Risk Premiums and Efficiency
in the Market for Crude Oil Futures*

Richard Deaves and Itzhak Krinsky™

The New York Mercantile Exchange's Crude Oil futures contract is investigated for the existence and nature of risk premiums and informational efficiency. During 1983-90, there is some evidence that short-term premiums were positive and covaried with recent volatility. As for efficiency, we find nothing inconsistent with weak-form efficiency, but some apparent violations of semi-strong efficiency. We argue that, for a number of reasons, such rejections should be interpreted with caution.

INTRODUCTION

The emergence of organized futures markets can be traced to 19th century Chicago and the opening of the Chicago Board of Trade in 1848.¹ Until the last 15 years, virtually all contracts were based on agricultural and metallurgical commodities, but of late futures markets have experienced a sharp expansion in their number as foreign exchange, interest rate, stock index and energy futures have begun trading. This is not surprising, given that the two principal benefits associated with futures markets are not unique to the farm and mining sectors. The first is the price discovery mechanism whereby market

¹The key differences between forward contracts and futures contracts are the latter are highly standardized with respect to quality, quantity and delivery; are traded on organized exchanges; and require a good will deposit (i.e. a margin) which varies (i.e. is marked-to-market) each day depending on price movements.

The Energy Journal Volume 13, Number 2. All Rights Reserved.

*The authors would like to thank the helpful comments of the editors and two anonymous referees. This research was funded by the Social Sciences and Humanities Research Council of Canada.
participants gain some insight, by observing futures prices, into consensus beliefs for the expected future spot price of the underlying commodity or asset. The second is the risk transfer mechanism, whereby those less disposed to bear the risk of price fluctuations can shift it to those less risk averse.

Trading in energy futures has grown rapidly during the past decade. The New York Mercantile Exchange (NYMEX) initiated the Heating Oil No. 2 contract in 1978, but it is the Crude Oil contract based on West Texas Intermediate that has achieved the most success. Trading began in 1983 and by January 1990 trading volume of over two million (1,000 barrel) contracts was witnessed. In this paper, a number of different econometric techniques will be employed in testing risk premium theory and (informational) efficiency for the NYMEX Crude Oil futures market during the highly volatile period of 1983-90. Preliminary studies of efficiency have been done by Bopp and Sitzer (1987), Dominguez (1989), Ma (1989), and Deaves and Krinsky (1991). This paper broadens the range of efficiency tests. In addition, to our knowledge, this paper stands alone in testing for the appropriate formulation of the risk premium inherent in Crude Oil futures prices.

We investigate risk premiums and efficiency in this study, because both concepts are important for those considering using the energy futures market, whether for hedging or speculative purposes, or a mix of the two. In fact, Deaves and Krinsky (1991) show that ex ante hedging efficacy is directly related to the magnitude of the risk premium, and the extent to which futures markets are efficient.

While hedgers transfer risk, speculators assume it. An investor who willingly assumes risk — in this case a speculator purchasing/selling a risky futures contract — does so only if he expects a reward sufficient for the risk borne. This comes in the form of the expectation of rising/falling prices. As explained below, however, only in an efficient market, can the gap between the current futures price and that expected to obtain at the end of the contract be viewed as a risk premium.

This risk premium theory of futures pricing, originating with Keynes (1930) and Cootner (1960), is, however, not without difficulty. The problem

---

2. As will become clear, this is not necessarily the same as saying futures prices are unbiased predictors of future spot prices.

3. The definition of informational efficiency differs from the better known notion of Pareto-production (or allocative) efficiency. Note that it is common to abbreviate informational efficiency to mere "efficiency" or the "efficient markets hypothesis." Also, "rational expectations" is an additional way to characterize this concept.

4. A speculator in futures markets is generally viewed as one who takes a position in futures without any position — whether current or anticipated — in the cash (or spot) market. On the other hand, a hedger will have such a position and uses futures to offset the price risk inherent in the cash market.
is that no universally accepted view exists on the magnitude (or even sign) of the risk premium in the case of commodity futures. Further, there is a now well-established view in the finance literature suggesting that risk premiums may be time-variant.

Inextricably linked to pricing theory is market efficiency. Financial markets are said to be efficient if observed prices conform to their theoretical counterparts. This implies that market participants, who analyze information relevant for determining security prices, are unable to determine any trading strategies that yield returns over and above what is appropriate for the risk borne. Positive returns are acceptable, but such excess returns are not.\(^5\)

The problem is that all tests of market efficiency are unavoidably joint hypothesis tests, with the null encompassing both efficiency and the assumed pricing model. If a test leads to rejection of the null, it may simply be that the pricing model utilized is inappropriate. In other words, the risk premium has been improperly specified. The only way to broach this conundrum is to consider whether such rejections could reasonably be due to inappropriate price modelling. If not, rejecting efficiency is defensible.

Despite the difficulties in testing efficiency, the great attention accorded it by financial economists attests to its importance, but it is more than an arcane academic matter. Efficiency should be an overriding concern for market participants as well. If prices can be shown to be set in markets which are inefficient, profitable (risk-adjusted) opportunities may be available. Also, as stated earlier, hedging efficacy is enhanced if futures markets are efficient. Inefficiency is reflected by a gap between the futures price and the expected spot price that is not necessarily equal to the risk premium. Thus, hedgers may either underpay or overpay for the service of risk transfer.

The paper proceeds by first investigating pricing and risk premiums, and then turning to the issue of efficiency. The organization is as follows. In the following section, the appropriate theoretical background on commodity futures pricing is provided. We begin by discussing an alternative model of commodity futures pricing, the theory of storage dating back to Kaldor (1939), Working (1949), Telser (1958), and Brennan (1958), which is argued to be inappropriate for investigating market efficiency. Next comes a further discussion of risk premium theory as applied to commodity futures. Since commodity futures contracts are risky assets, several models, prevalent in the finance literature, that attempt to explain risk premiums, will be investigated for their applicability to Crude Oil futures. These are the CAPM of Sharpe (1964), among others, and the ARCH-M of Engle, Lilien and Robins (1987). We then

---

5. See Fama (1970) for a full treatment of the efficient markets hypothesis. In the spirit of Grossman and Stiglitz (1980), one should be careful to factor in all costs (transaction and analysis-related).
perform tests on the existence, magnitude and time-variation of risk premiums, and to what extent such aforementioned models are capable of explaining returns. Next, using our findings on pricing as a backdrop, the efficiency of the NYMEX Crude Oil futures market is investigated. Finally, we conclude the paper by noting the main implications of the empirical results.

