Working Paper

INTERNATIONAL MONETARY FUND

IMF WORKING PAPER

WP/91/93

INTERNATIONAL MONETARY FUND

Research Department

Forecasting Accuracy of Crude Oil Futures Prices

Prepared by Manmohan S. Kumar *

Authorized for distribution by Peter Wickham

October 1991

Abstract

This paper undertakes an investigation into the efficiency of the crude oil futures market and the forecasting accuracy of futures prices. Efficiency of the market is analysed in terms of the expected excess returns to speculation in the futures market. Accuracy of futures prices is compared with that of forecasts using alternative techniques, including time series and econometric models, as well as judgemental forecasts. The paper also explores the predictive power of futures prices by comparing the forecasting accuracy of end-of-month prices with weekly and monthly averages, using a variety of different weighting schemes. Finally, the paper investigates whether the forecasts from using futures prices can be improved by incorporating information from other forecasting techniques.

JEL Classification Numbers:
A10, C22, C52, E37

* I am grateful to Bijan Aghevli, Peter Wickham, Charles Adams, William Perraudin and Blair Rourke for helpful comments. Raja Hettiarachchi provided excellent computing assistance.

Contents

Appendix Tables

Summary

This paper analyses the development of futures markets in crude oil and examines the accuracy of forecasts obtained using futures prices. Futures markets in crude oil have grown extremely fast during the last five years, and the volume of trade in futures transactions far exceeds the trade in the spot market. The depth and breadth of the futures markets suggest that forecasts obtained from futures prices are unlikely to be biased and are likely to provide a relatively accurate indication as to the future course of spot prices.

A number of empirical exercises are undertaken to evaluate the "unbiasedness" hypothesis, and the accuracy of the forecasts. An extensive dataset, covering the period from the inception of crude oil trading on the New York Mercantile Exchange to 1990, is utilized for this purpose. An analysis of the mean excess returns that could be obtained from holding futures contracts did not appear to suggest any systematic bias in the futures prices. This result complemented the results of the comparison of forecasts using futures prices with forecasts using a random walk model, which showed that the former provided more accurate forecasts for all forecast horizons. As the length of the forecasting horizon increased, however, the accuracy of both types of forecasts diminished markedly.

An analysis of intra-month futures prices suggested some marginal improvement in forecasting accuracy for distant horizons, compared with the end-of-the month prices. When weights on intra-month prices were determined endogenously, it appeared that the weighting scheme should be related to the length of the forecast horizon. Futures-prices forecasts were also more accurate compared with forecasts obtained from time-series models as well as judgmental and econometric forecasts. Combining forecasts from alternative techniques, however, yielded only a marginal improvement in terms of variance of forecast errors.

The empirical results strongly suggest that futures prices provide forecasts that are, in general, superior to those obtained from alternative techniques for short term horizons. For more distant horizons, their accuracy does diminish markedly; however, even for these horizons the futures forecasts are no worse, and are often better, compared with those obtained from alternative techniques.

I. Introduction

This paper undertakes an investigation into the efficiency and forecasting accuracy of crude oil futures prices. Efficiency of the market is analyzed in terms of the expected returns from trading in the futures contracts. Accuracy of futures price forecasts is analyzed by comparing it with the accuracy of forecasts obtained using a variety of other techniques including random walk and time series models, as well as forecasts using judgmental and econometric techniques. The paper also explores the predictive power of futures prices by comparing the forecasting accuracy of end-of-month prices with weekly and monthly averages, using a variety of weighting schemes. Finally, the paper examines the improvement in forecasting accuracy when futures prices are combined with forecasts from alternative sources.

The empirical analysis in the paper is based on the New York Mercantile Exchange's (NYMEX) futures contracts, which constitute by far the most active crude oil futures trading in the world. Trading in crude oil contracts started in April 1983; however, in the initial period, the market was relatively limited. Since 1985, activity has increased at a very fast pace indeed. While there is a growing futures market for Brent crude oil at the International Petroleum Exchange in London, until recently trading on this was limited, and contracts were available for only a few months ahead. On the NYMEX, in addition to crude oil, there are active futures markets in other energy products, including heating oil, gasoline, and natural gas. But the combined activity in these markets still falls short of the activity in crude oil futures.

The empirical exercise analyzes the behavior of futures prices from June 1985, by which time the crude oil futures market had become highly developed, to October 1990. The analysis considers maturities ranging from one month to nine months. Although during the last year there has been a sharp increase in the maturities available with contracts for up to nearly three years ahead, for most of the period under discussion the trading was confined to a period of around nine months to one year.

The discussion in the paper is organized as follows: Section II examines the emergence and growth of futures markets in crude oil trading during the last decade. Section III provides the empirical evidence on efficiency, while Sections IV and V provide evidence on forecasting accuracy, which compares the accuracy of futures forecasts with that of forecasts from a variety of alternative models; Section V also analyzes the improvement in forecasting accuracy which may result from combining futures forecasts with those from other techniques. A last section summarizes the main results and notes their implications for forecasting oil prices for different forecast horizons.

II. Futures Market in Crude Oil

1. Development of futures market

Futures markets serve two interrelated purposes: the first is a provision of an organized forum allowing agents to undertake hedging, or speculation. Their second purpose is that related to price discovery; the price of futures contracts provides a summary or consensus view, based on market trading, of the participants' expectations with regard to the future course of prices. For the market to perform either of these functions efficiently, there has to be sufficient activity in terms of the volume of contracts traded on a regular basis.

While there are a number of international forward and futures markets in crude oil, by far the largest futures market is on the New York Mercantile Exchange (NYMEX). Since the introduction of the oil futures contracts on NYMEX in April 1983, volume of trade in oil contracts has grown almost exponentially. In 1983, the average daily volume was around 1,200 contracts, (with each contract representing 1,000 barrels of oil). 1/ By the end of 1990, the daily average exceeded 100,000 contracts, or 100 million barrels a day, making crude oil futures one of the most heavily traded of any futures contract. At the same time, the total number of contracts outstanding has increased from around 5,000 contracts to nearly 280,000 contracts, equivalent to 280 million barrels. In comparison, total world oil production is about 65 million barrels a day. Furthermore, because the futures price is set in open trading, it is accepted as the benchmark from which almost all other prices for crude oil are calculated.

It is worth noting that there are futures contracts in oil products which predate the trading of crude oil contracts. For instance, gasoline futures were introduced on NYMEX in 1978 while heating oil futures were introduced in 1980. 2/ During 1978 crude oil futures trading was begun on the NYMEX, but because of insufficient volume of trade the contract was withdrawn. One of the main reasons given for insufficient trading was the lack of adequate volatility in crude oil spot prices. As the discussion below shows, the low volatility in turn was due to the market structure whereby spot trading was very limited; most of the oil traded was on the basis of contractual arrangements between oil producers and international oil companies.

The significance of contractual arrangements can be judged from the fact that until the late 1970s, almost 90 percent of the world's oil was sold under long-term contracts based on prices set by the major oil producers, and the other 10 percent was bought and sold informally between the international oil companies. Beginning with the late 1970s, however,

---

1/ Each barrel includes roughly 42 gallons of oil. See Appendix Table 1 for detailed information on the size of the market.
2/ For details, see Hirschfeld (1983).

there was a fundamental change, with this historic system of long-term, fixed price contracts negotiated between oil producers and the international oil companies being largely replaced with an open market system. By the early 1980s, 90 percent of the world's oil was available on the spot market. The move to the preponderance of spot market transactions seems to have been accompanied by considerably greater short-term volatility in oil prices in the recent years. 1/

Thus, on a short-term basis (day-to-day, or week-to-week), crude oil prices during the 1980s were considerably more volatile than during the 1970s, when the prices were stable often for months at a time. The incidence of sharp abrupt swings in prices, however, does not seem to have been increased any marked way by the move to greater market role. If anything, it could be argued that by adjusting prices continuously to the supply-demand imbalances, the extent of the abrupt changes may be lessened. 2/

The antecedents of the change from contractual to spot price transactions lay in two major developments during the 1970s: (i) the world's major oil producers exercised greater control over their oil fields, allowing them to sell their oil to whomever offered the best deal; (ii) major oil companies that previously had owned the oil fields were cut loose to bid for crude oil wherever it was available.

At the same time, the very high real interest rates of the early 1980s, which made the storage of oil an expensive activity, encouraged the formation of the oil futures market. The emergence of this market meant that it was much cheaper for a user to acquire claims on oil by buying a contract that assured delivery when needed than it was to purchase and store actual oil. As with almost all futures trading, actual delivery is made in only a small proportion of contracts: instead the contracts usually are bought and sold over and over again, often many times in the same day, as financial instruments that provide a form of price and supply insurance to users and producers.

2. Futures contracts

The futures trading began in 1983 with delivery for a period of up to six months ahead. Initially, even this period seemed likely to easily satisfy the needs of traders, and most of the transactions were concentrated in the first two or three months. Fairly soon, however, trading in distant months became equally active and there was demand for longer maturities. Trading was then extended to nine months and by 1989 to twelve months. Beyond nine months, however, trading volume remained limited. Since July 1990, the contracts have been extended almost continuously; first, they were extended to 15 months, then to 18 months, and by April 1991, contracts were

---

1/ See, for instance, Anderson et al (1990).
2/ See, for example, Hampton (1991).

available for up to three years ahead. For the purposes of comparing forecasting accuracy, the most recent lengthening of maturities could not be utilized. For the earlier periods also, most active trading was limited to a period up to nine months ahead, as the paper focuses on trading at these maturities.

A critical factor affecting the accuracy of the futures prices' forecast is the way the prices are actually determined. To ensure that the determination of prices is completely transparent, there are some very specific rules for arriving at the "settlement price" for the crude oil futures contracts. As discussed in Appendix I, the "settlement price" is a weighted average of the transactions prices, towards the end of the trading session. This is the price used in the following analysis of the futures prices' forecasting accuracy (see Appendix I for details).

The contracts are specified for a specific crude--namely, the "West Texas Intermediate" (WTI) to be delivered at a specific point--but the futures market rules allow for delivery of six other grades against the WTI contract. 1/ An adjustment factor is added by NYMEX to the WTI price. Buyers taking delivery of alternative grades may appear to be concerned about getting unexpected or less-valued oil. But for most substitute crudes, the buyers are actually indifferent. That is because the marginal benefit to the refiner of utilizing these grades generally exceeds their additional cost and the NYMEX adjustment factor. 2/ There is, in any case, a strong and systematic relationship between the West Texas price and prices of other crudes. So that from the West Texas price, one can infer forecasts of prices for other crudes (see Appendix II for a detailed analysis of this issue).

In any case, as noted earlier, the futures prices frequently set the benchmark for the pricing of other crudes in the spot market. It is also worth noting that while the futures prices are not entirely unconstrained in the sense that there are exchange-determined upper and lower limits to price changes in any given day, the constraints are seldom if ever binding. Even when they are binding, the decisions taken usually reflect almost entirely the market conditions (see Appendix I).

1/ The delivery point is the town of Cushing in Oklahoma, U.S.A. The six other deliverable grades include two Algerian grades, two Nigerian grades, a Norwegian grade, and UK Brent Blend.

2/ See, for instance, Petroleum Intelligence Weekly (1988). According to data provided by Petroleum Database Services, which has individual computer models of all U.S. refineries, the extra profit from five of the six deliverable crudes, relative to WTI, exceeded the NYMEX value adjustment by 15 to 70 cents a barrel during the second quarter of 1988. North Sea Brent Blend, the most readily available substitute crude, was the only grade that appeared unattractive to buyers at Cushing since it showed little or no extra profit compared to WTI, but the NYMEX adjustment method still penalized Brent with a slight premium.

