecasting equation*.

1.2 The optimal hedge

We show that there exists a unique hedging strategy, independent of the agent's utility function, which dominates all others in the sense of second-order stochastic dominance. The derivation of the optimal hedge is in two steps. We show first that the expected cost of delivery is a linear function of the current term structure of futures prices. The second step is to prove that the optimal hedge is the one that most closely hedges changes in the expected cost of delivery. All proofs are in Appendix A.

Lemma 1. *The expected cost of delivery is a linear function of current futures prices:*

$$
\mathrm{E}\left[F(T, T) \mid \mathbf{F}_t\right] = \alpha^{(n)} + \beta^{(n)} \cdot \mathbf{F}_{\mathrm{t}} \tag{3}
$$

where n is the largest integer $\leq T - t$, and $\alpha^{(\mathrm{n})}$ and $\beta^{(\mathrm{n})}$ are functions of the original α and β, as set out in Appendix A.

Call the right-hand side of Equation (3) Y_t. Assumption 6 extrapolates from the current set of N futures contracts to give a price for the next futures contract to be listed. The equation can be applied repeatedly to extend the term structure a further month. Y_t can then be interpreted as an estimate of the price at which a contract maturing at time T would trade if it were listed at time t. With this interpretation, Lemma 1 says that the expected future spot price is equal to the extrapolated futures price.

Consider the behavior of Y at the turn of the month. t is an integer and Y is a function of the futures contracts that mature at time t to $t + N - 1$. An instant later, at time t_+, Y is a function of the contracts that mature at $t + 1$ to $t + N$. In general this means that Y will jump. Since Y is a linear function of traded prices, it can be hedged perfectly intramonth through a static hedging strategy. It cannot be hedged over the month end because that jump reflects the difference between the price at which the long-dated contract trades and its expected price given the contemporaneous prices of the other contracts. This is economically plausible. The agent adjusts Y_t, the expectation at time t of the spot price at maturity, because of the arrival of new information. The information is the surprise in the long-dated futures price.

Lemma 2. *At the turn of the month when t is an integer and a new contract is listed, the extrapolated futures price Y_t jumps and the size of the jump is given by*

$$Y_{t+} - Y_t = \beta_1^{(T-t-1)} \varepsilon_t. \tag{4}$$

The characterization of the optimal hedge follows immediately:

Proposition 1. *The optimal hedge is to set*

$$\Delta_t = \beta^{(n)}, \tag{5}$$

where n is the largest integer $\leq (T - t)$. The agent's terminal wealth is given by

$$W = K - Y_0 - e \text{ where } e \equiv \sum_{n=1}^{T-N} \beta_1^{(T-n-1)} \varepsilon_n. \tag{6}$$

If the innovation in the forecasting equation has constant variance σ^2, then the error term has variance

$$\sigma^2 \sum_{n=1}^{T-N} \left(\beta_1^{(T-n-1)} \right)^2. \tag{7}$$

The optimal hedge can be implemented as follows: First, regress $F(t, t + N)$, the price of the new long-dated futures contract, on $\mathbf{F}_t$, the vector of current prices, to estimate β, then use the recursive relationships set out in Appendix A to compute $\beta^{(n)}$.

The hedge creates a synthetic long-term forward oil contract with an expected delivery cost equal to the expected future spot price. The hedge is not perfect. The error is a weighted sum of the surprises in the prices at which the new long-dated contracts come on stream. Any other hedging strategy has these same hedge errors, plus an additional contribution from imperfectly hedged changes in futures prices. The hedging strategy is optimal for any risk-averse agent. We will refer to this optimal strategy as the *linear hedging strategy*.

Further intuition about the optimal hedge can be obtained by recognizing that it involves hedging the expected future spot price at T. Now by the law of iterated expectations this can be written as

$$\mathrm{E}\left[F(T,T) \mid \mathbf{F}_0\right] = \mathrm{E}\left[\mathrm{E}\left[\ldots \mathrm{E}\left[F(T,T) \mid \mathbf{F}_{\mathrm{T-N}}\right] \ldots \mid \mathbf{F}_1\right] \mid \mathbf{F}_0\right].$$

Note that the expectations are all linear in the relevant futures prices. Going back in time, use the martingale assumption (Assumption 5) to replace the end-month price by the beginning-month price. Use assumption 6 to replace the price of a contract that has just come on stream by a linear combination of contracts that have been trading for months. Since the expectation is linear, hedge ratios are state independent, and the terminal payoff to the strategy is independent of the dynamics of the state variables.

1.3 Simple extensions to the model

No account has been taken of the time value of money. If short-term interest rates are not zero, but are equal to some known function of time $r(t)$, the results go through largely unaltered. The only difference is that it is necessary to “tail” the hedge[4] to account for the time value of money from when cash flows arise on the hedge until time T. Stochastic interest rates pose more of a problem, but provided that rates do not jump and rate changes are substantially uncorrelated with oil price changes, the model remains intact.

To see this let B_t denote the price at time t of a discount bond which matures at time T. Suppose that the agent invests any cash and does any borrowing by holding or shorting the discount bond. Reinterpret X_t as the face value of the agent’s holding of the discount bond at time t. Then the dynamics of X are no longer given by Equation (5) but by:

$$\mathrm{d}X_t = \frac{\Delta_t' \cdot \mathrm{d}\mathbf{F}_{\mathrm{t}}}{B_{t+\mathrm{d}t}} \tag{8}$$

If the bond price is not stochastic, or if it follows a Brownian diffusion with bond price changes which are locally uncorrelated with changes in the vector of futures prices, the results in Proposition 1 go through unaltered

[4] Tailing reduces the hedge ratios by a factor of $-\mathbf{exp} \int_t^T r(u)\, du$.

apart from the need to "tail" the hedge. The optimal hedging strategy is now to set:

$$\Delta_t = B_t \beta^{(n)} \tag{9}$$

where n is the largest integer $\leq (T - t)$.

Terminal wealth under this strategy remains as in Proposition 1. The extension of the model to take account of correlation between interest rate changes and changes in oil prices would be a much more substantial affair, and is not considered in this article.

We have examined the hedging problem when the commitment has the form of a "bullet" — to supply a quantum at some fixed future time. Given a "flow" commitment, where the commodity is to be supplied at a certain rate between two points in time, the optimal hedge is just the sum of the hedges for the individual deliveries. More complex commitments, for example, where the customer can choose when the commodity is supplied, are not linear in prices, and their treatment is beyond the scope of this article.

2. Speculation and Hedging

We have assumed that futures prices follow a martingale. This is restrictive. It means that there is no speculative motive for trading futures, only a hedging motive. How robust are our results to this assumption? We have shown that the optimal hedge can be obtained by regressing changes in the expected cost of delivery on changes in actual futures prices. Does this still hold if futures prices do not follow a martingale?

If futures prices have a predictable drift component, the optimal strategy will depend on the agent's preferences. We set the problem in an expected utility framework. To avoid complex dependence on wealth levels, constant absolute risk aversion is assumed. In place of the martingale assumption (Assumption 5), assume

Assumption 5′. *The agent's objective is to maximize*

$$\mathrm{E}\left[-\frac{1}{\lambda}\mathrm{e}^{-\lambda\tilde{W}}\right], \tag{10}$$

where λ is a risk-aversion coefficient and W is the agent's terminal wealth.

It is worth noting that the linear hedging strategy ($\Delta_t = \beta^{(n)}$) still leads to the agent locking in a price of Y_0 with a hedge error of e, whether or not futures prices follow a martingale process. What changes is that it is no longer necessarily the optimal strategy. So we can still implement the strategy and compute the corresponding hedge error even if futures prices have predictable drift components. Since the strategy is feasible it places a lower bound on the expected utility of the optimal hedge.

If there are predictable drift components, they can be exploited whether the agent has a long-term supply commitment to hedge or not. Call the optimal strategy in the absence of a supply commitment the *speculative strategy*, Δ^{S}. The corresponding expected utility of terminal wealth G^S will be a function of the cash holding X, the current level of prices $\mathbf{F}_t$, and time t. When taking on the supply commitment the optimal strategy changes. In addition to a speculative motive for trading there is also a hedging motive. The optimal strategy is now $\Delta^{\mathrm{S}} + \Delta^{\mathrm{H}}$. The following proposition shows that unless the long-term supply commitment provides a natural hedge against the risks of speculative trading, the incremental trading generated by the supply commitment, Δ^{H}, is independent of the drift.

