Exploration Activity, Long Run Decisions, and Roll Returns in Energy Futures*

Alexander David

University of Calgary

November 2012

Abstract

We present some evidence that firms' inventories as well as their long term expenditures on exploration and development (E&D) each help to predict the slope of the oil futures curve one year ahead. In addition we show that roll strategies in futures contracts conditioned on E&D expenditures rather than inventories have a stronger performance over the past 25 years. Historical data displays significant components of very low (less than once in six years) and very high (more than once every two months) frequency variation in prices, which supports the presence of both long and short term risk. Building on the work of Litzenberger and Rabinowitz (1995), we develop a theoretical model where firms change E&D expenses over the business cycle to manage the value of their extraction options. Firms optimally invest in short bursts when aggregate resource demand shifts from a stable to a volatile regime, and when their capital stock is far from the new optimum level. Such adjustments happen infrequently only when the stable regime has persisted for a fairly long period. The model is able to shed light on the stylized facts.

---

*I thank Jaime Cassasus, Lars Alexander Kuehn, Bryan Routledge, Chester Spatt, as well as participants at the Alberta Finance Institute Commodities Conference, Carnegie Mellon Finance Area, 2013 Western Finance Association Meetings, for their helpful comments. Address: Haskayne School of Business, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada. I am grateful to the SSHRC for a research grant. Phone: (403) 220-6987. E. Mail: adavid@ucalgary.ca.

Introduction

Recent years have seen the development of increasingly sophisticated technologies for the extraction of natural resources.[1] These new technologies brought to fruition by investments by the resource extraction industry have changed the current and expected future prices of resources and have important consequences for energy self sufficiently and stability of growth for North America. In this paper we ask if the investment in exploration and development (E&D) of resources has an impact or is affected by the keenly watched market statistics of current and future prices of the resource.

Of the most widely watched statistics in the futures market is the weak basis, which is the discounted value of the futures price less the spot price of the resource. We work with a closely related statistic, the relative basis, which is the weak basis divided by the current spot price of the resource. When this quantity is positive (negative) we say the futures market is in weak contango (backwardation). Of interest to practitioners and researchers is the economic information that determines the relative basis. The theory of storage (Kaldor (1939) and Working (1948)) implies that the futures relative basis is strongly positively related to inventories. We call this the "short-run" information about resource prices in the futures relative basis. However, it is now well recognized that the storage theory, though very useful, is unable to explain the basis in certain periods.[2] In addition in this paper, we argue that the futures relative basis also contains "long-run" information about resource prices, which has important implications for decisions such as the exploration and development of the resource extraction process. In particular, we will develop four stylized properties of oil futures prices that arise from the long-run risks faced by energy producers.

Stylized Fact 1: Long run decisions of oil and gas firms on E&D explains more of the variation in the futures relative oil basis than short term inventory decisions.

[1] In Alberta, Canada, new techniques have been developed to extract crude oil from bitumen using less water and energy and damage to the environment than previously envisaged. In the United States, new hydraulic fracturing technology has made the gas trapped in previously inaccessible shale rock now economically feasible to extract.

[2] In June 2009, a story on Bloomberg reported the following carry transaction conducted by J.P. Morgan: The spot price of oil was $550 per ton. At the time the August futures sold at $583 a ton. The storage cost at the time was about $9 per ton in a supertanker (an entire supertanker had to be purchased). The riskless rate was close to zero, so the futures basis was larger than the cost of carry. These facts are incompatible with the basic storage models.

Table 1 reports simple linear regressions at a quarterly frequency of the futures relative basis on inventory and the new capital raised by E&D firms as a share of U.S. GDP (see the data appendix for sources of data).[3] As can be seen in the top panel, while one quarter lagged inventory explains about 10 percent of the variation in the basis, lagged new capital for exploration explains explains about 18% of the variation in the relative basis (lines 1 and 2). When both variables are considered, we explain about 23% of the variation in the relative basis, and each variable remains significant. This suggests that both short and long run decision by firms are important determinants of the futures relative basis. The periods when the discounted futures price is higher than the spot price, inventory accumulates. In addition, firms raise more capital for E&D expenses in response perhaps to higher futures prices. However, as seen in the second panel, only E&D expenses can predict the basis four quarters ahead.

For the one quarter ahead forecast, the fitted values from the regressions along with the time series of the weak relative basis are shown in Figure 2. As seen, the fitted values of the right panel, which also use information in the NCS are much closer to the historical relative basis than the fitted values in the left panel. This is particularly true in periods of contango, which have frequently prevalent in the 2005-2010 periods. It is also notable that adding US production to the above regression does not increase our ability to explain the relative basis.