III. Efficiency of Oil Futures Markets

The first empirical issue to be analyzed is whether there is any bias in crude oil futures prices, that is, do these futures prices tend to exhibit any systematic behavior, which could mean that ex ante one could obtain positive returns from speculating in the futures market. If futures prices are to be useful in forecasting future spot prices, the expectation would be that there is no such bias. At any given time, there may be a number of factors which might lead to an expectation of an increase in the future spot price; equally, at other times, there could be factors leading to an expectation of a decline. So that on average, ex ante, there should be no systematic bias (see, for example, Kaminsky and Kumar (1990 and 1991)). There are a number of techniques for examining this efficiency, but the most rigorous and transparent is to examine the returns from holding futures contracts up to a given maturity.

Table 1 shows the mean excess returns from holding a futures contract and the corresponding t-statistic for the test of the null hypothesis of unbiasedness. The excess return is computed as the change in the contract's price between two different dates ranging from one to nine months, and covers the eight-year period 1983-90. Although mean excess return is positive for the first five forecast horizons, it is not significantly different from zero over any of the forecast horizons. (The tests of significance use standard errors corrected for autocorrelation using the "method of moments"). These results suggest, at least superficially, that the null hypothesis of a zero bias in the oil futures prices cannot be rejected. However, the evidence can be consistent with the presence of a time varying bias. Since there is evidence from other commodity markets that a time varying bias exists, it is useful to check whether there is such a bias in the oil futures market (see Kaminsky and Kumar (op.cit.)).

The procedure used for isolating such bias is to divide the sample into sub-periods according to whether the crude oil spot price was increasing or falling. The rationale for such a procedure follows from the results of a number of studies into the behavior of other commodity, and asset, futures prices. These results indicate that, in general, investors consistently underpredict the price of an asset when the asset is appreciating (for example, the U.S. dollar in the early 1980s) and systematically overpredict it when it is depreciating (as was the case after 1985, when the dollar started to depreciate). As Chart 1 shows, the spot oil price was increasing from January 1986 to January 1987 as well as from April 1988 to December 1989. The observations from these two sub-periods were pooled as were the observations from a period of declining spot prices. The results of computing the excess returns for the two periods of increasing and declining prices are given in Table 2. The realized excess returns during the period of increasing prices were significantly positive for the first four forecast horizons; for the other forecast horizons, the results were not significant. For the period of decreasing prices, there were again no significant excess

Table 1. Test of Unconditional Unbiasedness: Full Sample

Forecast Horizon (Months)	Excess Returns $(f_{t+i} - f_t)$	
	Mean	t-Statistic
1	0.0079	0.6135
2	0.0086	0.3620
3	0.0057	0.1723
4	0.0033	0.0772
5	0.0008	0.0159
6	-0.0020	-0.0305
7	-0.0039	-0.0590
8	-0.0069	-0.0957
9	-0.0107	-0.1369

Note: The t-statistics used standard errors of means corrected for autocorrelation (using "method of moments"). The i subscript refers to the forecast horizon (i=1 to 9). The sample is from September 1983 to October 1990.

CHART 1.

Crude Oil: Spot Prices and Excess Returns
(September 1984 - September 1990)

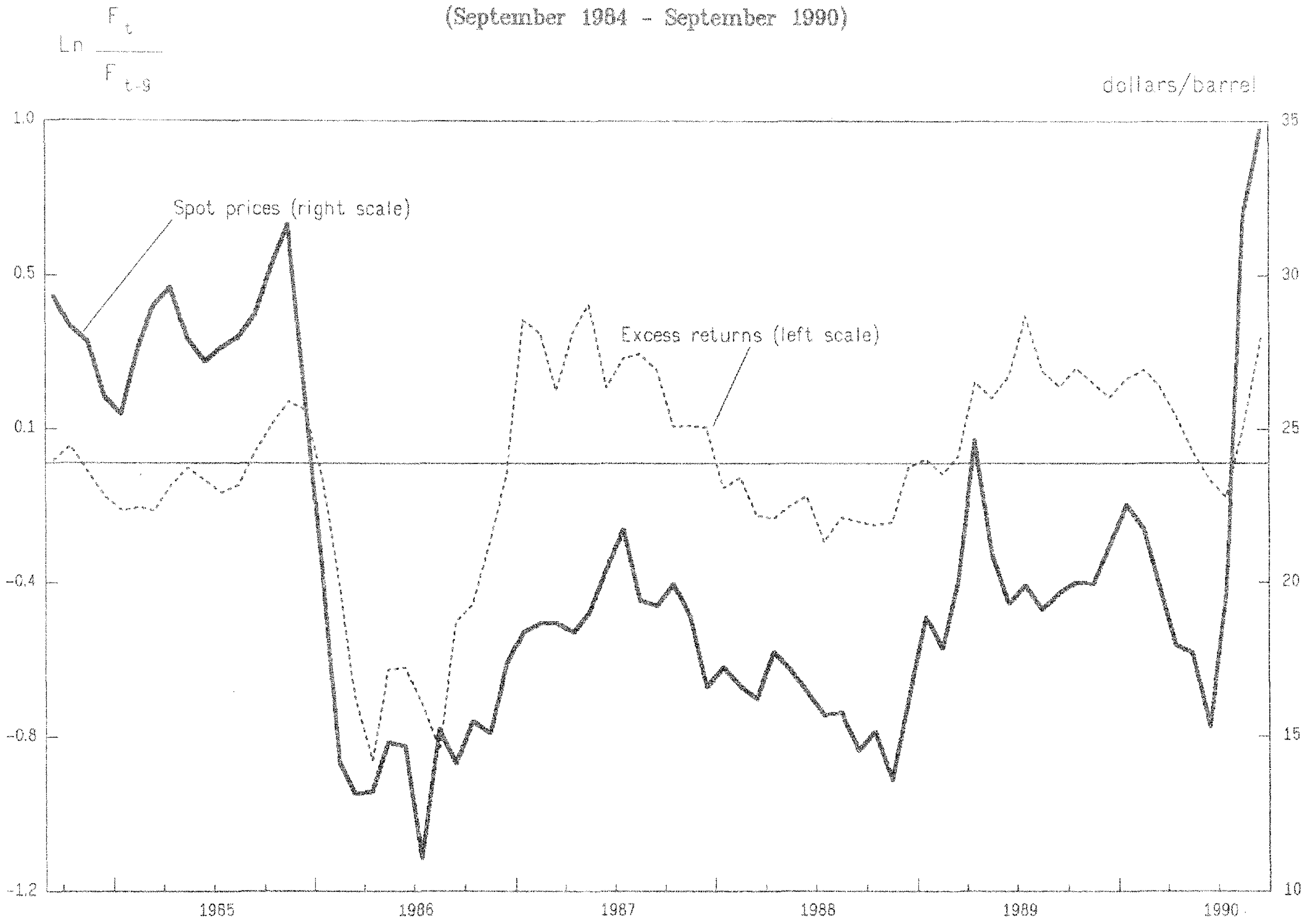


Table 2. Test of Unconditional Unbiasedness: Increasing and Decreasing Spot Prices

Movement of Spot Prices	Forecast Horizon (Months)	Excess Returns $(f_{t+i} - f_t)$ Mean	t-Statistic
Increasing	1	0.0460	2.3599
	2	0.0749	3.2646
	3	0.0996	4.1326
	4	0.1094	3.7480
	6	0.0721	1.1920
	9	-0.0042	-0.0413
Decreasing	1	-0.0016	-0.1294
	2	0.0003	0.0145
	3	0.0102	0.3352
	4	0.0276	0.7448
	6	0.0585	1.1846
	9	0.08944	1.2629

Note: The t-statisctics use standard errors of means corrected for autocorrelation (using "method of moments").

returns for any of the maturities. As has been noted in earlier studies, this evidence, weak and in general statistically insignificant, does not necessarily imply market failure or that the futures are necessarily biased predictors of future spot prices.

IV. Futures Price Forecasts

The accuracy of forecasts obtained from using futures prices is measured in relation to the spot price prevailing on the day of the maturity of any given contract. Forecasts using futures prices are generated using prices of contracts for delivery up to ten months ahead. Since oil contracts are terminated in the month preceding the contract month, forecasts up to nine months ahead are thereby obtained. More precisely, trading for any contract is terminated on the third business day, prior to the twenty-fifth calendar day of the month, preceding the delivery month. The reason for this apparently peculiar timing is that for delivery, the pipeline space must be reserved by the 25th of the month. 1/

1. End-of-month prices and random walk

In the initial exercise, futures prices on the last trading day of the month are utilized, under the premise that the latest price would incorporate the latest information and thereby provide the most accurate and up-to-date forecast. In a subsequent exercise, however, prices prevailing at different times during the month, weighted according to specified criteria, are also examined. Accuracy of forecasts using futures prices is investigated in relation to forecasts obtained from a number of different techniques. These include forecasts using the random-walk model, time-series models, judgmental methods, and econometric models.

Consider first a comparison of the forecasts from end-of-month futures prices with the forecasts from a 'random walk' model. The latter can be regarded as postulating that spot prices at period t are the best unbiased predictor of spot prices at any future period T. 2/ The rationale for this hypothesis is that a commodity such as crude oil is subject, on a day-to-day basis, to a large number of influences on its demand and supply, and on its price. Hence, at any given time, the spot price itself reflects the consensus view as to the current market situation and provides the best guidance as to the future course of prices.

1/ It is worth noting that the delivery date was changed in 1985 when it was based on the fifth day, prior to the twenty-fifth calendar day. An earlier study by Ma (1989) for the period 1984-86 apparently used the same delivery date. Given the extreme sensitivity of prices near the maturity date, the difference of even a couple of days can be important. For a somewhat different methodology, see, for instance, Dominguez (1987).

2/ An related model could be that of a random walk with drift; there was, however, no empirical support for the drift factor.

Table 3 provides the results of a comparison of the accuracy of the futures prices, with that of the random walk model. Forecasts are examined for a period ranging from one to nine months. Forecast error is defined as follows:

$$P^j_{t,T} - S_T \tag{1}$$

where $P^j_{t,T}$ is the forecast made at time t for a future spot price at time T by using the weighting scheme j (discussed below). Thus in the case of the futures forecast, the price at maturity is taken; for the random walk model

$$P^j_{t,T} = S_t$$

S_T is the realized spot price at time T; T-t denotes the forecast horizon, 1 *to 9 months*.

Three different criteria for comparing forecast accuracy are utilized. The first is the mean absolute errors (MAE). This is the absolute value of the deviation of the predicted value from the realized value. Secondly, and the main one, is the root mean square error, (RMSE), which attaches a higher weight to larger absolute errors. Thirdly, Theil's "U" statistic which, in addition, adjusts for trend changes, is utilized.

The results in Table 3 tabulate the forecast errors from using these three forecast accuracy criteria for one to four months, and six and nine months, respectively. These results suggest two main conclusions: the first is that for virtually all forecast horizons, and for the different accuracy criteria, futures prices provide more accurate forecasts than the random walk model. The second conclusion is that as the length of the forecast horizon increases, the accuracy of the forecasts, whether using futures prices or the random walk model, diminishes markedly. In general, however, as Charts 2 to 4 also emphasize, futures prices provide a fairly accurate forecast of future spot prices for up to a six-month forecast horizon.