Proposition 2. *If shocks to G^S, the expected utility of wealth from speculative trading, which occur when a new contract is listed, are independent of the innovation ε_t then*

(i) The optimal trading strategy is the sum of two strategies: a speculative strategy which the agent would follow if there were no forward commitment to hedge, and a hedging strategy which is the same as in the martingale case.

(ii) The break-even price at which the agent could write the long-term contract and maintain the same expected utility is

$$Y_0 + \frac{1}{\lambda} \log \mathrm{E}\left[\mathrm{e}^{\lambda e} \mid \mathbf{F}_0\right]. \tag{11}$$

If the hedge error is normally distributed this can be written as

$$Y_0 + \frac{1}{2}\lambda\, var[e]. \tag{12}$$

(iii) If the change in expected utility and the surprise in prices are generally positively (negatively) correlated the break-even price is lower (higher).

While it is possible to conceive of beliefs which would induce a correlation between ε_{t}, and contemporaneous changes in G^S, these beliefs tend to be rather elaborate. In the absence of correlation, the break-even price is the extrapolated futures price for the appropriate maturity plus a risk premium that depends on the agent's risk aversion and on the variance of the hedge error. The break-even price does not reflect the agent's beliefs about the drift in futures prices. This is because the agent is free to speculate on the drift by trading in short-dated contracts without having to write long-term contracts.

3. Empirical Evidence

We test the model using the crude oil futures contract traded on the New York Mercantile Exchange (NYMEX). Our data cover the period 1 July 1986–

31 December 1994. The contract is the most actively traded crude oil contract in the world with daily volume in 1994 averaging more than 100,000 contracts per day and open interest amounting to more than 400,000 contracts. The contract is for physical delivery of 1,000 barrels of crude and is based on West Texas Intermediate (WTI) grade. Contracts for the next 18 consecutive months are traded, together with four long contracts which start out with maturities of 21, 24, 30, and 36 months. Trading is by open outcry during the day, and electronically when the floor is closed.

Up to 1989, the longest-maturity contracts were 12 months, but there was very little activity in some of the longer-dated contracts. The dataset includes daily settlement prices for contracts maturing in each of the next nine months and for some longer contracts. Ninety percent of trading volume in the dataset is in contracts with four months or less to maturity; less than 3% is in contracts with more than nine months remaining. Although activity in the longer-dated contracts increases over time, volume in the longer-dated contracts remains relatively low (even in the last two years of the dataset only 3% of contracts traded had a maturity beyond 1 year).

The nature of the dataset imposes limitations on the types of tests it is possible to perform. The ideal would be to estimate the model, use it for hedging a series of long-term contracts, and compute statistics showing the performance of the hedge. Yet if "long-term" commitments means commitments extending over several years, it would require several decades of data to obtain a sufficient number of independent observations. With the available dataset we carry out a number of separate studies which together cast light on the validity and utility of the model.

The central assumption of the extrapolation model is the linearity of the forecasting equation. One concern is that nonlinearity, amplified by extrapolation, might make the model unreliable. To explore this, we examine the performance of the extrapolation model in a world where the linearity assumption is known to be false. In particular, we use two of the dynamic models of the term structure of futures prices discussed by Schwartz (1997) and show that the extrapolation model is well behaved.

Next, we test the model's performance in hedging a three-year contract using contracts with up to eight months to maturity, and also how well it hedges an eight-month contract using contracts of up to three months. In both cases the extrapolation model is compared with the Gibson and Schwartz (1990) two-factor model ("the GS model") which is used as a benchmark. The reason for looking at both the three-year and the eight-month horizon is that the former setting is probably more typical of the way in which the model would actually be applied, while tests on the eight-month horizon have more statistical power.

In the final set of tests we examine the validity of the assumptions of the extrapolation model, estimate the model parameters, and assess the quality of the hedging strategy.

3.1 Robustness of the model to nonlinearity

We examine how the extrapolation model is affected by nonlinearity. We assume that the world corresponds exactly to some dynamic term structure model where the forecasting equation is nonlinear, and evaluate how well the extrapolation model performs. The advantage of this approach is that it isolates the issue of nonlinearity from problems such us parameter uncertainty, which are considered separately.

Assume then that the world corresponds exactly to the GS model, where there are two state variables (the spot price and the instantaneous convenience yield) which follow a joint diffusion process. The price of a futures contract can be expressed as a function of the state variables, and any long-dated commitment can be hedged perfectly using just two futures contracts. The forecasting equation is exact: the price at which a new contract opens is a function of the prices of any two other traded contracts. But the key assumption of the extrapolation model is violated because the function is nonlinear.[5] The question we explore is whether, using realistic parameters [those estimated by Schwartz (1997; Table VI, column 2)] for the GS model, it is still possible to make effective use of the extrapolation model. In particular, how well can one hedge a five-year commitment using just two futures contracts with no more than nine months to maturity, rolling the position every two months?

The joint distribution of the five-, seven-, and nine-month futures prices is given by the model, so β_1 and β_2 can be estimated by regressing the nine-month on the five- and seven-month prices. The hedge ratios $\beta_1^{(60)}$ and $\beta_2^{(60)}$ can be computed and the hedged portfolio is

$$F(t, t+60) - \beta_1^{(60)} F(t, t+9) - \beta_2^{(60)} F(t, t+7).$$

The unconditional volatility of the hedged portfolio is then calculated using the GS model.[6]

The results are summarized in Table 1 (scenario I). The annualized volatility of the five-year price is 25.9%. The volatility of the hedged portfolio is 0.82%. The linear approximation does not prevent the hedge from removing almost all the risk.

The extrapolation model gives the optimal hedge as short 2.89 of the seven-month, and long 3.93 of the nine-month contract, so the hedge is net long approximately one barrel of oil. But the quality of the hedge is not driven just by the net exposure to the oil market. A simple hedge con-

---

[5] From Equation (18) in Schwartz (1997), the forecasting equation takes the form

$$F(t, t+N) = kF(t, t+m)^{-\theta} F(t, t+n)^{1+\theta},$$

where k and θ depend only on N, m, and n and the model parameters.

[6] The volatility of the hedged portfolio in general varies with the level of the convenience yield. The figures in Table 1 are root mean variances, where the variances are averaged across realizations of the convenience yield.

Table 1
Volatility of hedged portfolios in a Gibson–Schwartz world

	Two-factor model				Three-factor model	
Scenario	I	II	III	IV	V	VI
Volatility of 5-year contract:						
Unhedged	25.88%	18.27%	14.95%	11.67%	14.31%	11.66%
1-1 hedge	9.14%	11.32%	11.35%	6.91%	11.24%	7.58%
2 contract hedge	0.82%	1.02%	1.58%	0.64%	2.12%	1.51%
3 contract hedge	0.08%	0.13%	0.24%	0.07%	1.88%	0.89%
Volatility of 10-year contract:						
Unhedged	25.88%	18.27%	14.94%	11.66%	14.44%	11.80%
1-1 hedge	8.85%	11.74%	9.95%	6.17%	9.08%	6.73%
2 contract hedge	1.24%	1.30%	1.84%	0.71%	3.08%	2.26%
3 contract hedge	0.08%	0.13%	0.24%	0.07%	2.62%	1.71%

Hedged portfolios are constructed using the extrapolation method. Their volatility is then calculated in a Gibson–Schwartz world. The six scenarios use the parameter estimates made by Schwartz (1997; Tables VI and IX), and are derived from oil market data. The first four scenarios are based on Schwartz's two-factor model; the last two on his three-factor model. The first four scenarios differ in the period over which the estimates are made (1985–1995, 1990–1995, 1990–1995, and 1993–1996, respectively) and the data used (I and II use futures up to 9 months, III also uses a 17-month contract, while IV uses swap data). Scenarios V and VI use the same data sets as scenarios III and IV, respectively. In the 1-1 hedge the long-dated contract is hedged by one nine-month contract. The two-contract hedge uses seven- and nine-month contracts, while the three-contract hedge uses five-, seven-, and nine-month contracts.

sisting of one 9-month contract would have a much higher volatility of 9.1%.

The quality of the hedge is improved if more contract maturities are used. Using three contracts, of five, seven, and nine months, the volatility of the hedged portfolio falls to 0.08%.