COMMODITY FUTURES PRICING THEORIES

Theory of storage

According to the theory of storage, the return from purchasing a commodity at time \( t \) and selling it forward (using a futures contract) for delivery at time \( T \), should be equal to the cost of storage (principally interest foregone and warehousing) minus a convenience yield. More precisely,

\[
F_{t,T} - S_t = I_t + W_t - C_t
\]

where \( F_{t,T} \) is the log of the futures price at \( t \) for delivery at \( T \); \( S_t \) is the log of the spot price; \( I_t \) is the interest rate expressed between \( t \) and delivery \( (T) \); \( W_t \) is the (percentage) storage cost between \( t \) and \( T \); and \( C_t \) is the convenience yield over the same span. In other words the percentage basis, or (approximately) the log-difference between the futures price and the contemporaneous spot market (i.e. cash) price, is explained by financing and warehousing costs and the convenience yield.\(^6\) The convenience yield is usually described as the convenience of holding inventories because many commodities are used as inputs in the production process, or as the convenience of having inventories to meet unexpected demand. The storage theory predicts low convenience yield when inventories are plentiful and high convenience yields when stockout is more likely.\(^7\)

Tests of the theory of storage have been done by, among others, Telser (1958), Brennan (1958, 1986), Fama and French (1987) and Williams and Wright (1991).\(^8\) For energy futures, recent tests have been conducted by Cho

---

6. Normally the basis is expressed as the absolute difference between the futures price and the corresponding spot (or, to some, the negative of the latter).

7. Heinkel, Howe and Hughes (1990) provide a formal model which confirms one's intuition that higher inventory levels imply a smaller convenience yield. As well, two additional propositions are derived: convenience yield is increasing in marginal production costs and decreasing in spot price serial correlation.

8. Telser (1958) and Brennan (1986) provide empirical estimates of the convenience yield that are consistent with the theory. Fama and French (1987) provide evidence that marginal convenience yields vary seasonally for most agricultural and animal commodities but not for metals.
and McDougall (1990). They found that the basis tends to increase with total commodity stocks. This is consistent with the theory of storage since high stocks are associated with low convenience yield.

Unfortunately, the theory of storage is not overly useful for the purpose of determining theoretical prices and investigating market efficiency. Inspection of (1) elucidates this contention: \( C_t \) can be inferred from the basis and proxies for \( I_t \) and \( W_t \). There is however no way to know what would constitute an appropriate convenience yield. Therefore one cannot draw inferences on the appropriateness of \( F_{tT} \) at any point in time.

Risk premium theory

While the theory of storage relates current spot and futures prices, risk premium theory relates futures prices to anticipated future spot prices. Risk premium theory states that futures prices at \( t \) for delivery at \( T \), plus potentially a (positive or negative) risk premium, are equal to market expectations at \( t \) of the spot price for \( T \). More formally,

\[
F_{tT} + RP_t = E_{T}S_T
\]

where \( E_{T}S_T \) denotes the time-\( t \) expectation of the spot price for \( T \); and \( RP_t \) is the difference between the latter and the corresponding futures price. It is termed a risk premium since, if positive/negative it is the average reward for a speculator going long/short in (i.e. buying/selling) a futures contract at \( t \) and reversing his position just prior to delivery (\( T \)).

To see this, one should note that the futures price for a commodity with virtually no time till delivery must equal the cash price, or \( S_T = F_{T,T} \) and \( E_{T}S_T = E_TF_{tT} \). This allows us to rewrite (2) as

\[
RP_t = E_TF_{tT} - F_{tT}
\]

The right hand side of (3) is merely the expected return on a long futures position initiated at \( t \) and reversed just prior to delivery at \( T \) (\( E_R \)), or

\[
E[..]R_T = RP_t
\]

Thus the expected return to a long futures position held till the end of the

9. If a trader buys/sells a future contract and subsequently sells/buys an identical contract, the clearinghouse nets out his position to zero.
10. Were this not so, obvious arbitrage opportunities would be available.
contract is equivalent to the ex ante risk premium.

Note that futures returns are simply the sum of such expected returns (or risk premiums) and forecasting errors, as follows:

\[ R_t = F_{t,T} - F_{t,T} \]

\[ = R_P + M_t \]  

(5)

where \( M_t = S_t - E S_t = F_{t,T} - E F_{t,T} \) is the time-\( t \) forecast error of the time-\( T \) futures (or spot) price.

When should one expect this risk premium to be positive? To take the traditional view, if hedges are net short, or the majority of hedges possess (or anticipate possession of) the commodity, as in the case of an oil producer about to pump crude out of the ground, in order to insulate themselves from price fluctuations they sell (i.e. go short in) futures contracts. Long speculators (i.e. those buying futures contracts) are needed to supply price insurance, which implies that long positions should typically be rewarded by futures price increases. This means that risk premiums are positive. This pattern of futures price changes is called "normal backwardation." If, on the other hand, hedges are net long (e.g. oil consumers), it will be necessary for the risk premium to be negative in order to entice short speculators into taking sufficient offsetting positions. As a reward for risk-bearing, futures price declines should be the norm. This is termed "contango."

Not only will the sign and magnitude of the risk premium logically be related to the distribution of hedges, but also it will be a function of the market's determination of the price of risk and the amount of risk inherent in such positions. The simplest possibility to consider is that the risk premium is constant over time. In this case, the mean return could be used as the estimator for the time-invariant premium. Dominguez (1989) estimated mean returns for nearby to six-month out Crude Oil futures and found them to be insignificantly different from zero for a 1983-86 sample. One possible interpretation is that hedging activity was not predominantly on either the long or short side of the market.

The problem is that there is no reason to suppose that the premium will be constant over time. For one thing, the appropriate compensation for risk may depend on factors whose values change over time. After investigating the sign and magnitude of unconditional risk premiums, we turn to models that allow for such variation.

One such factor that might affect the risk premium is the return on a broad market portfolio, as in the capital asset pricing model or CAPM. Underlying this model is the prediction that a risky asset's ex ante return should increase with its correlation with the return on a broad portfolio of all risky assets. Since other forms of risk can be diversified away, they should not be
A time-varying relationship between risk and ex ante return is also a possibility, perhaps because the amount of risk inherent in speculative futures positions is perceived to change over time, as in the ARCH-M of Engle, Lilien and Robins (1987). In this model, volatility that has been high in the recent past leads to higher ex ante returns. We now discuss and estimate these models for Crude Oil futures.