To the extent that the random walk model may be regarded as "naive", an obvious next step would be to examine whether one can improve upon the accuracy of the futures prices by using other more sophisticated but readily available econometric or time-series models. Before considering these *alternatives, it is worth enquiring whether on an a priori basis, it is* appropriate to use anything but the latest price. One hypothesis would be that if the market is efficient, and the results discussed in Section III above suggest that it is, then it must be the case that the latest price embodies all the relevant information and provides the best possible prediction of the future spot price. Hence combining the latest price with other (preceding) prices would not lead to any increase, and may actually lead to a decline, in the forecast accuracy. An alternative hypothesis would be that if there is a time varying risk premia with a transitory and a permanent component, then an averaging procedure may dampen, or even cancel out, the temporary component. Such a procedure may, therefore, improve on

Table 3. Comparison of Futures Prices and Random Walk

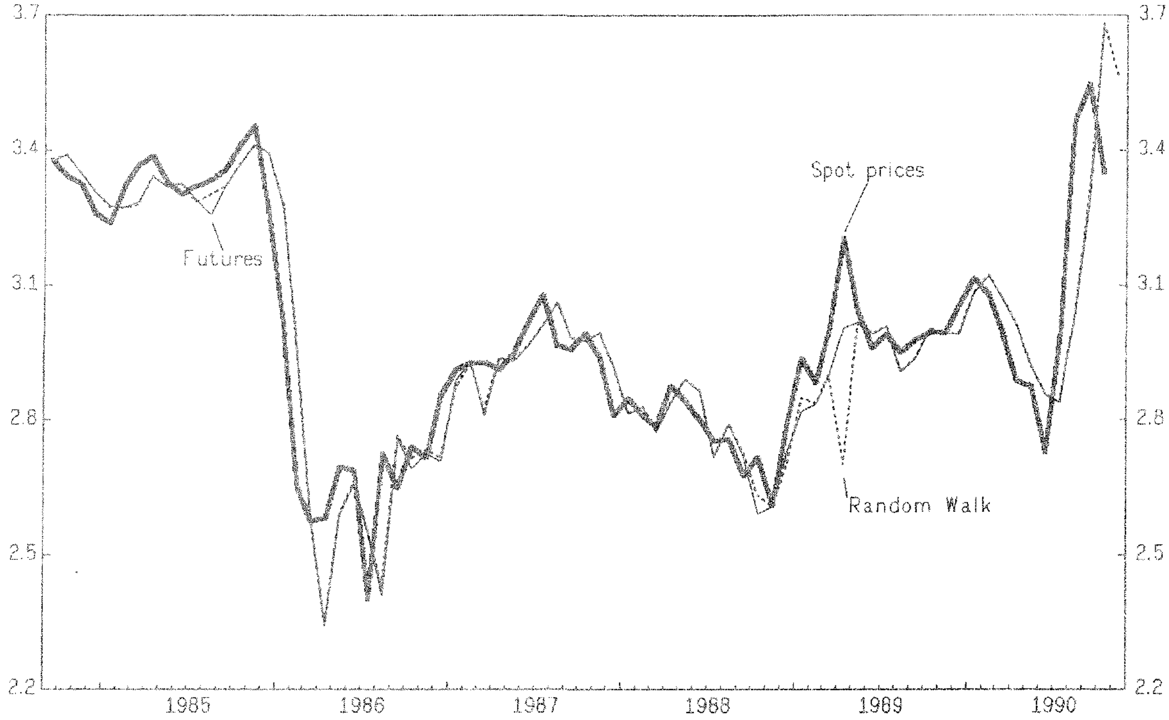
Forecast Accuracy Criterion	Horizon	Futures Prices	Random Walk
1. MAE	1	0.0870	0.0921
	2	0.1436	0.1438
	3	0.1722	0.1798
	4	0.1875	0.1925
	6	0.2260	0.2355
	9	0.2676	0.2836
2. RMSE	1	0.1104	0.1211
	2	0.1717	0.1799
	3	0.2080	0.2188
	4	0.2240	0.2378
	6	0.2573	0.2741
	9	0.2929	0.3172
3. Thiel's 'U'	1	0.0369	0.0404
	2	0.0573	0.0600
	3	0.0694	0.0730
	4	0.0747	0.0793
	6	0.0859	0.0914
	9	0.0977	0.1058

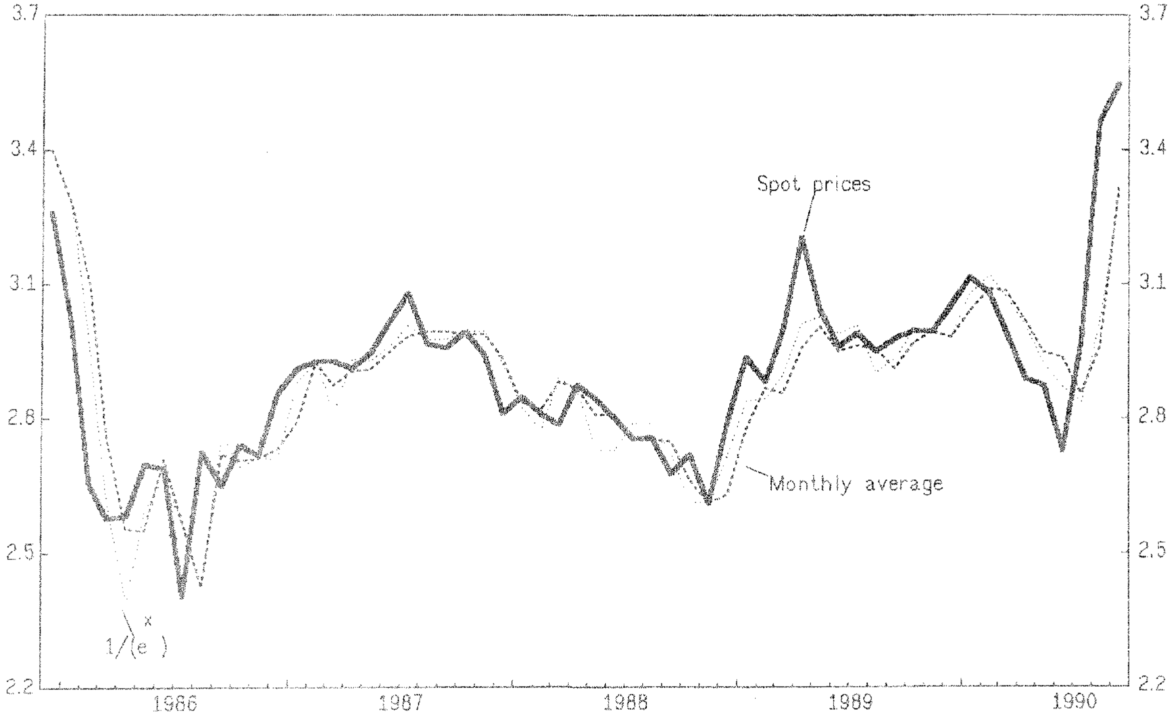
Note: Forecasts are for the period December 1985 to October 1990.

CHART 2.

Crude Oil: Spot prices and one month ahead Forecasts 1/

September 1984 - November 1990

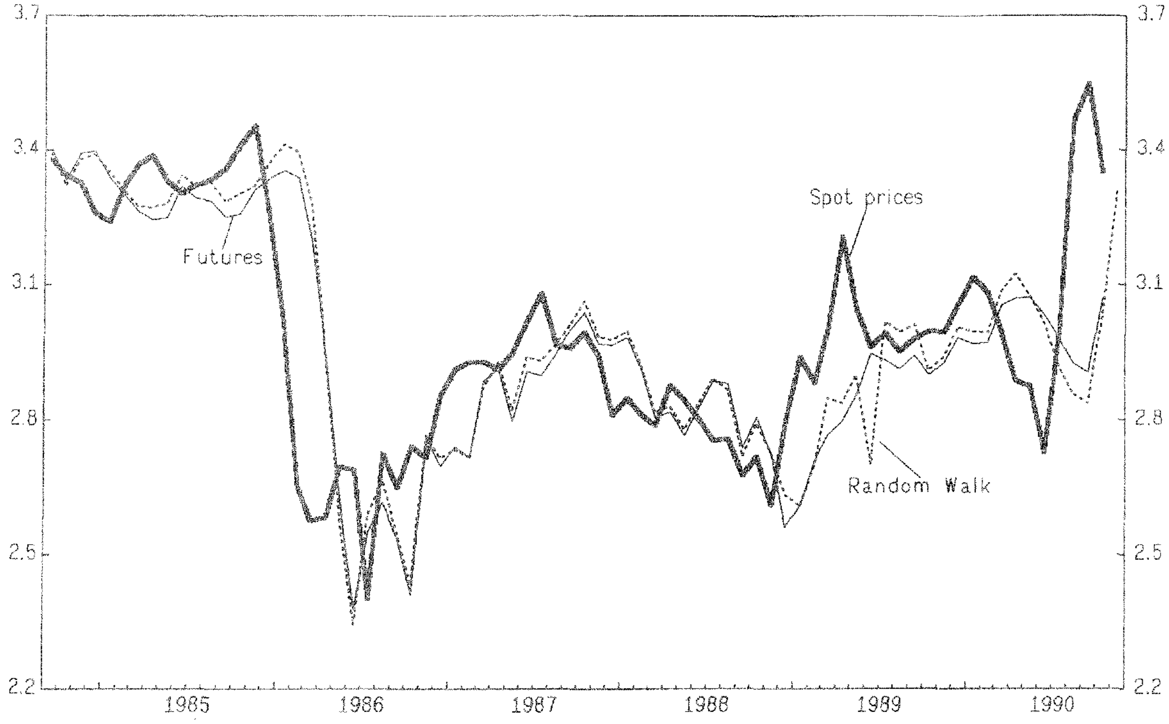


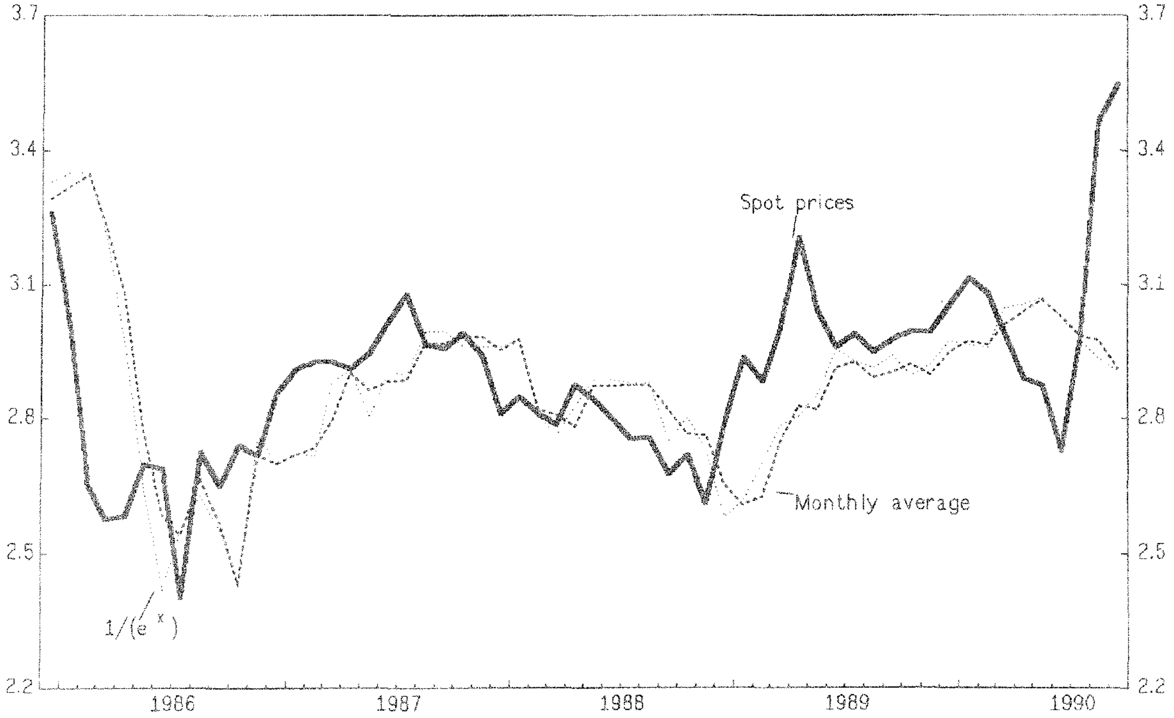


1/ Price (logarithms, dollars/barrel).

CHART 3.