Schwartz (1997) estimates the GS model using four different datasets and gets substantially different parameter values. As can be seen from Table 1, the volatility of the hedged and unhedged portfolios do vary, but for each set of estimates the same conclusions are apparent: the extrapolation method does succeed in removing the great bulk of the risk of a long-term commitment, the model's success does not derive mainly from hedging changes in the level of oil prices but comes also from hedging changes in the convenience yield, and a three-contract hedge comfortably improves on a two-contract hedge. The lower panel shows very similar results for the hedging of a 10-year commitment.

Schwartz (1997) also considers a three-factor term structure model, which incorporates stochastic interest rates as well as a stochastic spot price and convenience yield. Using his parameter estimates [Schwartz (1997); Table IX] we follow exactly the same procedure as before. The results are set out in the last two columns of Table 1. Again, the extrapolation model performs rather well, though somewhat less well than in the two-factor model case.

In Appendix B, we provide some theoretical insight into the divergence between the extrapolation model and the GS model. We show that the extrapolation model is likely to work best when the convenience yield is not too volatile and shocks to the convenience yield die away quickly.

We have shown that in a world where the key assumption of the extrapolation model is violated, the model gives an effective hedging strategy. This is encouraging evidence for the robustness of the extrapolation model, but says nothing about its merits relative to alternative models.

3.2 Replicating a three-year futures contract

We use the extrapolation model and the GS model to hedge a three-year commitment using futures contracts with no more than eight months to maturity. At month t the parameters of each model are estimated using prices up till that month. Using the model and the current price of the contracts to be used for hedging, the three-year futures price is estimated. We postulate an agent who sells one barrel of oil forward, for delivery at time $t+36$, at the model price, who then hedges it using short-dated futures contracts (up to eight months) as the model dictates. Eight months from the date the commitment is due, at month $t+28$, the final delivery price is locked in using an eight-month contract, and the profit or loss determined. If the model works perfectly, the profit will be zero.

The exercise is repeated each month. With the passage of time the volume of past data increases and the model can be estimated with increasing precision. With the dataset beginning in July 1986, the exercise starts in December 1988, using 18 months of data for model estimation. First delivery follows three years later, in December 1991. The exercise is repeated each month, with the last commitment made in August 1992, when there are six years of data to estimate the model. Thus the hedging exercise is repeated a total of 45 times. The results are set out in Table 2.

Three variants of the extrapolation method are considered. The first uses the three longest available contracts, with the position being rolled every month, so at the beginning of each month the hedge consists of contracts with six, seven, and eight months to maturity. The root mean square hedge error averages \$1.44/barrel over the period. The strategy requires a lot of trading — on average 488 contracts are traded per contract hedged.

The second strategy uses alternate months' futures contracts and is rebalanced every two months. It performs better. The error is \$1.07/barrel, and it only requires an average of 60 futures contracts to be traded per contract hedged. The reduction in trading volume is partly due to the lower frequency with which the hedge is rebalanced, and partly due to the smaller size of positions held at any time. The third strategy uses just two contracts and is rebalanced bimonthly. It performs only slightly worse than a three-contract hedge, and turnover is slightly reduced.

The GS hedge is implemented using contracts which initially have six and eight months to maturity and are rolled every two months. The parameters are estimated using a Kalman filter, as described in Schwartz (1997). On the

Table 2
Hedging a three-year commitment

	Hedge error		Turnover	Extreme cash position
	rms ($/bbl)	SD ($/bbl)	(# contracts)	($/bbl)
Extrapolation methods using:				
6, 7, 8 months	1.44	1.43	488	3.68 to −2.72
4, 6, 8 months	1.07	1.06	60	4.20 to −2.26
6, 8 months	1.16	1.16	48	4.12 to −1.50
GS model:				
6, 8 months	2.53	1.46	75	5.28 to −1.53
1-1 hedge				
8 months	3.75	2.54	15	6.90 to −2.26
No hedge	3.35	3.32	0	0.00

Hedge contract maturities are immediately after they have been rolled forward.

The table shows the performance of different strategies for hedging a three-year commitment using futures contracts of up to eight months. The hedge is run monthly from December 1988 for 45 months. For each strategy both the root mean square hedge error and the standard deviation of the hedge error are reported. Turnover is total number of futures contracts taken out (long or short) per contract hedged, averaged over the 45 months. The cash position is defined as the net cumulative margin paid on the futures position; for each hedge the maximum and minimum cash position is computed and the numbers are averaged over the 45 months.

With both the 1-1 hedge and the no hedge the long-term commitment is concluded at the price of the eight-month future; for the other models the contract is written at model price. The 1-1 hedge consists of a single eight-month contract that is rolled every two months.

basis of the estimated model parameters and the prices of the two contracts used for hedging at the beginning of each two-month period, the theoretical hedge ratios are computed. The ratios vary slightly over the two-month period as the slope of the term structure changes. In practice the changes are small and the effect on the hedging strategy is negligible, so the hedge is implemented as a static hedge within each period. The root mean square error of the GS hedge is $2.53/barrel. This is much higher than for the extrapolation hedges. The turnover is marginally higher than the bimonthly extrapolation hedge.

To provide some basis for comparison we also look at the risk from following a simple 1-1 hedging strategy, where a single eight-month contract is held for two months and then rolled forward; we also show what happens if the commitment is not hedged at all. For the 1-1 hedge and the no-hedge strategies, there is no model price for the three-year commitment, so the hedge profit is computed on the assumption that the contract is written at the eight-month futures price. This may overstate the hedge error since the eight-month price may not be an unbiased estimator of the three-year price. For these two strategies the standard deviation may be a better measure of hedging error than the root mean square.

The hedge errors are plotted in Figure 1. The top chart shows the results for the extrapolation hedges, while the lower chart shows the other strategies. The two charts are drawn to the same scale and show the hedging error

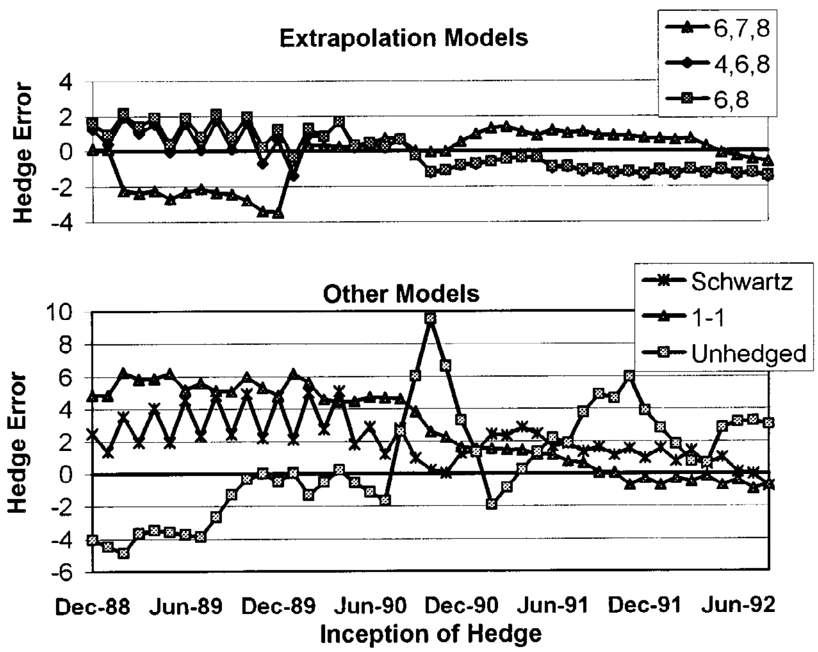

Figure 1
Hedging error from hedging a three-year commitment with futures contracts of eight months or less
A hedge is set up each month using the relevant model estimated on past data. The top panel uses the extrapolation method. "6,7,8" uses contracts which, just after the position is rolled each month, have six, seven, or eight months to maturity. Similarly for "6, 8" and "4,6,8" except that they are rolled bimonthly. The lower panel shows the effect of using the Gibson–Schwartz model, rolling bimonthly, with six- and eight-month contracts; also the effect of a single eight-month contract rolled bimonthly ("1-1") and no hedge.