For natural gas, as seen in the third and fourth panels we have similar results for the one quarter ahead forecast of the basis. For the four quarter ahead forecast, the sign on inventory becomes negative, while that of NCS remains positive and significant.

Stylized Fact 2: Long run decisions of oil and gas firms on E&D explains more of the variation in the WTI oil and natural gas futures roll returns than short term inventory decisions.

An intriguing aspect of returns on futures trading strategy is the roll return, which has been documented in Gorton and Rouwenhourst (2006) and Erb and Harvey (2006). The roll strategy

[3] Throughout this paper we look at the statistics of the one-year futures relative basis. While it would be of interest to study longer maturity futures, we are constrained by the lack of long historical times series on these longer term contracts. Two-year contracts started trading actively in mid 1990 and four-year contracts only in 1997. The correlations of the relative basis of the two-year and four-year contracts with the one-year contract are 99.7 and 98.7 over the subsamples, respectively.

involves basing a long/short decision based on the sign of the slope of the futures curve.[4] These authors find that the roll return is a very important component of total returns earned on commodity investing, and in fact, seems to dominate the risk premium on the commodities as a component of returns. Gorton and Rouwenhourst (2006) finds that sorting on the basis across commodities leads to a positive expected return, while Erb and Harvey (2006) finds the difference in roll returns is a large component of the cross sectional variation in excess returns across commodities.

In this paper we will study two alternative determinants of the roll trade return, inventories and long term E&D expenses. Table 2 provides regression of one year holding period returns of the roll trade on these two variables. The position is rebalanced quarterly and held for a year. The 'spot' price in these transactions is the nearest term futures price that has an active trade. As can be seen, both variables predict oil returns with R^2 of about 5.5 percent on this roll trade, which shorts (longs) the one year futures if the basis is positive (negative). Moreover, when put together, the total explanatory power almost doubles. For natural gas, only the E&D expenses are significant in forecasting returns.

To take advantage of the predictive power of these variables, we formulate new strategies, where the roll is not based on the current basis, but variables that predict the future basis. In particular, we short (long) the one year futures contract, if the variable is above (below) its historical mean. Table 3 shows some statistics based on these and related rolling strategies. Some useful points emerge: First, the return on always going long, and formulating the roll strategy on oil futures based only on the current basis are both positive, and of a similar magnitude of about 6.8 percent. Conditioning the roll on the inventories gives an average return of 3.4 percent (albeit with a higher Sharpe ratio), while conditioning on the E&D expenses, raises the average to above 15 percent. The strategies are by no means safe. Figure 3 provides the time series of the returns on all four strategies. As can be seen, the E&D variable provides useful timing information relative to inventories, as on several occasions, we see positive returns

[4] The following example from Erb and Harvey (2006) illustrates the intuition underlying the roll trade: On May 30th 2004, the July 2004 futures traded at $40.95, while the July 2005 futures traded at $36.65. Therefore, oil futures were in backwardation. Now, if the speculator goes long on the July 2005 futures and if the basis of $4.3 persists, then holding the contract for a year will lead to a gain of $40.95/36.65 - 1 = 0.117$. In this example, the futures position is not fully collateralized. Gorton and Rouwenhourst (2006) recommends that we subtract the riskless rate of return from this gross return for the collateral held in Treasuries.

when conditioning on the former strategy and negative on the latter strategy. Figure 4 shows the histograms of their returns. As can be seen, the skewness of the always long position in futures is negative, while on the conditioned roll strategies the returns are positive.

For natural gas, we again find that roll strategy conditioned on high E&D expenses provides the highest returns and Sharpe ratio.

Stylized Fact 3: It is natural to think of long-run risk as pertaining to events that happen less frequently. Figure 5 shows the variance frequency decomposition of the weak WTI relative basis $[e^{-r(t)}F(t) - S(t)]/S(t)$, where $r(t)$ is the one year Treasury Bill rate, $F(t)$ is the one year futures price, and $S(t)$ is the spot price. As seen in the figure, there is a large amount of variance of the weak relative basis that is of low frequency movements (every 6 years or less often). There is also a large proportion of high frequency movements (every 2 months or more often). Overall, the frequency decomposition is U-shaped, with these two extreme frequency movements explaining more of the variation in the relative basis than intermediate frequency movements. This is to be contrasted with the analogous decomposition for most macroeconomic series such as industrial production, which show a spike at a frequency such as 2.5 years (see e.g. Figure 6.5 in Hamilton (1994)). The lack of such a frequency suggests that the commodity pricing cycle is distinct from the business cycle, but we will examine this more in detail in this paper.