Crude Oil: Spot prices and three months ahead Forecasts 1/
September 1984 - November 1990



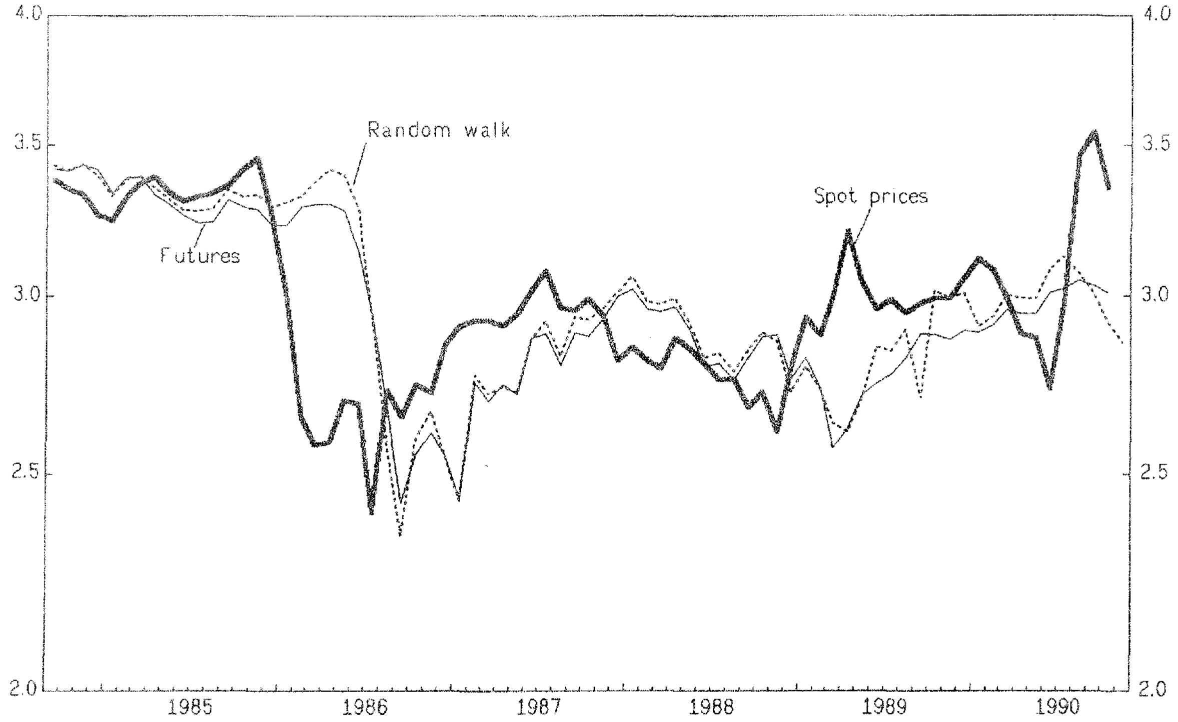


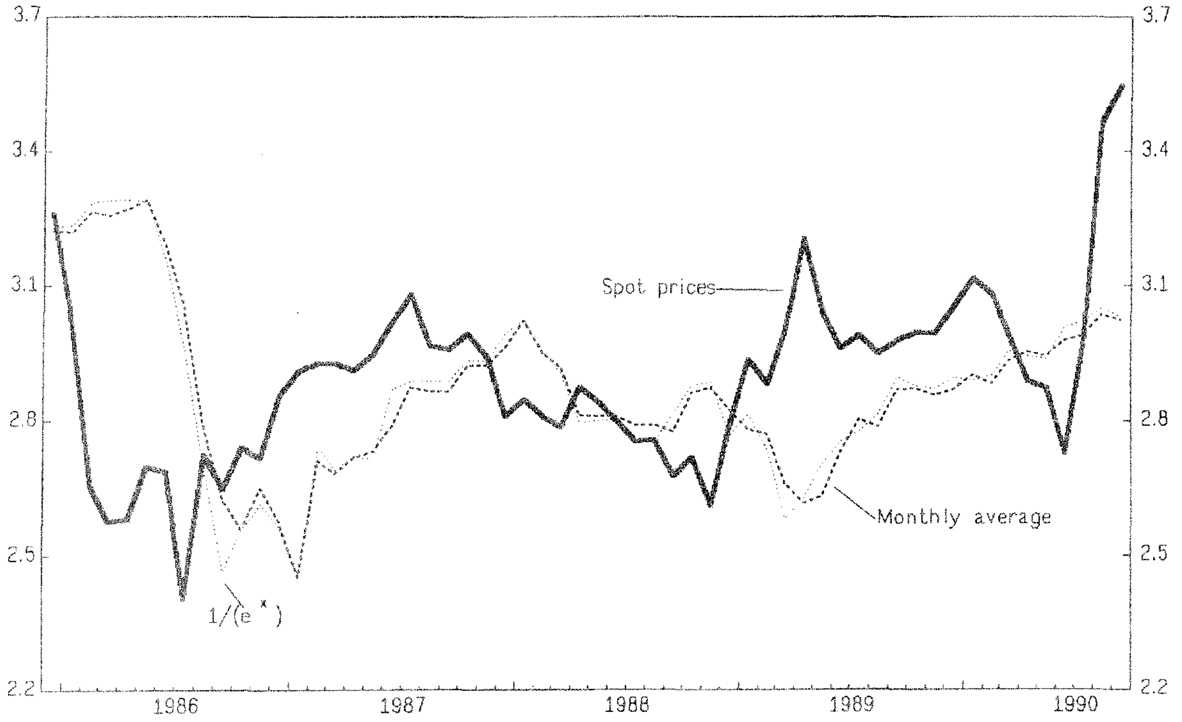
1/ Price (logarithms, dollars/barrel).

CHART 4.

Crude Oil: Spot prices and six months ahead Forecasts 1/

September 1984 - November 1990





1/ Price (logarithms, dollars/barrel).

the end-of-period prices. It might also be argued that to the extent that there are speculative bubbles in the market some sort of averaging procedure might improve the forecasts. In this regard the results of a number of recent studies into the efficiency of different asset markets are quite illuminating. Most of these studies suggest that even in markets that are regarded as approximately efficient, there may be evidence of speculative bubbles. For instance, Lo and Mackinley (1988) find a high degree of autocorrelation in weekly stock returns, while Fama and French (1988) and Poterba and Summers (1989) found similar results for returns spanning one year or more.

2. Predetermined weighting schemes

Given the above considerations, it may appear worthwhile to examine whether using prices prevailing during the month provide any incremental improvement in forecast accuracy of futures prices. The problem then is essentially an empirical one of choosing the optimal weighting scheme for combining different prices. The case of the end-of-month price would be simply a special case where the weight on intra-month prices is equal to zero. The end-of-month price is compared with three other schemes in which weights are also predetermined:

(i) Last five days' price: this is a simple average of the futures' (settlement) price on the last five trading days of any given month.

(ii) Monthly average: this is a simple average of the futures' price for each of the trading days in any given month. Since the number of trading days in the month can vary from 20 to 23 days, the average is based on closing prices on the last 20 trading days in the month.

(iii) Exponentially declining weights: The weights for the 20 trading days are computed according to the inverse of the exponential function; that is, the weights, W_i, are defined as follows: $W_i = 1/e^i$, $i=1,2...n$. Here 1 refers to the latest observation which is given the greatest weight, 2 refers to the last but one observation and so on.

In addition to the above four schemes, a scheme whereby weights were determined endogenously to minimize the forecast error (by using a maximum likelihood function) was also applied. (This methodology and its results are discussed presently.)

Consider first the results for the above four schemes, given in Table 4. For the one month ahead forecast, using end-of-the-month price yields the smallest forecasting error. According to the RMSE criterion, the monthly average gives forecast errors which are around 15 percent higher than the end-of-the-month price and the exponential scheme. It is worth noting, however, that monthly average forecast errors are still smaller than the errors generated by the random walk model (last column). As the forecast horizon increases, the difference in the forecast accuracy of the end-of-the month and exponential scheme vis-à-vis the monthly average declines monotonically, so that for the six-month ahead forecasts there is virtually no difference in the errors generated by the different schemes. As results in Appendix Table 2 indicate, this remains the case up to nine months forecast when the criterion is the MAE; however, when RMSE is the criterion, the monthly average has a negligible edge over the other schemes. It is also worth noting that regardless of the weighting scheme, the forecast errors using the random walk model are always larger than those using futures prices. Furthermore all forecast errors increase monotonically with the increase in the forecast horizon. (See Chart 5 for a comparison of the actual spot prices and forecasts for 9 month horizons respectively.)

3. Endogenous weighting scheme

The above four schemes apply predetermined weights to futures prices. An interesting extension would be to see if the weighting scheme could be determined endogenously by the data, and whether that reduces the out-of-sample prediction errors. This was examined by fitting the following equation to the data on spot and futures' prices:

$$\log S_T = \alpha + \sum_{i=1}^{k} \lambda^i \log P_{t,T} + \epsilon^t \tag{2}$$

Here S_T denotes the realized spot price at time T and $P_{t,T}$ denotes the futures price at time t for the contract maturing at T; α and λ^i are the estimated parameters. The equation was estimated separately for each of the nine forecast horizons. Given the non-linearity, the estimation procedure used a standard maximum likelihood function (MINDIS). Initially the equation was estimated using daily observations; that is, K = 20 and S_T was regressed on 20 variables constituting the prices at the end of each of the trading days in any given month. 1/ However, because of considerable collinearity between the daily observations, (for almost all the forecast horizons) it was not possible to achieve convergence. Therefore, estimation was undertaken using prices at the end of every fifth day (using data from

---

1/ As noted above, given the variable number of trading days, the last 20 days in the month were utilized.

Table 4. Comparison of Alternative Weighting Schemes for Futures Prices

Forecast Horizon (Months)	Forecasting Method				
	Futures prices				
	Last day	Last five days	Monthly average	Exponentially declining weights	Random Walk
1	0.1244	0.1268	0.1432	0.1248	0.1383
2	0.1984	0.1975	0.2067	0.1975	0.2057
3	0.2355	0.2356	0.2456	0.2351	0.2503
4	0.2527	0.2551	0.2558	0.2531	0.2716
5	0.2747	0.2773	0.2773	0.2751	0.2955
6	0.2929	0.2946	0.2945	0.2926	0.3143
7	0.3089	0.3086	0.3076	0.3081	0.3313
8	0.3207	0.3225	0.3204	0.3209	0.3447
9	0.3312	0.3301	0.3309	0.3311	0.3617

Note: Forecasts are for the period December 1985 to October 1990. Criteria for accuracy is the "Root Mean Squared Forecast Error."

1986 to 1988). Using the estimates of λ^i, normalized weights (summing to 1) were computed for each of the forecast horizons. These weights were then used to compute the out-of-sample forecasts and forecast errors for the period November 1988 to October 1990. In order to compare these forecasts with those obtained from using extraneous weights, forecast errors were recomputed for this period for the end-of-the-month, monthly average and the exponentially declining weights.