(measured in $/barrel) for a particular hedging strategy plotted against the month in which the hedge was started. The magnitude of the errors tends to decline over the period. In the case of the GS and extrapolation models, this could be explained by the improvement in parameter estimation which comes as more data are used to estimate the model, but it does not account for the improvement in the performance of the 1-1 hedge where no parameters are estimated. Part of the improvement is likely due to the easier hedging environment in the second half of the period when the shape of the term structure was less volatile.

Some of the plots show a sawtooth pattern in the early part of the period. This is due to bimonthly rebalancing. In any one month, a hedge which originated an even number of months ago might be long the eight-month contract and short the six-month. A similar hedge which originated a month earlier or later would be long the 7-month contract and short the five-month. Normally the hedge gain or loss would be very similar for these two positions. However, with the turmoil in the oil market in 1990–1991, there were large swings in prices which led to differences of up to $1/barrel between similar hedge portfolios in some months.

The cash flows that arise during the life of a hedge are of some interest. Table 2 shows the average across all the 45 months of the largest positive and negative cash balances on each strategy. Apart from the no-hedging strategy where there are no interim cash flows, it is movement in the level of oil prices rather than choice of hedging strategy that is the most important determinant of the size of interim cash flows.

The conclusions we can draw from this exercise are suggestive, but not conclusive. The 45 different profit figures for each hedge are not independent estimates of the error since the hedging periods overlap. The agent is actively managing each hedge over a period of 28 months, so the 45 data points do not even constitute two nonoverlapping periods. Furthermore, the Gulf War comes in the middle of the period and has considerable impact on the oil market. However, the results do provide evidence, which we confirm below, that the extrapolation model does reduce risk substantially, and appears to do better in this regard than the GS model.

3.3 Hedging an eight-month futures contract

To enhance the power of the tests we look at hedging an eight-month rather than a three-year commitment. Not only are there more nonoverlapping periods, but the market at eight months is sufficiently liquid that changes in the value of the hedge and of the underlying commitment can be compared on a monthly basis.

At the start of each month, the parameters of each model are estimated using all prior data. The hedge for an eight-month contract is constructed using futures with three months or less remaining. The position is held for a month and the hedge error measured. The procedure is repeated in subsequent months and the root mean square hedge error computed. The results are shown in Table 3.

For all methods the hedging error is very much larger in the early part of the period. In part this results from having less data earlier from which to estimate the parameter values, but in the main it is due to the great turbulence around the period of the Gulf War. So we report results both for the period as a whole, and also for the period from March 1991. The ranking of the methods is similar whichever period we look at. Both Gibson–Schwartz and the extrapolation method reduce risk compared with not hedging, the extrapolation method being more effective than Gibson–Schwartz. The difference between the two is statistically highly significant in the period excluding the Gulf War.

There is no great benefit in using a third contract in the extrapolation method. The ranking of the two and three contract hedges changes depending on the period chosen, and the differences are not statistically significant

Table 3
Hedging performance of Gibson–Schwartz and extrapolation models

	Extrapolation		Gibson–Schwartz		
	1,2,3[a]	2,3[a]	1,3[a]	2,3[a]	Unhedged
Root mean square error in hedging an eight-month future (cents/barrel/month):					
12/88–12/94	28.14	27.34	53.15	31.91	117.97
		(0.6403)	(0.0000)	(0.1605)	(0.0000)
3/91–12/94	9.35	10.13	23.48	16.20	69.74
		(0.3397)	(0.0000)	(0.0000)	(0.0000)

[a] Maturity of contracts used for hedging.

The root mean square error is computed by measuring the change in the price of an eight-month contract, hedged by contracts which initially have one to three months until maturity, over the following month.

Numbers in parentheses are *p*-values for the hypothesis that the hedging strategy is no worse than the extrapolation hedge using one-, two- and three-month contracts; they are computed following the procedure in Granger and Newbold (1986, p. 279).

in either period. Perhaps this is not surprising since the position taken in the one-month contract when it is used is very small.[7] In using Gibson–Schwartz, it is much better to use two- and three-month contracts for hedging than one- and three-month contracts; using the one month contract picks up short-term volatility that is not reflected in the eight-month price.

It might be argued that the relatively poor performance of the GS model is due to problems in estimating the parameters or in implementing the hedge. The hedge ratios in the GS model depend only on the futures prices, which are in our dataset, and the mean reversion rate of the convenience yield, κ. Even if the value of κ is chosen to minimize the hedging error, the hypothesis that the (2,3) GS model provides a better hedge than the extrapolation model is rejected at the 1% level in the post–Gulf War period.

To test whether the performance of the GS hedge could be attributed to the use of a static rather than a dynamic hedge as the model strictly implies, the hedge returns were reestimated as follows. At the beginning and end of each month the theoretical price of the eight-month future is estimated from the prices of the contracts used for hedging it. If the model is correct, and the hedge balanced dynamically, the change in the value of the hedge equals the change in the theoretical price of the eight-month contract over the month. It turns out that the hedge returns as measured in this way are virtually identical to the static hedge returns (there is a correlation of 0.997 between the two measures), and there is virtually no impact on the size of hedge errors.

The three-factor model described in Schwartz (1997) can also be used to hedge an eight-month contract with one-, two- and 3-month contracts, but the results are very poor. The absolute value of the hedge ratios required

[7] The holding of the one-month contract typically accounts for less than 10% of the gross number of contracts held in the hedge portfolio.

to hedge all three factors are so large (typical values are +11, −27, and +18 contracts, respectively, for every eight-month contract hedged) that the performance is no better than not hedging at all.

3.4 Implementing the model for long-term commitments

We now test the assumptions of the extrapolation model (notably Assumption 6), estimate the forecasting equations for different hedging instruments, and compute hedge ratios for different maturities. We also estimate the size of hedging error, taking into account the errors in estimating the model parameters.

Assuming that the longest-dated liquid contract suitable for hedging is the eight-month contract, the first step is to estimate the forecasting equation:

$$F(t, t+8) = \alpha + \sum_{i=l}^{8} \beta_i F(t, t+8-i) + \varepsilon_t. \tag{13}$$

Given the very weak evidence for stationarity in oil prices, it might seem problematic using ordinary least squares (OLS) for estimation. Augmented Dickey–Fuller tests for a unit root were generally inconclusive. For the longer-dated futures contracts we were unable to reject the existence of a unit root even at the 10% level; for the shorter contracts we could reject unit roots. But OLS is still efficient and consistent if the series form a cointegrating system [see Engle and Granger (1987)]. Using Johansen's (1991) method there is strong evidence of cointegration; the null that the series were not cointegrated was rejected at the 1% level.

The forecasting equation is estimated using OLS. The results are set out in Table 4. The very high correlation coefficients in the regressions are unsurprising. While oil prices are volatile, the shape of the term structure at any point in time is quite smooth. Virtually all the model's power in predicting the price of the new long-dated contract comes from the prices of the two or three longest-dated contracts trading. We therefore set β_4 to β_8 equal to zero, and take the forecasting equation in our base case to be as in the third column of Table 4. The error in forecasting the price of the new contract is less than 1.60¢/barrel.

We found no evidence of parameter instability, seasonality, or nonlinearity. The fifth and sixth columns of Table 4 estimate the regression over the two halves of the period separately. A Chow test was carried out to test the hypothesis that the true parameters are identical in the two halves. With an F-statistic of 1.84, the null cannot be rejected at conventional significance levels. We tested for seasonality by including monthly dummies in the regressions. Using a Wald test, the hypothesis that all the coefficients were zero could not be rejected even at the 20% significance level. We tested for nonlinear terms by regressing the eight-month price on both the five- to

Table 4
Estimation of the forecasting equation

Period	1986–1994	1986–1994	1986–1994	1986–1990	1990–1994	1986–1994
NObs	102	102	102	51	51	51
R^2	99.997%	99.996%	99.996%	99.997%	99.995%	99.968%
α	0.42	1.14	0.02	−0.47	12.46	0.10
β_1	2.078	2.116	1.839	2.054	2.015	
	(0.103)	(0.082)	(0.016)	(0.153)	(0.099)	
β_2	−1.232	−1.353	−0.840	−1.222	−1.236	1.979
	(0.215)	(0.150)	(0.016)	(0.280)	(0.172)	(0.089)
β_3	0.035	0.237		0.169	0.215	
	(0.196)	(0.069)		(0.129)	(0.076)	
β_4	0.331					−1.216
	(0.204)					(0.143)
β_5	−0.395					
	(0.179)					
β_6	0.229					0.232
	(0.105)					(0.057)
β_7	-0.047					
	(0.036)					
β_8	0.001					
	(0.009)					
Standard error (cents/barrel)	1.57	1.58	1.67	1.72	1.37	4.75

The forecasting equation is

$$F(n, n+8) = \alpha + \sum_{i=1}^{8} \beta_i F(n, n+8-i) + \tilde{\varepsilon}_n.$$

Standard errors in parentheses.

seven-month prices and on their squared values; we were again unable to reject the null at the 20% level.