RISK PREMIUM TESTS

NYMEX Crude Oil Futures Returns

In our empirical tests we use monthly observations of NYMEX Crude Oil futures settlement prices, with observation dates coinciding with the final trading date for each nearby (i.e. closest to delivery) contract.\textsuperscript{11} Trading in each contract terminates on the third business day prior to the 25th of the month preceding the delivery month. For example, late in September, the November contract was the nearby contract.\textsuperscript{12} Based on the evidence of Ma (1989) these terminal futures prices are reasonable proxies for spot prices.\textsuperscript{13}

Returns are calculated for one-month to five-month horizons as in (5).\textsuperscript{14} Though contracts trade for all future months for at least 12 months in the future, given that more distant contracts have much lower volume, it is preferable to concentrate on the more liquid contracts.

Unconditional Risk Premiums

If mean futures returns are significantly different from zero, we can infer that unconditional risk premiums are on average nonzero. We begin by calculating mean returns for NYMEX Crude Oil futures contracts and associated

---

\textsuperscript{11} Crude Oil futures prices were obtained from Techtools Inc., Los Altos, California.

\textsuperscript{12} Delivery initiation must be after the first calendar day of the delivery month with completion before the last calendar day of the delivery month. If delivery always takes place coincident with the end of trading, this implies that futures prices should converge to corresponding spot prices on the day of delivery. Since delivery can take place over a period of days, the question of which spot price should be used arises. According to Ma (1989), the evidence indicates that the futures price well predicts the spot price on the last day of trading.

\textsuperscript{13} In the context of Crude Oil, even the spot price is more akin to a one-month forward price. One must reserve space on the pipeline on or before the 25th for delivery in the following month, and delivery can be made over the entire month.

\textsuperscript{14} Of course this definition of futures returns is from the standpoint of a long speculator. The return to a short speculator is simply the negative of this.
t-statistics under standard OLS assumptions. The results are presented in the upper panel of Table 1 for the entire May 1983 to April 1990 sample period. Horizons of different lengths (1, 2, ..., 5 months) are examined. Overlapping observations are used for two- to five-month horizons. This is valid since the standard errors are estimated using the asymptotic covariance matrix suggested by Hansen and Hodrick (1980), which accounts for the moving average nature of the error terms.

For now ignoring the row labelled ARCH(1), we notice that all mean returns, though consistently positive, are insignificantly different from zero. At first glance, consistent with Dominguez (1989), it appears that unconditional risk premiums are unequivocally zero.

Table 1. Mean Crude Oil Futures Returns and ARCH Effect Tests
(Complete Sample: May 1983-April 1990)

<table>
<thead>
<tr>
<th>Forecast Horizon (Months)</th>
<th>Mean Return</th>
<th>T-statistic</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00589</td>
<td>0.49499</td>
<td>79</td>
</tr>
<tr>
<td>2</td>
<td>0.01055</td>
<td>0.62215</td>
<td>79</td>
</tr>
<tr>
<td>3</td>
<td>0.01239</td>
<td>0.62972</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>0.01205</td>
<td>0.53489</td>
<td>79</td>
</tr>
<tr>
<td>5</td>
<td>0.01690</td>
<td>0.43354</td>
<td>79</td>
</tr>
<tr>
<td>1 (ARCH(1))</td>
<td>0.01950</td>
<td>2.04426</td>
<td>78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Horizon (Months)</th>
<th>LM (ARCH(1))</th>
<th>LR (ARCH(2) vs. ARCH(1))</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.99</td>
<td>0.38</td>
<td>84</td>
</tr>
<tr>
<td>2</td>
<td>5.20</td>
<td>0.25</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>0.01</td>
<td>0.19</td>
<td>27</td>
</tr>
</tbody>
</table>

Note: Both the LM and LR test-statistics above are distributed as $\chi^2$ (1). The critical value for 5% is 3.84.
ARCH Effects

The latter calculation of means and t-statistics is based on the assumption of constant variance. It is being increasingly recognized that this may be invalid given the volatility clustering that researchers have documented of late in numerous asset markets.\textsuperscript{15} The idea is simple: large/small absolute price changes are often followed by further large/small absolute changes.\textsuperscript{16} Engle (1982) suggested that such autoregressive conditional heteroscedasticity (ARCH) effects could be modelled as follows:

\[
R_t = \gamma_0 + \sum_{i=1}^{m} \gamma_i x_{t-i} + \epsilon_t \\
\sigma_t^2(\epsilon_t) = \delta_0 + \sum_{i=1}^{q} \delta_i \epsilon_{t-i}^2 + u_t
\]

where the \(x_t\) are a set of variables explaining returns. Consider the simplest case where all \(\gamma_i\) are equal to zero, so that expected returns are constant.

To test for an ARCH effect of order q, the procedure is to regress the OLS squared residuals on q lags of themselves. The Lagrange Multiplier (LM) test-statistic, calculated as nR^2 (where n is the sample size), is asymptotically distributed as \(\chi^2(q)\). The second column of the lower panel of Table 1 provides these LM test-statistics for one-, two- and three-month horizons. Note that it is inappropriate to test for ARCH effects using overlapping observations because they will be present by construction. Therefore we do not examine return horizons beyond three months, since the number of observations would be small. The evidence quite strongly supports ARCH(1) for the one-month horizon, and, less strongly, for the two-month horizon. In addition, the third column provides likelihood ratio tests that, under the maintained hypothesis of ARCH(1), a second order ARCH effect exists as well. Clearly, there is no evidence of this being the case.

Based on these results it seems appropriate to make allowance for a first order ARCH effect for the one-month horizon, and this will be our practice below. We do not do so in the case of the two-month horizon, given that the ARCH effect appears weaker, and that the number of observations will be cut

\textsuperscript{15} See, for example, the review in Bollerslev, Chou, Jayaraman and Kroner (1990).

\textsuperscript{16} Volatility clustering implies \(\delta_i > 0\). Note that it is not obvious that such clustering is what one should expect. Hodrick (1987) suggests a counter example: if a large price change in the past was due to the resolution of uncertainty, then in the future one might expect smaller price changes. This effect notwithstanding, our evidence suggests that with Crude Oil futures volatility clustering dominates.
dramatically.17 Finally, returning to the upper panel of Table 1 the importance of the ARCH effect is illustrated. Note that the mean return is now significantly positive. On average, during 1983-90, a speculator who rolled over a long position in the nearby Crude Oil futures contract achieved a return of close to 2% per month.18 The obvious immediate interpretation is that on average short hedging activity predominated.