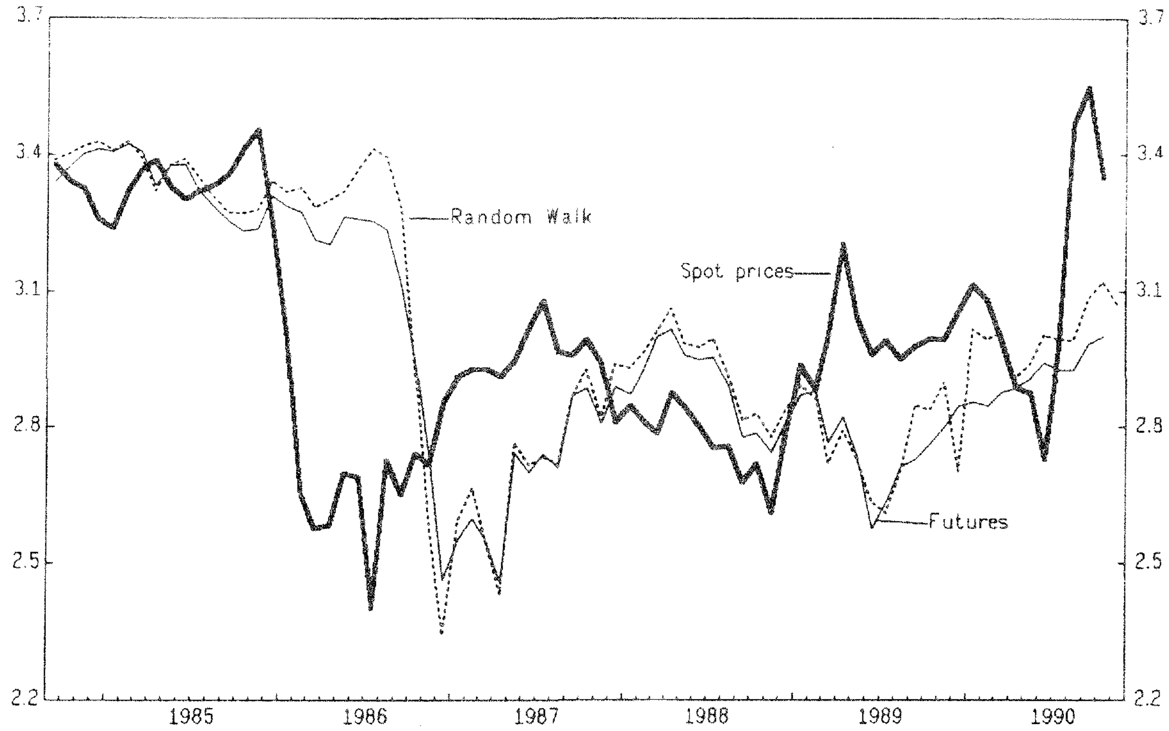
Results of the above exercise are provided in Table 5 and Chart 6. Consider first Chart 6, which compares the weighting scheme for the different forecast horizons as generated by the estimation, and the extraneous weights for the other schemes. (These other weights are, of course, invariant across forecast horizons.) As the chart shows, the pattern of weights is related to the length of the forecast horizon. The weights decline gradually with lag length when the horizon is distant, but for a nearby horizon the decline is very steep. For instance, when undertaking one month ahead forecast, according to the estimation, greatest weight should be placed on the latest futures price. But in undertaking a nine-month ahead forecast the average should be relatively more equally weighted. This result is fairly plausible and can be explained in a number of ways. One explanation may be in terms of the deficiency of lagged prices with respect to the information they contain. This deficiency is more serious relative to a short forecast than to a long forecast. That is, for the short forecast, the information (or the consensus) embodied in the latest price is likely to be much more important than the information in the lagged prices, and hence should be given the greatest weight. For the distant forecast, given the relatively greater uncertainty, somewhat greater weight can be given to the lagged prices. Some support for the above result, and the explanation, is found in a rather different study recently undertaken by Feinstein (1989). In that study, a volatility indicator for the stock market was constructed, by using weighted average of current and lagged implied volatilities from call options. That study reached very similar conclusions as to the relationship between weights and the forecast horizon.

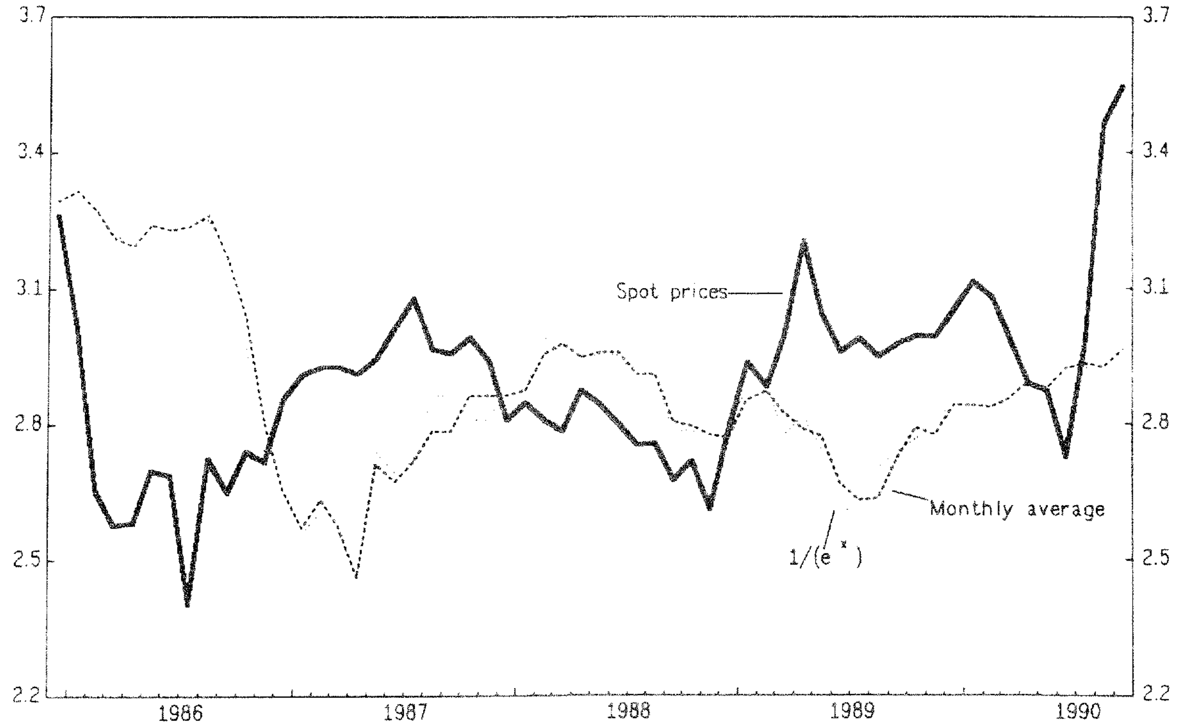
Next consider how the forecast errors using these endogenously determined weights compare with the errors generated by the other weighting schemes. The new weights were used to construct forecasts for the period November 1988 to October 1990. Given that these forecasts are for the out-of-sample period, it would not necessarily be the case that the forecast errors would be smaller than the other weighting schemes. This is borne out to some extent by the results in Table 5. Compared to the simple monthly average (Scheme II above), the forecast errors using the new weights are generally smaller. However, comparing the last-day forecast with the new weights forecast, shows that apart from the first and the third month, the former has marginally smaller errors.

CHART 5.

Crude Oil: Spot prices and nine months ahead Forecasts 1/

September 1984 - November 1990





1/ Price (logarithms, dollars/barrel).

CHART6: WEIGHTING SCHEMES

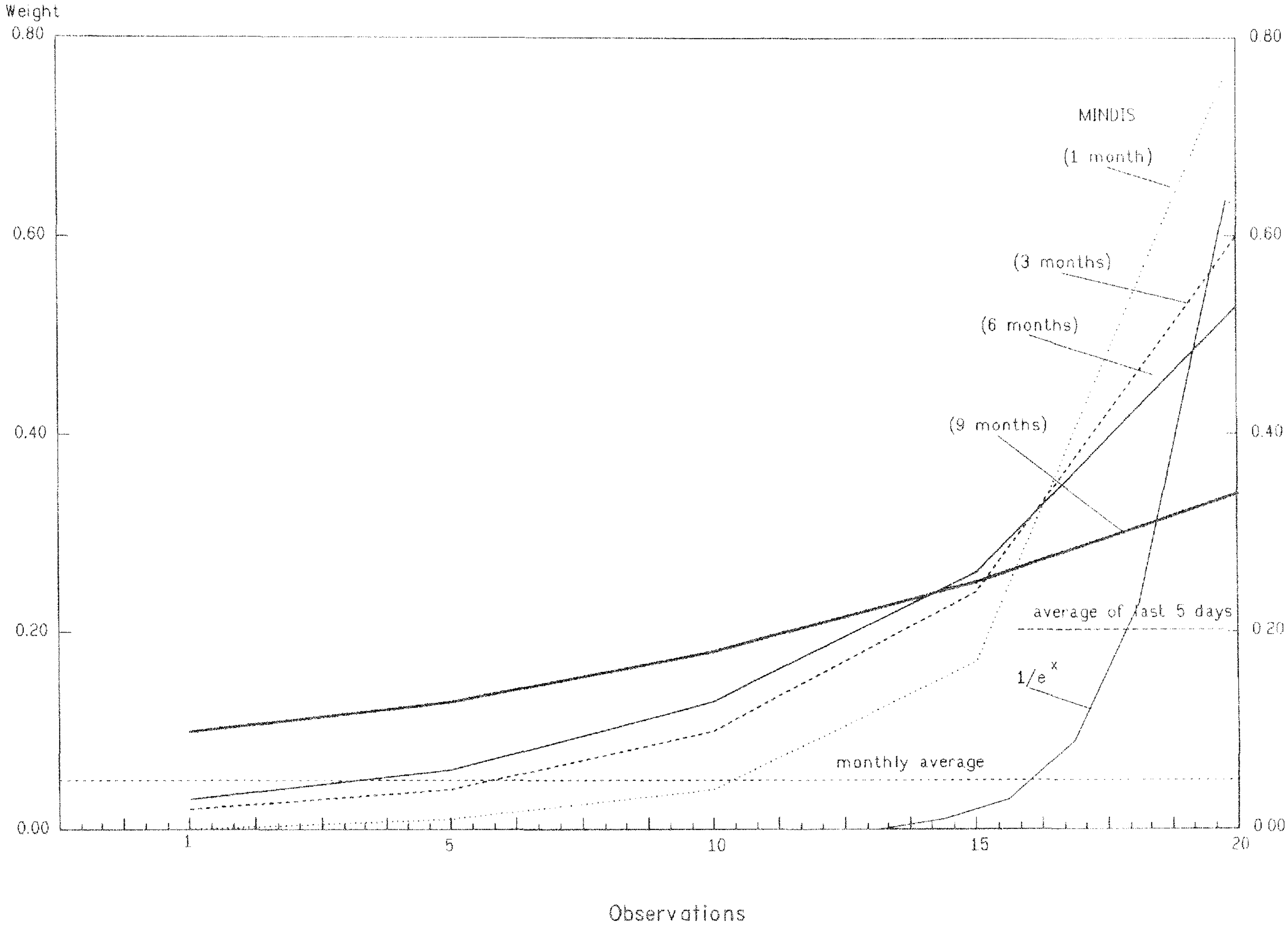


Table 5. Comparison of "MINDIS" Weighted Forecasts

Forecast Horizon	Root Mean Square Errors			
	"MINDIS" weighted	Last day	Monthly average	Exponentially declining weights
1	0.1385	0.1390	0.1622	0.1416
2	0.2189	0.2182	0.2288	0.2185
3	0.2454	0.2481	0.2436	0.2471
4	0.2509	0.2483	0.2534	0.2486
5	0.2644	0.2561	0.2683	0.2574
6	0.2751	0.2730	0.2758	0.2713
7	0.2830	0.2815	0.2816	0.2786
8	0.2893	0.2842	0.2884	0.2828
9	0.2900	0.2865	0.2916	0.2863

Note: "MINDIS" is a maximum likelyhood Algorithim for non-linear estimation. Estimation period was December 1985 to October 1988 and the forecasts were generated for November 1988 to October 1990.

V. Time Series, and Econometric and Judgmental Forecasts

This section discusses the relative accuracy of futures prices as compared to forecasts obtained from time-series and structural econometric models, as well as forecasts obtained using judgmental methods. A number of time series models were developed specifically for this study and forecasts using these models are discussed below; the other two types of forecasts were obtained from existing sources. Despite a plethora of forecasts for the crude oil price for the medium and long term (five to twenty years), the absence of systematic short-term forecasts is quite remarkable. The forecasts examined appear to be the best available. Both in the case of econometric models (given the inevitable "add-on" factors or adjustments), and judgmental forecasts, it is very difficult to ensure that, during the process of forecasting, information provided in the futures prices was not used. An effort was, however, made to select forecasts based on the "structural" factors and which were less likely to be influenced by the information embodied in the futures prices. However, in the case of both judgmental and econometric forecasts, the frequency at which forecasts have been made has often been limited. This, and a number of other issues, are examined in detail below.