The hedge error in the model was shown in Proposition 1 to be the weighted sum of the prediction errors. The time-series behavior of the error is therefore important. Returning to the base case (as in the third column of Table 4) the Durbin–Watson statistic is 2.22. Using a Lagrange multiplier test, there is no evidence of serial correlation in the residuals ($\chi^2 = 1.55$, p-statistic $= 21\%$).

We have already seen (in Section 3.2) that the extrapolation model seems to work well when rebalanced every two months, so the last column of Table 4 shows the regression of the eight-month price on the sixth, fourth, and second month. The equation was estimated using data from alternate months. The fit, unsurprisingly, is somewhat less good than the earlier regressions, as demonstrated by the slightly lower R^2 and the higher standard error. But this is offset by the fact that the hedge does not have to be rolled so frequently. There was no evidence of parameter instability, seasonality, nonlinearity, or serial correlation in the residuals.

Using the extrapolation methodology we compute the optimal hedge for different horizons in the base case. The results are set out in the top panel of Table 5. To interpret the table consider the last column in the top panel. If the agent has contracted to supply one barrel of oil in six years time and seeks to hedge using futures contracts with up to eight months to maturity, the optimal strategy is to go long 6.52 barrels of the eight-month, short 7.30 barrels of the seven-month, and long 1.55 barrels of the five-month. This implies a gross position of 15.38 barrels and a net long position of 0.77 barrels. The agent would need to revise the position monthly as the maturity of each futures contract shortens and as the maturity of the supply commitment reduces.

The optimal hedge involves very large long and short positions — to hedge an exposure of 1 barrel more than a year away requires a gross position in excess of 10 barrels in the futures market. Although the hedge looks stable — the optimal hedge is very similar for any T in excess of two years — maintaining it requires considerable trading. As the eight-month contract becomes a seven-month contract the agent must turn a long position into a short, and similarly for the other contracts. The volume of trading each month equals the size of the gross position.

Although the optimal hedge looks highly leveraged, the principal determinant of hedge returns is movements in the level of the oil price; at 72 months, for example the correlation between the return on the optimal hedge and on a single 8-month contract is 0.939. So the net exposure is a useful statistic to look at. As can be seen from Table 5, the net position, which starts at one barrel, remains close to unity for many months. This results from the fact that the estimated coefficients of the forecasting equation sum almost exactly to unity. It is consistent with models in which the price of oil is nonstationary.

Standard errors of the hedge ratios are shown in the Table 5.[8] The hedge error, as recorded in the bottom row, reflects both the rollover errors at the turn of the month as set out in Proposition 1 and the error in estimating the parameters of the model.

The collinearity of futures prices manifests itself in large standard errors in the estimated hedge ratios. The standard errors on the net hedge position are much smaller than on the gross. The standard errors increase steeply with the time horizon, though the net hedge position is still different from zero at conventional significance levels at 72 months. The hedge error increases with the time horizon, going from \$0.40/barrel at 24 months to \$1.18/barrel at 72 months.

The strategy of rebalancing the portfolio alternate months obviously has attractions in terms of containing transaction costs. So panel B of Table 5 shows the implications of restricting both the estimation of the forecasting

[8] The hedge ratios are functions of the coefficients of the forecasting equation. The standard errors are computed by approximating these functions by their first-order Taylor series expansions around their estimated value.

Table 5
Optimal hedging strategies

A. The optimal hedging strategy with monthly rebalancing						
Maturity (in months)	**8**	**9**	**12**	**24**	**48**	**72**
Hedge composition:						
Eighth month	1	2.12	4.70	7.65	7.26	6.52
		(0.08)	(0.43)	(1.29)	(2.23)	(2.97)
Seventh month	0	−1.35	−4.65	−8.50	−8.13	−7.30
		(0.15)	(0.76)	(1.90)	(2.78)	(3.52)
Sixth month	0	0.24	0.94	1.80	1.73	1.55
		(0.07)	(0.35)	(0.74)	(0.83)	(0.91)
Total (gross)	1	3.71	10.29	17.95	17.12	15.38
		(0.30)	(1.53)	(3.83)	(5.73)	(7.31)
Total (net)	1	1.00	0.99	0.95	0.86	0.77
		(0.00)	(0.01)	(0.07)	(0.20)	(0.30)
Hedge error ($/bbl)	0	0.02	0.09	0.40	0.84	1.18
B. The optimal hedging strategy with bimonthly rebalancing						
Maturity (in months)	**8**	**9**	**12**	**24**	**48**	**72**
Hedge composition:						
Eighth month	1	1.98	2.70	3.66	2.92	2.27
		(0.09)	(0.21)	(0.66)	(0.94)	(1.08)
Sixth month	0	−1.22	−2.18	−3.62	−2.92	−2.27
		(0.14)	(0.33)	(0.91)	(1.09)	(1.18)
Fourth month	0	0.23	0.46	0.85	0.69	0.54
		(0.06)	(0.13)	(0.33)	(0.33)	(0.32)
Total (gross)	1	3.43	5.34	8.13	6.53	5.08
		(0.29)	(0.67)	(1.86)	(2.33)	(2.56)
Total (net)	1	1.00	0.99	0.89	0.69	0.54
		(0.00)	(0.01)	(0.07)	(0.17)	(0.22)
Hedge error ($/bbl)	0	0.05	0.11	0.44	0.83	1.08
C. Risk of unhedged portfolio						
($/bbl)	0	1.16	2.20	4.02	6.35	8.29

The table shows the composition of the optimal hedge according to the maturity of the supply commitment. In panel A, the hedge is rolled over each month into futures contracts with six, seven, and eight months to expiration. The forecasting equation used is that estimated in column 3 of Table 4. In panel B, the hedge is rolled over every two months into futures contracts with four, six, and eight months to expiration. The forecasting equation used is that estimated in column 7 of Table 4. Figures in parentheses are standard errors; the standard errors are computed by approximating the quantities by linear functions of the regression parameters and using the covariance matrix from the regression equation. The hedge error in the bottom line of panels A and B includes the error in estimating the coefficients of the forecasting equation.

equation and the hedging strategy to contracts which mature on alternate months. The bimonthly strategy is highly correlated with the monthly strategy of panel A; in hedging a 72-month commitment, for example, the correlation between the two is 0.993 (this compares with a correlation of 0.939 with a strategy of holding a single contract). The strategies differ in the size of the hedge. The bimonthly hedge is net long 0.54 contracts at 72 months as against 0.77 contracts for the monthly hedge. But this appears to be the result of estimation differences rather than substance. The differences are well within the standard errors. Further, the bimonthly hedge was estimated using half the months in the sample; a bimonthly hedge estimated on the other months yields an estimate of the net position which is slightly higher than the monthly hedge.

The amount of trading required is much reduced by hedging bimonthly; at 72 months the bimonthly hedge requires trading only 5 contracts every two months against 15 contracts per month for the monthly hedge. A small part of this sixfold reduction in turnover comes from the lower net position held. The major gain comes from the smaller gross position required to hedge changes in the shape of the term structure, and the fact that positions are held for two months rather than one; together these factors reduce turnover roughly fourfold.

It is interesting to compare the standard errors of the two strategies. Some caution is needed inasmuch as each set of standard errors is computed on the basis that the model used is correctly specified. But taking the numbers at face value, it appears that rebalancing every two months does not reduce hedging efficiency to any appreciable extent, except possibly for very short horizons. This conclusion is supported by the high correlation between the two sets of hedging returns.

To put the hedging errors into context, panel C of Table 5 shows the error if the agent does not hedge at all[9] (until eight months before maturity when the position is hedged perfectly with a futures contract). The extrapolation method reduces the risk of an unhedged 72-month contract by 85% (with the hedge error dropping from \$8.29 to \$1.18/barrel); the gains are proportionately still larger at shorter maturities.