Unconditional Risk Premiums: Subperiods

Next we begin considering whether the risk premium may be time-variant. For example, is it possible that the risk premium may have been positive during some periods, but zero or negative during others? To take a somewhat crude approach in examining this possibility, the sample is divided into subsamples according to whether crude oil spot prices exhibited an increasing or decreasing trend. This method of obtaining the subsamples is based on the evidence on investors’ expectations in other futures markets. In the foreign exchange market, for example, it has been observed that investors consistently underpredict the value of an asset when the asset is appreciating, and systematically overpredict it when it is depreciating (Frankel and Froot (1987)).

Refer to Figure 1 for a plot of spot crude oil price movements during the entire sample period. Based on this recent price history, the period May 1983 to April 1990 is divided into four subperiods: May 1983 to November 1985, August 1986 to July 1987, August 1987 to November 1988, and December 1988 to April 1990.19 Table 2 presents the subperiod results for mean returns.20 During the second and the fourth subperiods, mean returns in Crude Oil futures were significantly positive for most of the four investment horizons considered.21 In the case of the third subperiod, returns were negative and, with the exception of one month horizon, significant at the 10% level. Only in the first subperiod are mean returns insignificantly different from zero. These results appear consistent with those of Frankel and Froot (1987).

17. Given the asymptotic nature of the estimation technique, it is preferable to avoid smaller sample sizes.
18. By "roll over" we mean a speculator initiates a long position in the nearby contract and simultaneously reverses it just prior to delivery and initiates a new long position in the second nearby contract (which immediately becomes the nearby contract).
19. December 1985 to July 1986 is omitted due to its small size. Also, during this period, oil prices plummeted.
20. Due to the small subsample sizes ARCH estimation is not attempted.
21. The investment horizon is restricted to a maximum of four months due to the low number of observations available in the subperiods.
Figure 1. Crude Oil Prices, 1983-1990

Table 2. Mean Crude Oil Futures Returns (Subsamples)

<table>
<thead>
<tr>
<th>Horizon (Months)</th>
<th>Sample Period</th>
<th>Mean</th>
<th>T-Statistic</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9/83-11/85</td>
<td>0.0102159</td>
<td>1.1701900</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.0146192</td>
<td>0.7892018</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.0140800</td>
<td>0.5494401</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.0142602</td>
<td>0.4995920</td>
<td>27</td>
</tr>
<tr>
<td>1</td>
<td>8/86-7/87</td>
<td>0.0580876</td>
<td>2.9214090</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.0903758</td>
<td>5.5455910</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.1305088</td>
<td>8.9374700</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.1530825</td>
<td>6.0188800</td>
<td>8</td>
</tr>
<tr>
<td>1</td>
<td>8/87-11/88</td>
<td>-0.0341863</td>
<td>-1.4640800</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>-0.0588102</td>
<td>-1.9942150</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>-0.0823586</td>
<td>-2.1139550</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>-0.0953055</td>
<td>-2.3609890</td>
<td>12</td>
</tr>
<tr>
<td>1</td>
<td>12/88-4/90</td>
<td>0.0212507</td>
<td>0.8096463</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.0547767</td>
<td>1.2939580</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.0860051</td>
<td>2.3885220</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.1199810</td>
<td>3.3030476</td>
<td>13</td>
</tr>
</tbody>
</table>
What is able to account for such patterns? A risk premium that changes from period to period is a possibility, as is market inefficiency. There may be a third possibility. This is related to the fact that expectations which look biased ex post, may actually be rational ex ante. This is a salient possibility that exists whenever prices follow a process that evolves over time. The process followed by crude oil prices may be a good candidate, in part due to unanticipated policy changes by OPEC.\textsuperscript{{22}} For example, during a period of rising prices (with OPEC exercising discipline), futures prices might systematically underpredict spot prices because market participants are cognizant that OPEC’s discipline might break down.

Normal returns and CAPM

It is now commonplace for financial economists to view risk in the context of the overall portfolio. In the context of a portfolio some (but not all) risk, where risk is viewed as the variability in return of an individual security held in isolation, can be eliminated by diversification. The distinction between diversifiable and non-diversifiable risk is highlighted in variants of the Capital Asset Pricing Model (CAPM). CAPM maintains that diversifiable risk is not priced (i.e. rewarded), while non-diversifiable risk, as reflected in a security’s beta is priced.

More formally, the expected return on a security over and above the risk-free rate (e.g. on government Treasury bills) is equal to the product of two elements, as follows:

\[ E[R_t] = RF_t + \beta[E[RM_t] - RF_t] \]

where \( RF_t \) is the return on a risk-free security at \( t \); \( RM_t \) is the return at \( t \) on the market portfolio (which can be thought of as a portfolio made up of all risky assets); and \( \beta \) is the covariance between the return on the market portfolio and the return on the security under consideration scaled by the market variance. The first element in (7) is the pure opportunity cost of invested funds while the second is a reward for risk-bearing. For futures the initial cash outlay is, for all intents and purposes, zero, since T-bills can be posted as margin. This implies the anticipated return is a pure risk premium, and, for futures, (7) simplifies to

22. See Deaves and Krinsky (1992) for a test of efficiency in Crude Oil futures that revolves around the information content of OPEC meetings.
\[ E[R_t] = \beta [E[R_{M_t} - R_F]] \] \hspace{1cm} (8)

Note that (8) is simply one specification of (4).

This formulation of the risk premium for futures markets has been investigated by a number of researchers. A key testable implication of CAPM as applied to commodity futures is that a commodity’s beta and mean return have the same sign, a relationship corroborated by Dusak (1973) (for three agricultural commodities) but then contradicted by Bodie and Rosansky (1980) for a much broader sample (23 commodities not including energy futures). Carter, Rausser and Schmitz (1983) countered that the proxy for the return on the market used in these studies (i.e. an index of common stocks) was inappropriate for commodity futures, since these measures omitted a weighting in aggregate wealth for commodity inventories. Using as their measure of the market an equally weighted average of the S&P 500 and the Dow Jones Commodity Futures Index they discovered betas for commodity futures to be indeed significantly positive. Marcus (1984), and Baxter, Conine and Tamarkin (1985) criticized this weighting scheme, and found betas were once again primarily insignificant when a more defensible 90% weight was assigned to the S&P 500.23

Betas for Crude Oil futures returns versus various measures of market returns are presented in Table 3. Five proxies for the market portfolio, namely the S&P 500 Index, the Dow Jones Commodity Index, and three weighted averages of these two indices are utilized. Note that all the estimated betas have a negative sign which, at least on the surface, indicates negative co-movement with aggregate returns.24 Still, in all cases, the betas are insignificantly different from zero. In the case of the one-month horizon, this is true whether or not the ARCH effect is taken into consideration.