1. Time-series models

Consider first a comparison of the accuracy of futures' price forecasts with the forecasts generated from time-series models. A major advantage of this technique is, of course, that the data requirements are very limited--only the data for the actual variable being modelled is needed, which in this case is simply the spot price. The main problem which arises is, however, to ensure that the forecasts obtained from this method correspond precisely to the forecasts implicit in the futures prices. This necessitates estimating models in such a way so that forecasts can be obtained for the day of the maturity of the contract. This in turn necessitates estimation of the models using daily spot price data. Apart from the considerations relating to the time period, the key issue is the specification of the model. In order to provide a rigorous test of futures' price accuracy, it is obviously critical that the time-series model should have the best specification possible. For this purpose, the estimation was initially undertaken using the most general form of the time series generating process--that is, the autoregressive integrated moving average process (ARIMA). The process is defined by the equation:

$$\phi(\beta)(1-\beta)^d q_t = \theta_0+\theta(\beta)a_t \qquad (3)$$

where $\phi(\beta)$ and $\theta(\beta)$ are operators in β of degree p and q, respectively, and the roots of $\phi(\beta)=0$ and $\phi(\beta)=0$ lie outside the unit circle. a_t is white noise process [$E(a_t)=0$ and var $(a_t)=\sigma^2$]. The process is thus of order

(p, d, q). This model is very general, subsuming autoregressive, moving average, and mixed auto-regressive-moving average models, and the integrated forms of all three. 1/

The preliminary analysis strongly suggested, however, that over the sample period, a more specific autoregressive moving average process (ARMA), that is, d = 0, was appropriate implying stationarity in the spot price series.) In such a case the roots of the polynomials lie within the unit circle. A number of models were estimated using this more restrictive specification. (The reasons for preferring the ARMA model was that it achieved as good a fit as an AR model but using fewer parameters.) The best fit was given by a first order autoregressive process and a moving average process of order two. 2/ That is, the equation is of the form

$$\tilde{P}_t = \phi_1 \tilde{P}_{t-1} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \qquad (4)$$

or more generally

$$\phi(B)\tilde{P}_t = \theta(B)a_t \qquad (5)$$

where $\tilde{P}_t$ is now the deviation of P_t from its mean.

The results of forecasts obtained using this model, and the comparable forecasts using futures prices, are shown in Table 6 using the absolute mean error and the root mean square errors for forecast horizons from one to nine months. As noted earlier, the increase in forecast error of the model as the length of forecast horizon increases from one to six months is quite marked; between six and nine months however, there is either no increase or even a marginal improvement. Compared to the futures prices, however, regardless of the forecast horizon, the ARMA model does less well, although it is only for the shortest forecast horizon that the difference is marked.

2. Judgmental forecasts and econometric models

Next consider the forecasts obtained using judgmental methods or econometric models. 3/ The objective is again to compare the accuracy of forecasts using these alternative techniques with that of forecasts obtained from futures prices. Clearly the comparison can be important given that the

---

1/ See, for instance, Box and Jenkins (1976) pp. 85-91, and Granger and Newbold (1986) pp. 25-28.

2/ See, for instance, Box and Jenbuis (1976) pp. 85-91 and Granger and Newbold (1986) pp. 25-28)

3/ In the tests below, given the 'adjustments' applied to forecasts obtained from econometric models, there was no attempt made to separate the econometric and the judgemental forecasts. c.f. McNees (1990).

Table 6. Comparison of ARMA Model and Futures Prices Accuracy

Forecast Horizon (Months)	ARMA Model		Futures Prices	
	MAE	RMSE	MAE	RMSE
1	0.065	0.102	0.053	0.068
2	0.101	0.126	0.082	0.110
3	0.120	0.146	0.108	0.141
4	0.142	0.180	0.131	0.167
6	0.180	0.207	0.159	0.197
9	0.174	0.204	0.161	0.192

Notes: The ARMA model was estimated using daily spot price data, starting with 2nd January 1986 to end-December 1986, to get the first set of forecasts for 1 to 9 months ahead. The model was then reestimated for 1st February 1986 to end-January 1987 for the next set of forecasts, and so on.

futures prices are readily available while the alternative techniques can require considerable investments in model building, following developments in the oil industry, etc.. In theory, such a comparison should not pose any problems: given the ubiquitous nature of oil price forecasting models and forecasts, one would simply take these forecasts and compare them with the futures prices. More formally, denote by

$P^j_{t,T}$ the price forecast obtained at time t for T, using the

forecasting technique j; and

$P^f_{t,T}$ the price of the futures contract at time t, for maturity at T;

and S_T the spot price at T.

Then the comparison would simply be between

$$(P^j_{t,T} - S_T) \text{ and } (P^f_{t,T} - S_T) \tag{6}$$

using any one of the forecasting accuracy criteria.

In practice, however, unlike the time-series model estimated above, there are a number of important constraints which arise in making this comparison.

The first of these constraints arises because in the alternative forecasts T refers not to any specific day, as in the futures forecast, but to a specific period such as a year, or at best a quarter. (Unfortunately forecasts on a quarterly basis are available only for a very short period.) For the above comparison to be undertaken, therefore, it is necessary to transform $P^f_{t,T}$ so that it is comparable to $P^j_{t,T}$. One, and perhaps the only procedure which could be used for this is to obtain an average of prices for contacts maturing in the year for which the forecast is made. Thus one would have a sequence of maturities, T-1, T-2, . . . T-11--where T now refers to the end of the calendar year for which the alternative forecast is made. Even where this can be done, given that futures contracts were only available from 1983 onwards, it would mean that a maximum of only eight observations could be available for the analysis.

However, a second constraint then arises from the fact that the forecasts are often made several months prior to the beginning of the calendar year to which they apply. This means that the forecasting horizon T-t is usually in excess of fifteen months or so. Given that until September 1990, the most distant contracts were only for a period up to fourteen months, it would not be possible to have the contracts which mature throughout the year. Indeed for the period before 1987, when contracts were

available for even shorter periods, it may be possible to have only the first few months of the year for which the forecast is made. 1/

Often, it is also quite difficult to ascertain precisely when the alternative forecasts were made; that is, while for futures, the time t is transparent, it can be highly uncertain for the alternative forecasts. Indeed for the latter, the time at which the forecast is made can often only be guessed at, with a considerable margin of error.

Finally, it is important that the alternative forecasts, especially the judgmental ones are not influenced by the path of futures prices themselves. In practice, of course, it is impossible to ensure this; all that can be done is to ascertain the extent to which a formal model, or formal analysis, is utilized in making the forecasts.

In view of the above considerations, forecasts were obtained from three main sources: annual forecasts published by the Petroleum Economist, the United States Energy Administration, and the World Bank. 2/ Analysis was undertaken, separately, using each of the three sets of forecasts, for the period 1984 to 1991. In general, the forecasting errors using futures prices were smaller, although not markedly so, compared to the forecasts from the alternative techniques. An examination of the individual forecasts also showed that on an annual basis, for the majority of years, the futures prices were more accurate. However, in view of the earlier considerations, these results should be treated with caution. But they do suggest that where futures prices are available, they can provide forecasts which are no worse, and are often better, compared to those obtained from judgmental or econometric methods.

3. Combined forecasts

A final exercise was undertaken to examine whether the forecasts obtained from using futures prices can be improved on, in terms of the variance of the forecast errors, by combining them with forecasts obtained from alternative sources. There is a considerable theoretical literature, and empirical evidence, to show that in many practical situations combined forecasts can outperform individual forecasts in terms of error variance. 3/ Following Granger and Newbold (1986) this reduction in error variance can be seen readily: Let f_n and t_n be the futures and time series

---

1/ These two factors are particularly relevant when considering comparison made, for instance, by Choe (1990).

2/ The forecasts by the USEA have often been in terms of constant dollars--they were converted into nominal dollars using the expected inflation rate.

3/ For a succinct summary of this literature, see, for instance, Granger and Newbold (1986) pp. 266-276.

forecasts of the spot price S_n at time n, and e_n^f and e_n^t be the two forecast errors, respectively. That is,

$$e_n^f = S_n - f_n$$

$$\text{and} \quad e_n^t = S_n - t_n. \qquad (7)$$

The combined forecast is $C_n = Kf_n + (1-K)t_n$, where K and (1-K) denote the weights, and the combined forecast error e_n^c is

$$e_n^c = S_n - C_n = Ke_n^f + (1-K)e_n^t \qquad (8)$$

and the error variance is

$$\sigma_c^2 = K^2 \sigma_f^2 + (1-K)^2 \sigma_t^2 + 2K(1-K)\rho\sigma_f\sigma_t \qquad (9)$$

where ρ is the correlation between two forecast errors.

It can be seen readily that $\sigma_c^2 < \min(\sigma_f^2, \sigma_t^2)$ unless either ρ is exactly equal to σ_f/σ_t or to σ_t/σ_f. If either equality holds, then the variance of the combined forecast is equal to the smaller of the two error variances.

The main issue here is the choice of weights to be applied to the different forecasts. Using the forecasts from the time series model, the following weighting scheme was applied: 1/

$$K_n = \frac{\sum_{t=n-v}^{n-1} e_t(2)^2}{\sum_{t=n-v}^{n-1} \left(e_t(1)^2 + e_t(2)^2\right)} \qquad (10)$$

where $e_n^j = X_n - f_n(j)$, $j = 1,2$

$f_n(j)$ are the two forecasts of X_n from futures and time series models, respectively, and e_n^j are the two forecast errors; n denotes the number of forecasts and Kn is the estimate of the weight to be applied at time n; V was varied between 1 and 3 to obtain optimal weights which were constrained to lie between zero and unity. The results of the exercise showed only a marginal improvement in terms of reduction in error variance

---

1/ For a rationale of this type of weighting scheme, see Granger and Newbold (*op. cit.*), p. 269.

for forecasts horizons ranging from six to nine months. 1/ These results can be interpreted in the context of the earlier findings of the close correspondence between the two sets of forecasts (futures and time series), since, as has been emphasized, the combination is most likely to result in significant improvement only when individual forecasts are very dissimilar in nature.

VI. Summary and Conclusions

This paper has analyzed the forecasting accuracy of crude oil futures prices by comparing these forecasts with those obtained from a variety of other techniques. It emphasized the increasing depth and breadth of the oil futures markets and the role that these factors were likely to play in price discovery. The empirical analysis focused on the behavior of the futures prices from the inception of the market to end-1990, thus using the most extensive sample data ever utilized to examine these issues. The following key results were obtained:

(1) There did not appear to be any systematic bias in crude oil futures prices. This was shown by an analysis of the mean excess returns which could be obtained from holding of futures contracts.

(2) A comparison of end-of month forecasts obtained from using futures prices with forecasts from a random walk model showed that the former provided more accurate forecasts for all forecast horizons. However, as the length of the forecasting horizon increased, the accuracy of forecasts, whether using futures or the random walk model, diminished markedly.

(3) An analysis of intra-month futures prices showed some marginal improvement in forecasting accuracy compared to end-of-the month prices.

(4) When weights are determined endogenously, it appears that the weighting scheme is related to the length of the forecast horizon; the weights decline gradually with lag length when the horizon is distant, but for the nearby forecasts the decline is very steep.

(5) A number of time series models were estimated and the forecasts from them compared with the futures prices. In general the model-based forecasts have larger errors compared to the forecasts using futures prices. A similar result was obtained when comparing judgmental and econometric forecasts.

1/ Thus, for instance, for the six-month ahead forecast, the error variance using futures and time series models were 0.03577 and 0.04267 respectively (using V=1). The error variance of the combined forecast was 0.03447.