We now investigate whether the model can be applied using just the most liquid contracts, those with up to three months to maturity. The forecasting equation, where the three-month contract is expressed as a linear function of the three shorter contracts, is misspecified. The Durbin-Watson statistic is 1.53 and there is strong evidence of serial correlation in the residuals (using a Lagrange multiplier test, the chi-square is 6.06, with a p-statistic of 1.4%). There is also evidence of a time trend. But most significantly from the viewpoint of the model, there is evidence of nonlinearity. The price of the three-month contract depends on the squared difference between the first- and second-month prices. Table 6 shows the forecasting equation with these additional terms. It also shows that the violations of our assumptions,

[9] The error is computed as

$$\sqrt{\sigma^2 \sum_{n=1}^{T-N} \left(\beta_1^{(T-n-1)}\right)^2 + \sum_{n=1}^{T-N} \left(\beta^{(T-n-1)} \cdot \mathrm{cov}(d\mathbf{F} \cdot d\mathbf{F}') \cdot \beta^{(T-n-1)}\right)}.$$

The first part is the unhedgable error in the extrapolated price which occurs when the hedge is rolled over, while the second is the variance of the change in the extrapolated futures price over the course of each month. The actual covariance matrix is used as an estimate of the true one. Implicitly the methodology assumes that changes in the futures price vector are serially independent, and their covariance is constant. Estimates made using the bimonthly model coincide with the estimates in Table 5, which use the monthly model, to within 15%.

Table 6
Forecasting equation using near-month contracts

	Full sample	**Excluding 9/90 to 1/91**
NObs	102	97
α	0.379**	0.189**
	(0.071)	(0.040)
β_1	1.736**	1.854**
	(0.068)	(0.069)
β_2	−0.802**	−0.975**
	(0.089)	(0.108)
β_3	0.043	0.109*
	(0.029)	(0.042)
γ	0.111**	0.150*
	(0.032)	(0.058)
δ	0.00070	0.00046**
	(0.00037)	(0.00017)
ρ	0.528**	0.194*
	(0.101)	(0.095)
Standard error	0.0486	0.0362
Adjusted R^2	0.99979	0.99974
Durbin–Watson	2.09	1.54

* significant at 5% level; ** significant at 1% level.
The equation estimated is

$$F(t,t+3) = \alpha + \beta_1 F(t,t+2) + \beta_2 F(t,t+1) + \beta_3 F(t,t) + \gamma\,\{F(t,t+2) - F(t,t+1)\}^2 + \delta t + \varepsilon_t,$$

where the error term itself follows an AR(1) process:

$$\varepsilon_t = \rho\varepsilon_{t-1} + \eta_t$$

Standard errors in parentheses.

which were not observed in the previous regression using the eight month contract, do not result from the unusual conditions in the oil market around the time of the Gulf War in 1990–1991, but are clearly significant in the rest of the data.

The existence of a time trend, and in particular any uncertainty in measuring the magnitude of the trend, will create additional uncertainty in the expected price at which long-term commitments can be locked in, but the extrapolation method remains optimal. The existence of positive autocorrelation in the error specification will make the aggregate hedging error under the extrapolation method larger than it would be otherwise. If the nature of the serial correlation can be correctly identified, it should be possible to correct the error estimate to take account of it. Nonlinearity in the forecasting equation is more problematic since it suggests that the optimal hedge should vary according to the level of prices. That would require a major extension to the model, which is beyond the scope of this article. Of course the extrapolation method can still be used with very short-dated contracts; and despite its theoretical problems, it may give, as we have seen in Section 3.3, quite good results in practice. But we cannot claim that it is optimal, and it would be unwise to put any reliance on estimates of hedging errors, analogous to those in Table 5, which are made from within the model.

The extrapolation model appears to be well specified for oil futures contracts at around eight-months maturity. We have looked at two rival specifications, both involving three hedging contracts, one of which requires rebalancing alternate months. Both strategies remove the bulk of the price risk in a long-term commitment, with the bimonthly strategy having an advantage in terms of lower transaction volume. The model fits less well at the very short end of the term structure.

4. Conclusions

Futures exchanges offer a range of contracts of different maturities which are designed to manage risk. An agent wishing to use futures contracts to hedge a long-maturity exposure has a difficult problem in deciding how best to do it. This article has set out a robust and simple way of addressing the problem.

The empirical results suggest that one could expect to remove about 85% of the risk (as measured by the standard deviation of the hedge error) of a six-year oil supply commitment by hedging with medium-maturity futures contracts (up to eight-months). This is after allowing for errors in the estimation of the parameters of the model. We also used the extrapolation model to hedge an eight-month contract using contracts with three months or less to maturity. Despite evidence that the assumptions of the model are violated for these maturities, it substantially outperforms the Gibson–Schwartz model.

The robustness of the model derives from the paucity of assumptions made about the processes driving the term structure of commodity prices. No assumptions are made about the number of state variables, about the processes they follow, or about the way risk is priced. The key assumption is that the expected price at which the new long-dated contract starts trading is a linear function of the prices of existing contracts. While strictly inconsistent with some models of the term structure of commodity prices, evidence was presented suggesting that it can still give good results in practice.

An attractive feature of the model is the close relation between parameter estimation and the computation of hedge ratios. The model parameters are estimated by cross-sectional regression of prices of contracts of different maturities. The regression coefficients are chosen to minimize the difference between the actual and expected opening price of the new long-dated contract. Since it is these differences which, aggregated every time the hedge is rolled, constitute the error in the hedge, parameter estimation and hedging strategy are closely linked. This contrasts with the more traditional approach where the parameters are estimated by fitting the entire process. The model is then used to compute the optimal hedge ratio. The more direct approach is less vulnerable to problems arising from model misspecification.

In applying the model one would wish to choose suitable hedging instruments, taking into account the depth and liquidity of the market and the size of transaction costs. In this article we have started to explore alternative formulations, using different sets of contracts for hedging and different rebalancing frequency. The model assumes that the forecasting equation is correctly specified, so our results have been suggestive rather than definitive. It would be useful to set the model in a framework where different formulations could be compared. It would also be interesting to extend the empirical work in this article to look at markets on other assets.

Appendix A

Proof of Lemma 1. The proof of the Lemma is by induction:

(i) *Lemma 1 holds for $t = T$:*

$$\mathrm{E}\left[F(T,T) \mid \mathbf{F}_{\mathrm{T}}\right] = F(T,T),$$

so set

$$\alpha^{(0)} = 0;\ \beta_k^{(0)} = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{otherwise.} \end{cases} \tag{A1}$$

(ii) *If Lemma 1 holds for $t = T - n$ for some integer n, it holds for $t = T - n - \eta$ for any $0 < \eta < 1$:*

$$\begin{aligned} \mathrm{E}\left[F(T,T) \mid \Omega_{T-n-\eta}\right] &= \mathrm{E}\left[\mathrm{E}\left[F(T,T) \mid \mathbf{F}_{T-n}\right] \mid \mathbf{F}_{T-n-\eta}\right] \\ &= \mathrm{E}\left[\alpha^{(n)} + \beta^{(n)'} \cdot \mathbf{F}_{T-n} \mid \mathbf{F}_{T-n-\eta}\right] \\ &= \alpha^{(n)} + \beta^{(n)'} \cdot \mathbf{F}_{T-n-\eta}. \end{aligned} \tag{A2}$$

The first line is the law of iterated expectations; the second is true by hypothesis; the third follows because of the martingale assumption.