ARCH-M model

The ARCH-in-mean or ARCH-M model of Engle, Lilien and Robins (1987)), has recently become popular for modelling ex ante returns for a number

23. Breeden (1980) and Hazuka (1984) examined, using a consumption-based CAPM approach, futures returns. In this formulation, the relevant covariance is of commodity returns and changes in aggregate consumption. Their success in explaining commodity futures returns was very much mixed.

24. In contrast to Carter, Rausser and Schmitz (1983), we find, for the most part, not an increase but a decrease in betas by introducing the Dow Jones Commodity Futures Index into the market portfolio.
Table 3. Calculation of Betas

<table>
<thead>
<tr>
<th>Horizon (Months)</th>
<th>S&amp;P Beta</th>
<th>Beta with .25 Commodity Index Weight</th>
<th>Beta with .50 Commodity Index Weight</th>
<th>Beta with .75 Commodity Index Weight</th>
<th>Dow Jones Commodity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.07817</td>
<td>-0.13649</td>
<td>-0.23631</td>
<td>-0.30528</td>
<td>-0.23498</td>
</tr>
<tr>
<td></td>
<td>(-0.34565)</td>
<td>(-0.46522)</td>
<td>(-0.62343)</td>
<td>(-0.72106)</td>
<td>(-0.64240)</td>
</tr>
<tr>
<td>1 ARCH(1)</td>
<td>-0.10563</td>
<td>-0.14757</td>
<td>-0.21945</td>
<td>-0.29865</td>
<td>-0.21906</td>
</tr>
<tr>
<td></td>
<td>(-0.50472)</td>
<td>(-0.54704)</td>
<td>(-0.63954)</td>
<td>(-0.85869)</td>
<td>(-0.77883)</td>
</tr>
<tr>
<td>2</td>
<td>-0.19941</td>
<td>-0.29754</td>
<td>-0.42747</td>
<td>-0.43166</td>
<td>-0.24939</td>
</tr>
<tr>
<td></td>
<td>(-0.81033)</td>
<td>(-0.85789)</td>
<td>(-0.88530)</td>
<td>(-0.78678)</td>
<td>(-0.57021)</td>
</tr>
<tr>
<td>3</td>
<td>-0.19941</td>
<td>-0.11032</td>
<td>-0.24385</td>
<td>-0.33921</td>
<td>-0.04510</td>
</tr>
<tr>
<td></td>
<td>(-0.81033)</td>
<td>(-0.17694)</td>
<td>(-0.28013)</td>
<td>(-0.42199)</td>
<td>(-0.10737)</td>
</tr>
<tr>
<td>4</td>
<td>-0.03566</td>
<td>-0.10037</td>
<td>-0.23811</td>
<td>-0.34394</td>
<td>-0.26760</td>
</tr>
<tr>
<td></td>
<td>(-0.07131)</td>
<td>(-0.13570)</td>
<td>(-0.23267)</td>
<td>(-0.37113)</td>
<td>(-0.48927)</td>
</tr>
<tr>
<td>5</td>
<td>-0.10426</td>
<td>-0.19106</td>
<td>-0.33969</td>
<td>-0.41033</td>
<td>-0.28740</td>
</tr>
<tr>
<td></td>
<td>(-0.16854)</td>
<td>(-0.22060)</td>
<td>(-0.30729)</td>
<td>(-0.44697)</td>
<td>(-0.57647)</td>
</tr>
</tbody>
</table>

Note: T-statistics are supplied below the coefficient estimates in brackets.

of financial assets. As its name suggests, this methodology is related to the simple ARCH model. The difference is that this model provides for an increase in ex ante return with recent volatility. The following model for ex ante premiums or returns is hypothesized:

\[
R_t = \gamma_0 + \gamma_1 \sigma_t^2(e_p) + \epsilon_t
\]

\[
\sigma_t^2(e_p) = \delta_0 + \sum_{i=1}^{g} \delta_i \tau_{t-i} + u_t
\]  \hspace{1cm} (9)

The intuition of the ARCH-M is that speculators require a reward for return unpredictability (i.e. volatility), so the ex ante premium is assumed to be linearly related to some function of the variance. In this model, an increase in the conditional variance will be associated with an increase or decrease in the conditional mean of \( R_t \). The direction of the relationship logically depends on

25. For examples, see Bollerslev, Chou, Jayaraman and Kroner (1990).
the distribution of hedgers. If hedgers are net short, which means that speculators are net long, \( \gamma_1 \) will be positive; whereas if hedgers are net long, which means that speculators are net short, \( \gamma_1 \) will be negative. A variant of this model has recently been applied to futures by Melvin and Sultan (1990). They found that ex ante premiums tend to increase (i.e. normal backwardation is more prevalent or contango less so) the greater is the volatility in gold prices.

In considering the potential success of this model in explaining ex ante returns, we use the one-month horizon only. Table 4 provides a summary of the key results. All models estimated fit into the following framework:

\[
R_t = \gamma_0 + \gamma_1 \sigma^2(e_t) + \gamma_2 R_{Mt} + \epsilon_t
\]

\[
\sigma^2_t(e_t) = \delta_0 + \delta_1 \epsilon_{t-1}^2 + \nu_t
\]

(10)

Note that (10) is a special case of (6). The two variables potentially explaining returns are some function of residual variance (i.e. the ARCH-M effect) and market returns \( (R_{Mt}) \). Consistent with previous findings, the ARCH order considered is \( q = 1 \). The first two rows—the estimation of mean returns with and without ARCH(1)—are provided for the purpose of comparison.

Rows 3 and 4 refer to two ARCH-M estimations. The only variable explaining ex ante returns is some function of conditional variance. In the first case the function of the error variance is simply the standard deviation, and in the second case it is the variance itself. For both specifications likelihood ratio tests indicate significant modelling improvement by allowing the previous month's volatility to explain this month's ex ante return.

Rows 5 and 6 repeat the estimation of rows 3 and 4, but, in addition, the market return—half in the S&P 500 and half in the Dow Jones Commodity Index—is included as another potential explanatory variable of normal Crude Oil futures returns. The intention is to investigate whether including this additional variable in any way degrades the significance of the ARCH-in-mean effect. Clearly it does not.

The table also provides in the last column an \( R^2 \)-like measure, \( R^2_a \), for

26. Nothing says the distribution of hedgers need be stationary. Furthermore, if the variance of the return series is governed by an ARCH process, the unconditional density function of returns will tend to have fat tails. See Bollerslev, Chou, Jayaraman and Kroner (1990) for a discussion.