(6) Combining forecasts from the time series models and the futures' prices yielded only a marginal improvement in term of variance of forecast errors.

The above results clearly suggest that crude oil prices provide forecasts which are, on average, superior to those obtained from alternative techniques for short-term horizons. For more distant horizons, their accuracy does diminish markedly; however, even for these distant horizons the futures forecasts are no worse, and are often better, than those obtained from alternative techniques.

Appendix Table 1. Crude Oil: Size of Market
(New York Mercantile Exchange)

As of End of March	Open Interest				Volume
	Total	Nearest month	6-Month contract	9-Month contract	
1983 (end July)	5,426	1,644	149	--	1,163
1984	28,717	1,050	1,050	300	3,924
1985	44,165	20,927	621	244	7,737
1986	73,842	29,260	2,169	3,318	19,520
1987	150,829	60,680	4,213	2,442	38,015
1988	198,457	75,546	9,394	3,025	28,749
1989	235,721	87,196	11,830	7,872	139,641
1990 (end May)	277,750	75,579	15,070	5,782	109,917

Source: New York Mercantile Exchange.

Note: Contracts are for 1,000 barrels.

Appendix Table 2. Comparison of Alternative Weighting Schemes for Futures Prices: Mean Absolute Errors

	Forecasting Method				
	Futures prices				
Forecast Horizon (Months)	Last day	Last five days	Monthly average	Exponentially declining weights	Random Walk
1	0.0870	0.0892	0.0981	0.0874	0.0921
2	0.1436	0.1422	0.1488	0.1426	0.1438
3	0.1722	0.1701	0.1702	0.1713	0.1798
4	0.1875	0.1902	0.1933	0.1878	0.1925
5	0.2135	0.2135	0.2145	0.2129	0.2162
6	0.2260	0.2245	0.2267	0.2250	0.2355
7	0.2438	0.2435	0.2438	0.2432	0.2537
8	0.2557	0.2557	0.2560	0.2552	0.2677
9	0.2676	0.2656	0.2688	0.2673	0.2836

Note: Forecasts are for the period December 1985 to October 1990.

Rules for Pricing of Crude Oil Futures Contracts

This appendix outlines the procedures for determining the "settlement price" (SP) for crude oil futures contracts on the New York Mercantile Exchange. The SP is a daily price at which the clearing house clears all trades and settles all accounts between clearing members for each contract month. Since the SP is used to determine both the margin calls and invoice prices for deliveries, there are some very precise rules for its determination. 1/ By the same token, it is the best guide to the market's views as to the future course of prices.

There are essentially two sets of rules, contingent on the volume of trade, for determining the SP. One set of rules applies if at the opening of business on any trading day, a given delivery month has more than 10 percent of the total open interest for all delivery months of the futures contracts. 2/ The second set of rules applies if the volume criterion is not met. These two sets of rules are considered below:

A. When the volume criterion is met:

(i) The SP is the weighted average of the transactions prices during the closing range; this range is defined as the last five minutes of trading before the end of the trading session. The weights are given by the number of contracts traded. For instance, suppose for January 1992 delivery, during the trading range n_1 contracts are traded at price p_1 and n_2 contracts for p_2. The settlement price would then be equal to $(p_1n_1 + p_2n_2)/N$, where $N = n_1 + n_2$.

The reason for having this procedure is that in the so-called "open cry" system of trading in the futures markets, at any given time there would be a range of prices at which transactions would be occurring. To call any one of those prices the "settlement" price would thus be quite arbitrary. The procedure adopted ensures that a set of representative prices is taken; by taking only the last few minutes of trading the objective is to have the prices reflect the latest information available to the market.

1/ These rules are set out in the New York Mercantile Exchange's Rule Guide. The rules for energy contracts (for crude oil, gasoline as well as fuel oil) are given by Rule 6.52. The rules are set by the Exchange and approved by the Commodity Futures Trading Commission.

2/ Open interest is defined as the total number of futures contracts, long or short, that have been entered into and not yet liquidated by an offsetting transaction or fulfilled by delivery. The term is interchangeable with "open contracts" and "open commitments".

(ii) If there are no transactions in the closing range, the SP is the last trade price, unless a bid higher or offer lower than the last trade price is made in the closing range. Such higher bid or lower offer is then called the SP.

B. When the volume criterion is not met:

(i) For these delivery months, the SP is the price relationship between any given delivery month and the current delivery month. The price relationship itself is based on the last "spread transaction" executed in the closing range between such months. Spread transaction is a trade involving the simultaneous purchase of one futures contract against the sale of another futures contract. 1/ For instance, on February 1, 1991, a trader may sell March 1991 contract, the current delivery month, and buy March 1992 contract. The difference in the prices of these two contracts would thus determine the "price relationship" and the SP for March 1992.

(ii) If there is no such spread transaction in the closing range, the relationship would be established by the last such spread transaction executed that day unless a bid higher or offer lower than the last transaction is made in the closing range, in which case the last bid or offer for such spread is the SP.

(iii) If there are no spread transactions and no bids or offers made during any particular trading day, the spread differential for that day is taken to be the spread differential of the settlement prices for the preceding business day.

In addition to the above two sets of rules, there is a provision in the Rules that allows the Settlement Price Committee to establish the SP under specific circumstances. 2/ There are essentially two such circumstances: (a) if the SP, determined according to either set of rules, is inconsistent with transactions that occurred during the closing range in other delivery months; or (b) if the SP is inconsistent with market information known to the Committee. In either of these two circumstances the Committee may establish the SP at a level consistent with other transactions or market information. In such an event the Committee is required to prepare a written record of the basis for any SP so established.

1/ There are a number of different types of spread transactions: the intra-market spread--consisting of buying one month and selling another month in the same commodity; the intercommodity spread--consisting of a long position in one commodity and a short position in a related commodity; and the intramarket spread--consisting of buying a commodity at one exchange and selling the same commodity at another exchange. For the determination of the crude oil SP, it is the first of these spread transactions which is relevant.

2/ This committee consists of three members including a floor trader, a floor broker, and an oil market expert.

It appears that after the outbreak of hostilities in the Middle East, the Committee had to intervene a number of times to set the SP, especially for some of the distant months for which trade volume was very limited. It should be noted that the Committee may determine the SP for one month with the SP for the following month determined by rules (A) or (B) and the Committee again determining the SP for the month following that. Given that the decisions of the Committee have serious financial implications for traders and other users of the futures market, the settlement prices so determined invariably reflect a consensus view, not only of the Committee members but also of the major traders.

The Relationship Between the "West Texas" and Other Crude Oil Prices 1/

This Appendix analyses the relationship between the spot price of "West Texas Intermediate" oil and the "average" price of crude oil. The latter, as used by the International Monetary Fund, is an arithmetic average of three other spot crude prices including "Dubai", "UK Brent", and "Alaskan North Slope" prices. The movement in the differential between the average and the West Texas price, in addition to being of some interest in its own right, also has a direct bearing on the projection of the average price. The Appendix first discusses some a priori considerations relating to the behavior of the differential and then presents some statistical evidence based on daily data on the various crude prices from November 1988 to November 1990, in all over 500 observations. The results indicate that, in general, there is a fairly clear negative relationship between the differential and the level of "West Texas" spot price.

1. Magnitude of the differential

This section examines some of the factors likely to determine the magnitude of the differential between the West Texas crude oil price (WT) and the "average" price (AP) based on the three other crudes and changes in it overtime. These factors include quality differences between the different types of crudes, the degree of substitution between them, the extent to which substitution possibilities change with changes in the price and transportation costs. Each of these factors is discussed in turn below:

1.1 Quality differences: There are several dimensions along which the four crudes differ, with the two most important being the sulphur content of oil and its weight. 2/ In terms of refinery operations, the lighter the oil and the lower its sulphur content, the easier it is to refine it to produce a wide variety of end products. Conversely, the heavier oil would be more expensive to refine and the range of end products available from it would be relatively limited. This difference means that, other things given, the lighter oil will be at a premium relative to the heavier oil.

Of the four crudes, WT is the lightest and has the lowest sulphur content, followed by U.K. Brent; the third in terms of this ranking is Dubai with Alaskan being the heaviest. For any given level of demand, it might be

---

1/ The analysis focuses on the relationship between the "average" price of crude oil as used in the forecasts of the International Monetary Fund, and the West Texas price.

2/ The combination of these two characteristics determines the "sweetness" of the oil--the lighter the oil and the lower the sulphur content, the "sweeter" it is.

expected that the relative f.o.b. prices of the four crudes would reflect their relative quality. This would suggest that the WT price would be higher than the AP. 1/ There may be, of course, circumstances relating to the capacity of specific refineries, or a sudden upsurge in the demand for a particular end product, which lead to an increase in the demand for one of the heavier crudes, increasing its price relative to the others. Any such factor is, however, likely to be transitory with the normal ranking of the prices reestablishing itself within a short period of time.

1.2 Substitution possibilities: The quality differences noted above suggest that in general the differential will be positive (that is WT price will exceed AP). This does not mean, however, that the differential will be constant over time regardless of the level of prices. The extent to which the differential changes with respect to the level of crude prices depends on the extent to which substitution possibilities themselves change. Suppose, for instance, that at a relatively low price substitution between the different crudes is small. Initially, in such a situation an exogenous increase in demand for WT would increase its price, without affecting the prices of the other crudes, leading to an increase in the differential. As WT price increases, however, the substitution possibilities may change: that is, it may become more profitable to buy the cheaper crudes and make adjustments in the refinery operations than to continue to buy a specific crude. In such a situation, the differential in proportionate terms (that is, relative to WT price) would narrow. Indeed if the switch to the cheaper grades is large enough, say due to supply constraints on WT, the absolute differential may also decline.

The above argument suggests that in a period of scarcities with prices relatively high, consumers are less likely to be concerned about quality or grade differences than in a period of relative abundance and relatively low prices. 2/ It might also be argued, however, that it is not only the level of prices but also their rate of change which is important. If prices are rising very fast, consumers would want to ensure their supplies quickly in anticipation of continuing increase. In such a situation also, they may pay less attention to the precise grade, leading to a negative relationship between the differential and the rate of change of prices.

1.3 Transportation costs: As the names of the crudes suggest, each of them originates in a very different geographical location. The price quotations for all four are f.o.b.(although there is a marked difference in

---

1/ Of course, to the extent that one-third of the weight in the average price is that of Brent, which is similar to WT, the differential between WT and the average would be somewhat less than if, say, an average of only Dubai and Alaska was used.

2/ This is a phenomenon found in a very wide variety of primary commodity markets.

the distance between the oilwells and the f.o.b. points). 1/ There will thus be a difference in the f.o.b. prices depending solely on how far the f.o.b. point was from the location of the major consumer markets. For instance, the farther the supply source from the market, other things given, the lower is likely to be the f.o.b. price. On the basis of this argument, it might be expected that since WT is closest to a major consumer market, there would be a positive differential between it and the average price.

To the extent that transportation costs reflect changes in the prices of crudes, some association between WT and the differential might also be expected. For instance, if there is an increase in the demand for all crudes, then an equi-proportionate increase in all crude prices, and in transportation costs, will lead to an increase in the differential in absolute terms but not in proportionate terms. If transportation costs do not increase proportionately, however, the c.i.f. price of, say, the Dubai crude will become less than the WT c.i.f. price. This will lead to a relative increase in the demand for Dubai crude leading to an increase in its f.o.b. price. Thus, the transportation cost factor means that even if all crudes were perfect substitutes, an increase in demand and prices will lead to a narrowing of the differential in both absolute and proportionate terms.

2. Empirical evidence

The statistical evidence on the relationship between the differential and the average price is based on daily observations on the various prices from November 1988 to December 1990. The reason for using daily observations, rather than at a lower frequency, is that the crude spot oil market (as well as the associated futures market) has been highly volatile over the last several years, so that the differential itself is likely to have varied considerably even from day to day.