(iii) *If Lemma 1 holds for $t = T - n_+$ (that is just after the shortest future matures and a new contract is listed) it holds for $t = T - n$:*

$$\begin{aligned} \mathrm{E}\left[F(T,T) \mid \Omega_{\mathrm{T-n}}\right] &= \mathrm{E}\left[\mathrm{E}\left[F(T,T) \mid \mathbf{F}_{\mathrm{T-n+}}\right] \mid \mathbf{F}_{\mathrm{T-n}}\right] \\ &= \mathrm{E}\left[\alpha^{(n-1)} + \beta^{(n-1)'} \cdot \mathbf{F}_{\mathrm{T-n_+}} \mid \mathbf{F}_{\mathrm{T-n}}\right] \\ &= \alpha^{(n-1)} + \sum_{k=2}^{N} \beta_k^{(n-1)} F(T-n, T-n+N+1-k) \\ &\quad + \beta_1^{(n-1)} \left(\alpha + \sum_{k=1}^{N} \beta_k F(T-n, T-n+N-k)\right). \end{aligned} \tag{A3}$$

The first line is the law of iterated expectations, and the second is true by hypothesis. The third uses Assumption 6 to replace $F(T-n^+, T-n+N)$ by its expectation at time $T-n$. Assembling terms the proposition holds with

$$\begin{aligned}
\alpha^{(n)} &= \alpha^{(n-1)} + \beta_1^{(n-1)}\alpha \\
\beta_k^{(n)} &= \begin{cases} \beta_{k+1}^{(n-1)} + \beta_1^{(n-1)}\beta_k & \text{if } k < N \\ \beta_1^{(n-1)}\beta_k & \text{if } k = N. \end{cases}
\end{aligned} \tag{A4}$$

Proof of Lemma 2. The change in Y at the turn of the month is caused by the difference between the expected opening price of the new N-month contract and its actual price. Since the weight of the contract in S_n is just $\beta_1^{(T-n-1)}$, Lemma 2 follows immediately.

Proof of Proposition 1. If the agent follows some strategy Δ_t, her terminal wealth will be

$$\begin{aligned}
W &= K - Y_T + \int_{t=0}^{T} \Delta_t' \cdot \mathrm{d}\mathbf{F}_t \\
&= K - Y_0 + \int_{t=0}^{T} \left\{ \Delta_t' \cdot \mathrm{d}\mathbf{F}_t - \mathrm{d}Y_t \right\}.
\end{aligned} \tag{A5}$$

But from Lemmas 1 and 2:

$$\int_{t=0}^{T} \mathrm{d}Y_t = \int_{t=0}^{T} \beta^{(n)'} \cdot \mathrm{d}\mathbf{F}_t + \sum_{n=1}^{T-N} \beta_1^{(t-n-1)}\tilde{\varepsilon}_n, \tag{A6}$$

so the agent's terminal wealth can be written as

$$W = K - Y_0 + \int_{t=0}^{T} \left\{ \Delta_t - \beta^{(n)} \right\}' \cdot \mathrm{d}\mathbf{F}_t - \sum_{n=1}^{T-N} \beta_1^{(t-n-1)}\tilde{\varepsilon}_n. \tag{A7}$$

The integral and summation are both sums of independent mean zero random variables. The best the hedge can do is to set the integral equal to zero, hence the proposition. Any alternative strategy adds noise, and hence is dominated in the sense of second-degree stochastic dominance.

Proof of Proposition 2. Define $G^{\mathrm{S}}(\mathbf{F}, X, t)$ as the expected utility of terminal wealth of an agent with no commitments who has cash X at time t when the term structure of futures prices is $\mathbf{F}$, and who follows an optimal hedging strategy. Let that strategy be Δ^{S}.

It follows immediately that

$$G^S(\mathbf{F}, X, t) = -\frac{1}{\lambda}\mathrm{e}^{-\lambda X} f(\mathbf{F}, t),$$

where $f(\mathbf{F}, T) = 1$.

$$G^S(F, X, t) = \mathrm{E}\left[G^S\left(\mathbf{F} + \mathrm{d}\mathbf{F}, X + \Delta^{S'} \cdot \mathrm{d}\mathbf{F}, t + \mathrm{d}t\right) \mid \mathbf{F}_{\mathrm{t}} = \mathbf{F}, X_t = X\right] \qquad \text{(A8)}$$

$$\Delta^S = \underset{\Delta}{\arg\max}\ \mathrm{E}\left[G^S(\mathbf{F} + \mathrm{d}\mathbf{F}, X + \Delta' \cdot \mathrm{d}\mathbf{F}, t + \mathrm{d}t) \mid \mathbf{F}_{\mathrm{t}} = \mathbf{F}, X_t = X\right].$$

The first equation reflects the agent's exponential utility function; the second says that if the agent follows her chosen strategy, expected utility is a martingale, and the third equation says that the chosen strategy is optimal.

Now define the function

$$G^H(\mathbf{F}, t) \equiv \mathrm{e}^{\lambda(Y_t - K)}\mathrm{E}\left[\exp\left(\lambda \sum_{n=n^*}^{T-N} \beta_1^{(n)'} \varepsilon_n\right) \mid \mathbf{F}_{\mathrm{t}} = \mathbf{F}\right], \qquad \text{(A9)}$$

where n^* is the smallest integer $> t$.

The significance of G^H is as follows: if futures prices follow a martingale then the expected utility of terminal wealth of an agent who has committed to supply one barrel of oil at time T and price K is $-G^H/\lambda$. The optimal hedging strategy in such a case is given by

$$\Delta^H = \beta^{(\mathrm{n})}. \qquad \text{(A10)}$$

To prove Proposition 2 we need to show that the expected utility function of an agent with a supply commitment when futures prices are not martingales is

$$G^* \equiv G^S G^H \qquad \text{(A11)}$$

and the optimal hedging strategy is

$$\Delta^* \equiv \Delta^S + \Delta^H. \qquad \text{(A12)}$$

G^* satisfies the boundary condition

$$G^*(\mathrm{F}, X, T) = G^S(\mathbf{F}, X, T)G^H(\mathbf{F}, T) = -\frac{1}{\lambda}\mathrm{e}^{\lambda\{F(T,T) - K - X\}}. \qquad \text{(A13)}$$

Note that

$$G^H + dG^H = \exp\left\{\lambda \Delta^{H'} \cdot \mathrm{d}\mathbf{F}\right\} G^H(\mathbf{F}, t), \qquad \text{(A14)}$$

except when t is an integer, and that when t is an integer $\mathrm{E}[\mathrm{d}G^H \mid \mathrm{F}_t] = 0$.

Under trading strategy Δ G^* evolves as follows during the month (t noninteger):

$$\begin{aligned} &G^*(\mathbf{F} + \mathrm{d}\mathbf{F}, X + \Delta' \cdot \mathrm{d}\mathbf{F}, t + \mathrm{d}t) = \\ &G^S\left(\mathbf{F} + \mathrm{d}\mathbf{F}, X + (\Delta - \Delta^H)' \cdot \mathrm{d}\mathbf{F}, t + \mathrm{d}t\right) G^H(\mathbf{F}, t). \end{aligned} \qquad \text{(A15)}$$

From this, two things follow immediately: first, the optimal trading strategy is Δ^*; second, under this strategy G^* is a martingale.

All we now need to prove the first part of the proposition is to show that at the turn of the month, when t is an integer, G^* is a martingale. But we know that G^S and G^H are both martingale over the month end, so G^* wil also be martingale provided that:

$$\mathrm{E}\left[\mathrm{d}G^S\,\mathrm{d}G^H \mid \mathbf{F}_t\right] = 0 \text{ for integer } t. \tag{A16}$$

This proves the first part of the proposition.

For breakeven $G^*(\mathbf{F}, X, 0) = G^S(\mathbf{F}, X, 0)$, so $G^H(\mathbf{F}, 0) = 1$. The second part of the proposition then follows from Equation (A8).

Appendix B: Relationship Between the Extrapolation Model and Dynamic Term Structure Models

In a k-factor dynamic term structure model there are k factors which follow a joint diffusion process. Using the standard no-arbitrage arguments and specifying the form of the price of risk, one can in general express any derivative claim as a function (implicit or explicit) of the state variables. So, in particular, there is a *pricing function* which relates the price of a futures contract to its maturity and to the value of the state vector **S**. Suppose that $N \geq k$, so that there are at least as many futures contracts trading at any one time as there are factors. Then the value of the state vector (which may not be directly observable) can be inferred exactly from prices of the traded contracts (which can be observed directly) by inverting the pricing function, under quite weak conditions.[10] The price of any other contingent claim can also be expressed in terms of the prices of traded contracts. In particular, the opening price of a contract is a function of the prices of the other contracts which are trading. For any integer t,

$$F(t, t+N) = h(\mathbf{F}_t), \tag{B1}$$

where $h(\cdot)$ is some function which depends on the model. Comparing this with Assumption 6 shows two major differences. Assumption 6 has a noise term and is linear in $\mathbf{F}_t$, while Equation (B1) holds exactly and may not be linear. Now, the linearity assumption (Assumption 6) holds exactly (with zero error) if the pricing function can be written as a sum of no more than N multiplicatively separable functions of the state vector and time to maturity, in the following way:

$$F(t, n) = \alpha(n-t) + \sum_{i=1}^{N} f_i(\mathbf{S}) g_i(n-t), \tag{B2}$$

where $\alpha(\cdot)$ and the $\mathrm{g_i}(\cdot)$ are functions of the time to maturity and $\mathrm{f}_i(\cdot)$ is a function of the state variables.