27. These and the use of the log of variance are the most commonly used functions in the literature. We investigated the latter as well but found much less explanatory power.

28. Note that the likelihood ratio (LR) equals \( 2\ln(L_u/L_r) \) where \( L_u \) is the likelihood of the unrestricted estimation and \( L_r \) is the likelihood of the restricted estimation. It is distributed as \( \chi^2(q) \) where \( q \) is the number of constraints.

29. Changing the weights in the market portfolio has no appreciable effect on these results.
Table 4. ARCH-M Model

<table>
<thead>
<tr>
<th>Row</th>
<th>Model</th>
<th>$\gamma_0$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\delta_0$</th>
<th>$\delta_1$</th>
<th>$L$</th>
<th>$R^2_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>No ARCH</td>
<td>0.00657</td>
<td>-</td>
<td>-</td>
<td>0.01115</td>
<td>-</td>
<td>136.36</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.5495)</td>
<td>(9.1832)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>ARCH(1)</td>
<td>0.01950</td>
<td>-</td>
<td>-</td>
<td>0.00536</td>
<td>0.64111</td>
<td>143.90</td>
<td>0.176</td>
</tr>
<tr>
<td></td>
<td>(2.0443)</td>
<td>(4.6812)</td>
<td></td>
<td></td>
<td>(2.5362)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>ARCH(1)</td>
<td>-0.01818</td>
<td>0.45259</td>
<td>-</td>
<td>0.00343</td>
<td>1.00268</td>
<td>146.31</td>
<td>0.225</td>
</tr>
<tr>
<td></td>
<td>(1.2472)</td>
<td>(2.5842)</td>
<td></td>
<td></td>
<td>(2.9342)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>ARCH(1)</td>
<td>0.00836</td>
<td>1.13089</td>
<td>-</td>
<td>0.00380</td>
<td>0.99393</td>
<td>146.85</td>
<td>0.236</td>
</tr>
<tr>
<td></td>
<td>(1.1265)</td>
<td>(2.0124)</td>
<td></td>
<td></td>
<td>(3.0569)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5)</td>
<td>ARCH(1)</td>
<td>-0.01975</td>
<td>0.46088</td>
<td>-0.21767</td>
<td>0.00352</td>
<td>0.96558</td>
<td>146.55</td>
<td>0.230</td>
</tr>
<tr>
<td></td>
<td>(1.2861)</td>
<td>(2.5461)</td>
<td>(0.6791)</td>
<td></td>
<td>(2.8789)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>and Market Return</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td>ARCH(1)</td>
<td>0.00739</td>
<td>1.16378</td>
<td>0.15299</td>
<td>0.00341</td>
<td>0.98148</td>
<td>147.02</td>
<td>0.239</td>
</tr>
<tr>
<td></td>
<td>(0.9528)</td>
<td>(2.0111)</td>
<td>(0.4940)</td>
<td></td>
<td>(3.0090)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>and Market Return</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: T-statistics are supplied below the coefficient estimates in brackets.

each estimation, where $R^2_m$ is defined as $1 - \exp(-LR/n)$ with LR being the likelihood ratio for the joint significance of the explanatory variables and n being the sample size. The $R^2_m$'s for the ARCH(1)-M models are a little less than .25. Such "fits," it should be noted, suggest that time-variation in required short-term Crude Oil futures returns is able to explain a fair percentage of return variability.

TESTS OF MARKET EFFICIENCY

Testing efficiency

Armed with some insights on pricing we next turn to tests of market efficiency. To review, efficiency implies that, based on an analysis of the

30. See Magee (1990) for details.
relevant information set, one is unable to earn excess returns, where the latter are defined as deviations of actual returns from normal (i.e. risk-adjusted) returns.\textsuperscript{31} Econometrically, this means that excess returns have a mean of zero and are uncorrelated with any information available to investors ex ante. If such correlations can be found, then investors, capitalizing on this knowledge, can take appropriate futures market positions (either long or short) in order to earn excess returns. Consider the following regression:

\[
EXR_t = \beta_0 + \sum_{i=1}^{N} \sum_{j=1}^{K} \beta_{ij}X_{it-j} + \epsilon_t,
\]

where \(EXR_t\) is the excess return at \(t\). Market efficiency implies that \(\beta_0 = \beta_{i1} = 0\) for all relevant \(X_t\).

The basic strategy of efficiency tests is to search for such correlations, but all tests of efficiency are, of necessity, tests of special cases of the hypothesis, since it is only possible for researchers to focus on subsets of the information set at any given time. With this in mind, the hypothesis is categorized into the weak, semi-strong, and strong versions, with the alternative versions only differing with respect to the relevant subset.\textsuperscript{32} As one moves toward the strong version of the hypothesis, the information set is continually expanded and ultimately contains the entire information set. That is to say, if futures markets are efficient in the strong sense, then they are also efficient in the weak and in the semi-strong senses.

Given the joint hypothesis nature of efficiency tests and the possibility of time-variation in risk premiums, the results of regressions like (11) can be interpreted in different ways. For example, the regularity that returns can be partly explained using public information, or, synonymously, that they are predictable, may mean that Crude Oil futures markets are inefficient. Another possible interpretation is that ex ante returns are both time-variant and explained by such publicly known variables.

\textsuperscript{31} Note that there is a possible distinction between ex ante and normal returns. It is possible that unexpected events may lead to changes in returns merely because the appropriate compensation for risk has changed. As will become clear, this is true in the case of CAPM, but not for the simple ARCH-M.

\textsuperscript{32} The weak form of market efficiency claims that prices in a market fully reflect all information contained in the history of prices. Markets are efficient in the semi-strong sense if prices fully reflect all publicly available information. Finally, the strong version claims that market prices reflect all information, public and private. Private information includes information possessed by, for example, corporate insiders and government officials. This form is seldom tested since few have any illusions that insiders are unable to profit (albeit improperly) from their private information. Also, this information is not available to researchers.
One's prior is crucial in this regard. Raynauld and Tessier (1984) on finding that, for several agricultural commodities, returns were correlated with some macroeconomic and market-specific variables clearly known by market participants ex ante, under their maintained hypothesis of market efficiency, concluded that risk premiums varied with time and were dependent on such factors. On the other hand, Kaminsky and Kumar (1989) concluded that markets were not totally efficient, when they found that returns (for seven primarily agricultural futures) were in some cases correlated with lagged returns for other commodities and with macroeconomic information.