It is also worth noting that prior to mid-1989, the average price for the IMF's WEO forecasts was based on a weighted average of different crudes. Since then a simple unweighted average has been used. For the purpose of the analysis below, a consistent, unweighted, series for the average price was computed, based on the three crudes discussed earlier.

Appendix Table 3 provides some basic statistics on the mean prices and their variances for each of the crudes, the average price, and the differential between the average and the WT price. In order to isolate any seasonal factors, the values are presented on a quarterly basis as well as for the period as a whole. There are three features worth noting in this

---

1/ West Texas is f.o.b. Midland Texas; Dubai is f.o.b. Dubai; U.K. Brent is U.K. Brent ports and Alaskan North Slope is f.o.b. Gulf of Mexico ports. (See, for instance, Petroleum Market Intelligence (1990).) This suggests that both Alaskan and UK Brent f.o.b. prices would contain elements reflecting transportation costs.

table: first, the remarkable volatility in all crude prices even before the events in the Middle East which affected the prices in the second half of 1990. Prices increased by nearly a third between the fourth quarter of 1988 and second quarter of 1989, fell by nearly 10 percent in the following year, and nearly doubled following the Middle East crisis. Secondly, there does not appear to be any statistically significant difference in the volatility of different crude prices (according to the standard F-test) or any marked seasonal component. Thirdly, although the turning points in the prices are virtually identical, the magnitude of the change in prices is different. This is reflected in a marked variation in both the absolute differential, as well as in the differential as a percentage of the average price. Thus the differential varied between nearly 19 percent in quarter 2, 1990, to barely 4 percent in the last quarter.

Next consider the relationship between the differential and the average price. An analysis of the behavior of these variables using daily data suggested a fairly clear negative relationship between the percentage differential and the average price. To obtain a quantitative indication of this relationship, a first order correlation matrix was computed both for this set of variables as well as the three crudes constituting the average price and the WT price. The results of this exercise are provided in Appendix Table 4. The first half of the matrix shows that the daily prices of different crudes are virtually perfectly correlated. This does not, of course, imply that prices change in an equi-proportionate manner. This can be seen readily in the second half of the correlation matrix which shows a statistically significant correlation between the differential in absolute and percentage terms (DI and DIP respectively) and WT and AP. It is also worth noting that there is somewhat of a closer relationship between the level of prices and the differential in percentage terms than with the differential in absolute terms, supporting the discussion in Section 2. The last section of the matrix considers each of the three crudes constituting AP separately. The divergence of each of these three prices from AP is computed (this is given by the variables DALAP, DDUAP and DUKBAP, respectively); these variables are then correlated with the crude prices as well as with the differentials. The results further highlight the marked differences in the three crude prices. For instance, when WT increases, the divergence between Dubai and AP also increases; this is the reverse of the situation with regard to the other two crude prices.

In order to obtain an indication of the quantitative relationship between the differential and WT, a regression analysis was undertaken with DI or DIP as the dependent variables and WT (or AP) as the independent variables. The standard OLS estimation of this regression would, however, be inappropriate. This is because since DI (and DIP) are defined in terms of AP and WT, the error term in the regression would be correlated with the explanatory variables, yielding biased and inconsistent estimates of the parameters. Consistent estimates can be obtained by using the Instrumental Variable (IV) technique, as is done below. The main instruments used were the lagged values of the independent variables themselves. Since

preliminary analysis had shown a high degree of autocorrelation in the error term, this was also taken into account in the regression analysis.

The results of the estimation are provided in Appendix Table 5. All results indicate a highly significant negative relationship between the differential and WT as well as AP. The results for the absolute differential suggest, for instance, that a dollar per barrel increase in WT price would be accompanied, on average, by a ten cent fall in the differential. Similarly the results for the proportionate differential suggest that a 10 percent increase in the price would be accompanied, on average, by a two percentage point fall in the differential.

Two additional tests were undertaken to further explore the statistical relationship between WT price and the differential. One set examined whether there was any threshold effect; the results indicated that for WT price below $20 a barrel, the differential declined at a slightly slower rate compared to a price above $20 but this result was not statistically significant. Secondly, the relationship between the differential and the rate of change in WT price was examined. Here the results indicated that with the level of WT price taken into account, there was a weak negative relationship between the differential and the rate of change of the WT price.

The above results suggest that when using WT futures price to forecast the average spot price, it would be appropriate to adjust the differential by taking into account the expected WT price. The results provide an indication of the magnitude of the adjustment which may be necessary. It should be emphasized, however, that the results provide an indication of the average adjustment; given the considerations noted above, exceptional market developments concerning any of the four crudes should also be taken into account in computing the precise value of the adjustment.

Appendix Table 3. Crude Oil Prices and Differentials, 1988-90 1/

(In U.S. dollars per barrel)

	1988	1989				1990				Average
Crude Oil	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	1988-90
Alaska	13.14	16.55	18.14	16.91	18.40	19.48	14.88	24.81	30.85	19.14
	(1.49)	(1.13)	(1.28)	(0.61)	(0.86)	(1.01)	(1.19)	(7.03)	(3.21)	(5.45)
Dubai	11.49	14.60	16.62	15.32	17.08	17.30	13.91	25.74	33.13	18.18
	(0.97)	(0.98)	(1.03)	(0.56)	(0.45)	(1.13)	(0.79)	(8.74)	(4.83)	(6.80)
U.K. Brent	14.14	17.24	18.53	17.48	19.33	19.79	16.39	26.08	34.76	20.22
	(1.22)	(1.06)	(1.15)	(0.46)	(0.72)	(1.23)	(0.82)	(7.46)	(3.35)	(6.12)
Average	12.92	16.13	17.76	16.57	18.27	18.86	15.06	25.54	32.91	19.18
	(1.21)	(0.99)	(1.04)	(0.51)	(0.61)	(1.09)	(0.87)	(7.73)	(3.66)	(6.10)
West Texas	15.20	18.21	20.32	19.20	20.31	21.71	17.88	26.32	34.19	21.35
	(1.27)	(1.09)	(0.91)	(0.78)	(0.69)	(1.14)	(1.15)	(6.58)	(3.28)	(5.49)
Differential	2.28	2.08	2.55	2.63	2.04	2.86	2.82	0.78	1.28	2.17
	(0.29)	(0.42)	(0.75)	(0.39)	(0.37)	(0.43)	(0.45)	(1.33)	(1.04)	(0.97)
Differential	17.78	12.95	14.54	15.86	11.18	15.20	18.73	4.97	4.1	12.88
(percentage)	(2.72)	(2.65)	(4.82)	(2.12)	(2.10)	(2.36)	(2.77)	(7.06)	(3.34)	(6.01)

1/ Based on daily observations; data for 1988 Q4 are for November 1st to December 30th 1988, and 1990 Q4 are for October 1st to November 30th 1990. All prices are in U.S. dollars per barrel except for differential, which is both in dollar/barrel and percentage change. Numbers in brackets are the standard deviations.

Appendix Table 4. Crude Oil Prices and Differentials: Correlation Matrix

	AL	DU	UKB	AP	WT	DI	DIP	DALAP	DDUAP	DUKBAP
AL	1.0									
DU	0.98	1.0								
UKB	0.98	0.99	1.0							
AP	0.99	1.0	1.0	1.0						
WT	0.99	0.98	0.99	0.99	1.0					
DI	-0.64	-0.70	-0.67	-0.67	-0.57	1.0				
DIP	-0.80	-0.82	-0.81	-0.82	-0.74	0.92	1.0			
DALAP	-0.46	-0.60	-0.59	-0.56	-0.59	0.56	0.64	1.0		
DDUAP	0.79	0.87	0.81	0.82	0.80	-0.69	-0.76	-0.73	1.0	
DUKBAP	-0.54	-0.48	-0.40	-0.47	-0.48	0.24	0.37	-0.23	-0.50	1.0

Note: Correlation are based on 517 daily observations for the period October 1988 to November 1990. All prices are in dollars per barrel. AL, DU, UKB and WT denote prices of Alaskan, Dubai, U.K. Brent and West Texas crudes, respectively. AP is the average price of the first three crudes. DI and DIP denote the differential between AP and WT in absolute and percentage terms. The last three variables DALAP, DDUAP and DUKBAP denote the divergence of each of the three crudes from AP.

Appendix Table 5. Instrumental Variable Estimates

Dependent Variable 1/	Explanatory Variables 2/			Instruments	DW	R	n
	Constant	AP	WT				
DI	4.51 (16.61)	-0.12 (-9.08)		AP(-1)	2.32	0.12	517
DI	4.51 (11.83)		-0.11 (-6.36)	WT(-1)	2.34	0.01	517
DIP	0.69 (17.12)	-0.20 (-14.25)		AP(-1)	2.32	0.28	517
DIP	0.76 (12.48)		-0.21 (-10.55)	WT(-1)	2.33	0.08	517

1/ The dependent variables are DI and DIP, absolute and the proportionate differential, respectively. AP and WT denote average price and West Texas price; the instruments are lagged values of these variables. All regressions are corrected for first-order serial correlation by using he maximum likelihood iterative technique. T-ratios are in brackets.

References

Anderson, G.H., Bryan, M.F., and Pibe, C. (1990), "Oil and the Economy", Consensus, Vol. XX, No. 49.

Bopp, A. E. and Sitzer, S. (1987) "Are Petroleum Futures Prices Good Predictors of Cash Value", The Journal of Futures Markets, Vol. 7, No. 6.

Cho, D.W and McDougall, G.S. (1990) "The Supply of Storage in Energy Futures Markets", The Journal of Futures Markets, Vol. 10, No.6, 611-621.

Choe, B (1990) "Rational Expectations and Commodity Price forecasts", PRE, WPS 435, World Bank, June.

Clemen, R. T. (1989) "Combining Forecasts: A Review and Annonated Bibliography", International Journal of Forecasting, 5.

Dominguez, K. M. (1987) "The Volatility and Efficiency of Crude-Oil Futures Contracts", in Oil and Money: Coping with Price Risk Through Financial Markets, Harvard International Energy Studies.

Fama, E.F. and French, K.R. (1988) "Permanent and Temporary Component of Stock Prices", Journal of Political Economy, Vol. 98, pp. 246-274.

Feinstein, S. P. (1989) "Forecasting Stock Market Volatility Using Options on Index Futures", Economic Review, Federal Reserve Bank of Atlanta, May, pp. 12-30.

Hampton, M. (1991) "Cycling Towards Low Prices", Petroleum Economist, April 1991.

Hirschfeld, D. J. (1983) "A Fundamental Overview of the Energy Futures Market", The Journal of Futures markets, Vol. 3, No.1.

Kaminsky, G. and Kumar, M.S. (1990) "Efficiency in Commodity Futures Markets", IMF Staff Papers, September.

Kaminsky, G. and Kumar, M.S. (1990) "Risk Premia in Commodity Futures Markets", IMF Working Paper 90/116, December.

Lo, A.W. and Mackinlay, A. C. (1988) "Stock Market Prices Do Not Follow Random Walks: Evidence From A Simple Specification Test", Review of Financial Studies, Vol. 11, pp. 41-66.

McNees, S. K. (1990) "The Role of Judgement in Macroeconomic Forecasting Accuracy", International Journal of Forecasting, October.

Ma, C. W. (1989) "Forecasting Efficiency of Energy Futures Prices", The Journal of Futures Markets, Vol. 9, No. 7.

Petroleum Intelligence Weekly, June, (1988), Petroleum Database Services.

Porteba, J. M. and Summers, L.H. (1989) "Mean Reversion in Stock Prices: Evidence and Implications", Journal of Financial Economics.

Serletis, A. (1991) "Rational Expectations, Risk and Efficiency in Energy Futures Markets", Energy Economics, April.

Walton, D. (1991) "Backwardation in Commodity markets", The Commodities Analyst, Vol. 1, Goldman Sachs, New York, June.