To see that this is true, find the α, $\beta = \{\beta_j \mid j = 1, \ldots, n\}$ which solve the equation set:

$$\begin{gathered} g_i(N) = \sum_{j=1}^{N} \beta;\, g_i(N-j) \text{ for } i = 1, \ldots, N \\ \alpha = \alpha(N) - \sum_{j=1}^{N} \beta_j \alpha(N-j) \end{gathered} \tag{B3}$$

(if F can be expressed as the sum of fewer than N multiplicative terms, the solution

[10] For the pricing function to be invertible, the mapping of the state vector into the price space has to be 1-1.

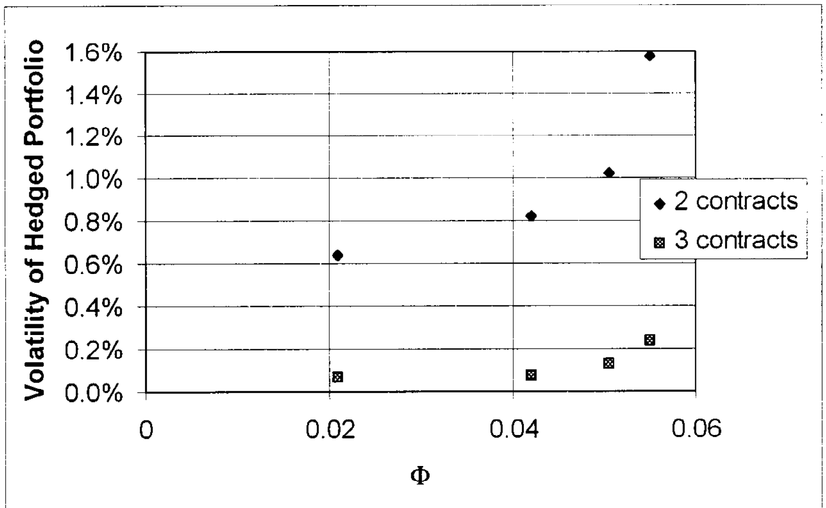

Figure 2
Effectiveness of the extrapolation method in a Gibson–Schwartz world
The extrapolation method is used to compute the optimal hedge for a five-year contract in a Gibson–Schwartz world. Either two or three hedging contracts with nine months or less to expiration are used. The graph shows the instantaneous volatility of the hedged portfolio as a function of $\Phi = \sigma_2^2/\kappa^3$, where σ_2 is the volatility of the convenience yield and κ is its rate of mean reversion. The different observations correspond to scenarios I–IV in Table 1. The conjecture is that the lower the value of Φ, the closer the model is to linearity and the better the extrapolation method works.

will not be unique and there will be an infinite set of hedging strategies), then it follows immediately that

$$F(t, t+N) = \alpha + \sum_{i=1}^{N} \beta_i F(t, t+N-i). \tag{B4}$$

Equation (B2) may not hold exactly for the model of interest for any finite value of N. Reasoning informally, we may conjecture that the more closely the approximation holds, the better the extrapolation model will perform. To explore the implications of approximating the pricing function in this way, consider the GS model. The futures price is a function of S, the spot price, and d, the convenience yield:

$$F(t, n) = S\mathrm{e}^{A(n-t)-B(n-t)d}, \tag{B5}$$

where $A(\cdot)$ and $B(\cdot)$ are functions of time to maturity and depend on the parameters of the diffusion process and the price of risk. Using a Taylor's series expansion

$$F(t, n) = \sum_{i=0}^{\infty} f_i(\mathbf{S}) g_i(n-t) \tag{14}$$

where

$$f_i(\mathbf{S}) \equiv S\frac{(\overline{d}-d)^i}{i!}; \tag{B6}$$

$$g_i(n-t) \equiv e^{A(n-t)-B(n-t)\overline{d}}B(n-t)^i$$

$$\text{and} \quad \overline{d} \equiv \mathrm{E}[d]$$

This has the form of Equation (B2) with $N=\infty$. To measure the impact of ignoring higher-order terms, note that the ratio of successive terms in the series is

$$\frac{f_i(\mathbf{S})g_i(n-t)}{f_{i-1}(\mathbf{S})g_{i-1}(n-t)} = \frac{(\overline{d}-d)}{i}B(n-t). \tag{B7}$$

The unconditional variance of the convenience yield, d, is $\sigma_2^2/2\kappa$, where σ_2 is the volatility of the convenience yield and κ is the rate at which it mean reverts. $B(\tau)$ is equal to $(1-e^{-\kappa\tau})/\kappa$, which increases from 0 to $1/\kappa$ as τ goes from 0 to infinity. So the variance of the right-hand side of Equation (B7) is of order σ_2^2/κ^3, a ratio we denote by ϕ. We conjecture that the approximation of the infinite series by just N terms will be better the smaller ϕ is. The extrapolation model will work best when the convenience yield is not too volatile and any shocks die out quickly.

To test the conjecture that the quality of the hedge depends on ϕ we make use of the fact that Schwartz (1997) gives four different sets of estimates for the model parameters based on different sample periods and using different contract maturities. The statistic ϕ varies from 0.021 to 0.055, with a value of 0.042 in the base case. There is a reasonable correlation between the level of ϕ and the volatility of the hedged portfolio, as shown in Figure 2.

References

Baesel, J., and D. Grant, 1982, "Optimal Sequential Futures Trading," *Journal of Financial and Quantitative Analysis*, 17, 683–695.

Bessembinder, H., 1991, "Forward Contracts and Firm Value," *Journal of Financial and Quantitative Analysis*, 26, 519–572.

Brennan, M. J., 1991, "The Price of Convenience and the Valuation of Commodity Contingent Claims," in Lund, D. and B. Oksendal (eds.), *Stochastic Models and Option Values*, Elsevier Science Publishers, New York.

Ederington, L. H., 1979, "The Hedging Performance of the New Futures Markets," *Journal of Finance*, 34, 157–170.

Engle, R., and C. Granger, 1987, "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251–276.

Froot, K. A., D. S. Scharfstein, and J. C. Stein, 1993, "Risk Management: Coordinating Corporate Investment and Risk Management Polilcies," *Journal of Finance*, 48, 1629–1658.

Gibson, R., and E. S. Schwartz, 1990, "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims," *Journal of Finance*, 45, 959–976.

Grant, D., 1984, "Rolling the Hedge Forward: An Extension," *Financial Management*, 13, 26–28.

Granger, C. W. J., and P. Newbold, 1986, *Forecasting Economic Time Series* (*2nd ed.*), Academic Press, New York.

Howard, C. T., and L. J. D'Antonio, 1991, "Multiperiod Hedging Using Futures: A Risk Minimisation Approach in the Presence of Autocorrelation," *Journal of Futures Markets*, 11, 697–710.

Johansen, S., 1991, "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Auto-regressive Models," *Econometrica*, 59, 1551–1580.

McCabe, G. M., and C. T. Franckle, 1983, "The Effectiveness of Rolling the Hedge Forward in the Treasury Bill Futures Market," *Financial Management*, 12, 21–29.

Mello, A., and J. Parsons, 1995, "Maturity Structure of a Hedge Matters: Lessons from the Metallgesellschaft Debacle," *Journal of Applied Corporate Finance*, 8, 106–120.

Ross, S., 1995, "Hedging Long Run Committments: Exercises in Incomplete Market Pricing," working paper, Yale School of Management.

Schwartz, E. S., 1997, "Presidential Address: The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging," *Journal of Finance*, 52, 923–973.

Smith, C. W., and R. M. Stulz, 1985, "The Determinants of Firms' Hedging Policies," *Journal of Financial and Quantitative Analysis*, 20, 391–405.