The tests of energy futures along these lines have primarily been in the context of the predictive ability of futures versus time series and structural models. Bopp and Sitzer (1987), using Heating Oil futures for 1980-85, and Ma (1989), using Heating Oil, Leaded Gasoline and Crude Oil futures up to September 1986, found evidence that relevant information could be utilized to improve upon the spot price forecasts inherent in futures prices. Alternative interpretations are possible: energy futures markets may not have been totally efficient; risk premiums were predictably time-variant; or some combination of both of these was true.

How is one to proceed given such difficulties in interpretation? Given the unsettled nature of the pricing debate, categorical conclusions are out of order. Still, it seems valid to examine whether such return predictability could be reasonably viewed as reflective of a time-varying risk premium, and this will be our general approach.

Turning to our own methodology, since we have no evidence rejecting both a conditionally or unconditionally zero risk premium — except in the case of the one-month return — it appears reasonable, with this one exception, to equate excess returns and returns. An alternative approach is also utilized for the one-month return horizon. The ARCH(1)-M effect that we have demonstrated suggests that we should calculate excess returns as the difference between R, and returns predicted using our estimated ARCH(1)-M equation.

Persistence in excess returns

As discussed before the tests differ in the information set employed. We begin with weak-form efficiency tests which examine whether there is any persistence in futures returns. In tests of weak form efficiency, the information

33. This is because ex ante returns are zero if risk premiums are zero.
34. Results presented here are in terms of the specification using the standard deviation as the appropriate function of the variance in the ARCH-M equation, but results were robust to using the variance as an alternative specification.
set conditioned on is simply the set of lagged returns. Crude Oil futures returns (using one- to three-month horizons) are regressed on five lags of returns.\textsuperscript{35} In addition, for the one-month horizon, excess returns (calculated using ARCH(1)-M) are regressed on own lags. If any of the coefficients are statistically significant, this means that excess returns can be predicted (at least within-sample) on the basis of their history.

In Table 5 we present test-statistics for the null hypothesis that all slope coefficients are zero. There is no evidence against the null hypothesis at conventional levels of significance. To conclude, the weak-form efficiency test results must be viewed as unequivocally in favor of market efficiency. Volatility may be persistent, but such persistence does not extend to returns.

Table 5. Test of Weak Form Market Efficiency

<table>
<thead>
<tr>
<th>Number of Lags</th>
<th>Forecast Horizon (Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 (ARCH(1)-M)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

\textsuperscript{Notes: The critical values for the } \chi^2 \text{ test (5\%)} \text{ are 11.05 (5 DF), 9.49 (4 DF), 7.81 (3 DF), 5.99 (2DF), and 3.84 (1 DF).}

A comprehensive test of semi-strong form efficiency

Next we undertake a comprehensive semi-strong form efficiency test. We include as possible explanatory variables of Crude Oil futures excess returns a fairly obvious set. The variables conditioned on are contemporaneous and

\textsuperscript{35} In this case it makes no sense to use overlapping observations since returns will be autocorrelated by construction.
lagged values of the basis and a set of world macroeconomic variables.\textsuperscript{36}

The basis is one obvious market-specific magnitude followed by hedgers, speculators and other traders. Fama and French (1987) investigated the contemporaneous basis as a possible determinant of returns. Their empirical strategy was to decompose the (percentage) basis into the sum of a risk premium and the negative of the expected percentage change in the spot price; or, ex post, the sum of a return and "negative a" change in the spot price; and regress both of these components on the basis.\textsuperscript{37} Adding up constraints implies that variation in the basis is attributable to either variation in returns or price changes.\textsuperscript{38} For a variety of commodity futures (not including energy futures), they succeeded in detecting time-variation in returns in some commodity futures which exhibited strong seasonality in their price movements. Since their maintained hypothesis was market efficiency, their conclusion was that, in the case of these commodities, there was time-variation in risk premiums.\textsuperscript{39} Ma

\textsuperscript{36} Contemporaneous is meant in the sense of contemporaneous to the beginning of the horizon over which returns are measured.

\textsuperscript{37} The ex ante decomposition of the basis is as follows:

\[ S_t - F_{t,T} = [E_t S_T - F_{t,T}] + [S_t - E_t S_t] . \]

The following ex post analogue is used for estimation purposes:

\[ B_{t,T} = [S_T - F_{t,T}] + [S_t - S_T] , \]

where \( B_{t,T} \) signifies the \( T \)-horizon basis at \( t \). Of course the first element on the right hand side is merely the futures market return, while the second is the negative of the realized percentage change in the price of the spot commodity. Both components are regressed on the basis, as follows:

\[ S_t - S_T = \theta_0 + \theta_1 B_{t,T} + u_t \]

\[ R_t = \lambda_0 + \lambda_1 B_{t,T} + w_t . \]

These equations can be interpreted as follows. If \( \theta_1 \) is positive, the basis observed at time \( t \) contains information about the change in the spot price from \( t \) to \( T \). That is to say, the futures price has power to forecast the future spot price. As for the second equation, if \( \theta_1 \) is positive, either risk premiums are time-variant (which is Fama and French's view), or markets can be concluded inefficient.

\textsuperscript{38} Adding up constraints ensures that \( \theta_1 + \lambda_1 = 0 \) and \( \theta_1 + \lambda_1 = 1 \). Fama and French (1987) discuss some econometric issues involved here. One problem is that it is very difficult to conclude that both slope coefficients are nonzero (even if theoretically they are), especially if the basis has low variability relative to the two dependent variables. The inability of Ma (1989) and Deaves and Kinsky (1991) to detect a relationship between the basis and returns (referred to below) may potentially be due to this.

\textsuperscript{39} Our view is that a better case could have been made for concluding market inefficiency. Indeed, since they make no attempt to argue that such time-variation in premiums is in any sense a justifiable reward for risk-bearing, this seems preferable.
(1989) and Deaves and Krinsky (1991) applied this latter approach to Crude Oil futures. Both papers concluded that the contemporaneous basis was unable to explain Crude Oil futures returns, a result quite consistent with efficiency.\footnote{Ma (1989) found significant explanatory power for one of four horizons.} Here we extend this approach by utilizing lags of the basis as well as the contemporaneous basis.

The world macroeconomic variables used in explaining one- to five-month excess returns are inflation rates (i.e. growth rates in the OECD Consumer Price Index); growth in OECD industrial production; OECD narrow (i.e. M1) monetary expansion; and changes in U.S. short-term interest rates (as proxied by changes in U.S. three-month T-bill rates).\footnote{These particular variables are chosen because a number of existing studies suggest that they should have an impact on commodity prices and, therefore, presumably on returns on NYMEX Crude Oil futures. For example, see Kaminsky and Kumar (1989). Data for OECD countries are obtained from Main Economic Indicators (Historical Statistics): 1969-1988, Department of Economics and Statistics, OECD. The data for 1989 and 1990 are obtained from Main Economic Indicators (Monthly). Interest rates are obtained from the Federal Reserve Bulletin: Board of Governors of the Federal Reserve System, Washington, D.C., (various issues).} Further, since the timing of the release of these data varies among countries and announcement types, we have experimented with different lag structures.

Table 6 contains test-statistics where the null hypothesis is that all slope coefficients have no explanatory power in explaining Crude Oil futures returns. The results suggest that we can, for the most part, reject the efficient markets hypothesis when three and four lags are included. When two lags are included the results are mixed, but efficiency cannot be rejected if only a single lag is used.

Hypothesis tests that focus on single variables indicate that lagged values of OECD industrial production have the most explanatory power, mostly entering with a negative sign at two and three lags. The lag structure and the typical sign of the coefficients both require consideration.

The negative sign on OECD industrial production is without doubt puzzling. The "real activity" hypothesis of Fama (1982) may contain the kernel of a possible explanation. Empirically, Fama found a strong negative relationship between inflation and real activity growth rates. He argued that to the extent that real activity was exogenous this type of stagflation was due to the fact that "most of the variation in real money demand in response to variation in real activity is accommodated through inflation rather than through nominal money growth (p. 228)." Though oil of course is simply a single price, what we are witnessing may be the same result at the micro level. Of course another possible explanation of the negative sign is merely sample-specific statistical artifact. This question will have to be addressed by additional data and further research.
Table 6. Semi-Strong Tests using Macroeconomic Information

<table>
<thead>
<tr>
<th>Forecast Horizon (Months)</th>
<th>1</th>
<th>1 (ARCH(I)-M)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Variables (With 4 Lags)</td>
<td>Chi-Square value</td>
<td>65.10369</td>
<td>49.54958</td>
<td>36.28051</td>
<td>42.14544</td>
<td>68.81379</td>
</tr>
<tr>
<td>(DF)</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>All Variables (With 3 Lags)</td>
<td>Chi-Square value</td>
<td>34.38123</td>
<td>28.30141</td>
<td>24.99339</td>
<td>27.95743</td>
<td>24.88902</td>
</tr>
<tr>
<td>(DF)</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>All Variables (With 2 Lags)</td>
<td>Chi-Square value</td>
<td>20.36452</td>
<td>13.07626</td>
<td>13.74392</td>
<td>11.48882</td>
<td>20.27936</td>
</tr>
<tr>
<td>(DF)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>All Variables (With 1 Lag)</td>
<td>Chi-Square value</td>
<td>7.62696</td>
<td>6.26874</td>
<td>3.94270</td>
<td>5.08759</td>
<td>4.11574</td>
</tr>
<tr>
<td>(DF)</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Notes: The dependent variable in these regressions is Crude Oil futures returns (using one- to five-month horizons). The right hand side variables are various lags of OECD inflation, growth in industrial production, and monetary expansion, changes in U.S. T-bill rates, and the basis. DF denotes degrees of freedom for the $\chi^2$ test. The critical values for the $\chi^2$ test (5%) are 31.41 (20 DF), 25.00 (15 DF), 18.31 (10 DF), and 11.07 (5 DF).

A possible interpretation of the pattern of significance with respect to lags is that Crude Oil futures returns are not so much affected by output, as by public announcements of output. As stated earlier, releases tend to lag the periods that they refer to by varying amounts of time. It is a truisim in financial markets that, assuming markets are efficient, only the unanticipated component of the relevant announcement, or "surprises," should have an impact on asset prices and returns. Expectations should already have been embedded in market prices. We have undertaken no such a decomposition, but it is straightforward
to show that actually announced economic variables will be correlated with such surprises. To demonstrate inefficiency it would be necessary to prove that returns were predictable based on previously made announcements, but, given the uncertainty as to when the announcements were actually made, it is not possible for us to do this. Thus, on this basis, rejection of efficiency is not an obvious conclusion. \footnote{42}

There are three additional reasons for caution. First, out-of-sample forecasting reliability has not been demonstrated. There is a salient difference between finding correlations within sample and using these relationships for predictive purposes out of sample. Whether the latter is possible requires future investigation which is beyond the scope of this paper. Second, to calculate excess returns one must factor in the relevant transaction costs. Given the low commissions and market impact costs for such highly liquid contracts as Crude Oil futures, this is probably only a minor concern. \footnote{43} Finally, as mentioned at length above but noted here for the sake of completeness, all tests of efficiency are tests of a joint hypothesis, and rejection may be merely due to the use of an inappropriate pricing model.

CONCLUSION

The New York Mercantile Exchange's Crude Oil futures contract has enjoyed tremendous success since its inception in 1983. Proper utilization of it as a speculative or hedging vehicle requires a knowledge of pricing relationships and market efficiency, an endeavor attempted herein.

We investigated risk premiums and efficiency in this study, because both concepts are important for traders in the energy futures market. Speculators can profit from any departures from market efficiency that they succeed in detecting. Hedging efficacy is positively related to the magnitude of the risk premium, and the extent to which futures markets are efficient.

During 1983-90, there is some evidence that short-term unconditional premiums were positive, as evidenced by positive mean unconditional one-month returns when ARCH estimation was conducted. This implies that typically a short hedger (such as a producer) may have had to accept a price below which would have been available at spot in order to induce long speculators to accept price risk. That is, such a hedger had to pay a risk premium. As for time-

\footnote{42} It should be granted that no attempt has been made here to specify a structural model which fully accounts for the relationships among oil prices, macroeconomic variables and other relevant variables (such as inventories and technological change). This is an exercise beyond the scope of this paper.

\footnote{43} Deaves and Krinsky (1992) cite 1% as an absolute upper bound.
variation, no evidence was produced that futures returns could be explained by returns on a broad market portfolio, but, employing the ARCH-M methodology, we did find evidence that futures returns covaried with recent volatility. This suggests that the risk premium paid by hedgers has varied with market conditions.

As for efficiency, our findings were inconclusive. No evidence was produced consistent with weak form inefficiency, but our comprehensive semi-strong form efficiency test utilizing contemporaneous and lagged values of the basis and a set of world OECD macroeconomic variables did produce results that suggested possible inefficiencies. Still, hasty conclusions against efficiency may be unwarranted, given the anomalous negative sign on the OECD industrial production coefficients, the ambiguity as to exactly when the OECD information was truly available to the public, and the absence of demonstrated out-of-sample forecasting reliability. It goes without saying that further research would be fruitful in this regard.

REFERENCES