Stylized Fact 4: As seen in Figure 6 the realized volatility of WTI oil prices and its relative basis have a U-Shaped relation. Volatility is high for extreme backwardation or extreme contango, each of which are relatively low frequency events.

In the remainder of the paper we will build a model of the long run decision making of resource producing firms that in equilibrium will lead to futures prices having these three stylized properties. Following Brennan and Schwartz (1985), another way of looking at the relative basis is the convenience yield embedded in futures prices.[5] The convenience is the benefit that accrues to the owner of the resource but not to the long counterparty of the futures contract. The convenience yield is normally associated with the storage model, where the resource held can be used to enhance the profits of the holder who sells into the market at

[5]It is common practice to write $F(t) = S(t)\, e^{(r+u-y)(T-t)}$, where $F(t)$ is the forward price at date t, $S(t)$ is the spot price, r is the riskless rate, u is the proportional storage cost, and y is the implicit convenience yield.

strategic points of times, in particular when the resource is scarce (see e.g. Routledge, Seppi, and Spatt (2000)). We instead follow Litzenberger and Rabinowitz (1995) (LR) in endogenously determining the convenience yield of the *unextracted resource* held in the ground. In this setup, if the resource price is low, the resource is not extracted and hence the convenience yield is the expected present value of the *price protection services* associated with storing the resource in the ground, rather than holding a forward contract. Clearly, the holder of the futures contract cannot make the optimal extraction decision.

LR argue that several properties of empirical futures prices can be explained by the above mentioned optionality of unextracted reserves. In particular, the futures relative basis is negatively related to the prices of put options on the spot, which is the discounted expected value of the price protection services.

At odds with data for most commodities (see e.g. Figure 2), the LR model implies backwardation all the time when extraction costs increase at the riskless rate. Having the growth rate of extraction costs fluctuate enables this model to have both backwardation and contango. For example, if the rate of extraction costs increases more rapidly than the riskless rate, then producing firms in compensation will demand a higher futures price. We assume that the growth of extraction costs depends on the accumulated capital in the industry, which can be augmented sequentially by E&D expenses. This directly implies that measures of E&D should empirically affect the relative basis. We formulate a model in which firms in resource extraction optimally develop the capital that determines their future extraction costs, and find that firms optimally choose higher E&D in periods when the futures market relative basis is larger, which is the observed stylized fact.

A key feature of our model is the shocks to consumers' demand for the resource. We assume as a standard Markov chain structure for these shocks, with alternating low volatility ("stable") and high volatility ("volatile") regimes. We show that the optimal E&D investment process and its slow impact on the accumulated capital of the resource producing firms helps explain the two additional stylized facts. In a persistent particular economic regime, firms' capital moves slowly towards its optimal level in that regime. All else equal, this implies a direct effect that extraction costs increase slower in stable times leading to more backwardation.

However, there is additional impact, which is asymmetric over the business cycle. Because investment is constrained to be positive, and capital can only be freed at the rate of depreciation, upon a transition from volatile to stable, capital can only adjust slowly downward. In this period, investment essentially comes to a halt, and actual capital becomes temporarily far away from the new optimum even though it slows moves towards it. When the regime switches to being volatile again, the capital stock is far below its optimal level so that extraction costs are large, but because of the high volatility, call options increase in value accompanied by rapid investment. In such periods, due to the high extraction costs the the futures market goes into steep contango. In following periods, as the capital stock builds towards its optimum, firms react by cutting more production, which leads to an increase in the current price, and a decline in the relative contango. As more capital builds, extraction costs decline, and the futures market returns to backwardation.

This is consistent with the first stylized fact. This asymmetric adjustment process leads to low frequency variation of the relative basis since the investment effect above happens only for a brief period when the economy switches from the stable to volatile regime, which is the second stylized fact. The difference in the models' basis across regimes also makes the basis somewhat predictable in the model, which leads to roll returns when conditioned on investment, as in the data. It is important to note that the roll return in our model arises despite not having a risk premium. Thus our model provides a rationale for roll trade returns not based on risk. This contrasts our work from a recent paper by Baker and Routledge (2012), which reports that risk premiums on commodities are also related to the futures basis. Finally, the U-shape relation between the relative basis and price volatility happens when this above effect is combined with the assumed high volatility of demand shocks in the strong regime.

Our model contributes to the literature on resource extraction and storage. Most existing models have either one of these features. Models of storage assume exogenous extraction decisions (e.g. Deaton and Laroque (1992) and Routledge, Seppi, and Spatt (2000)). Models with endogenous extraction or production of resources on the other hand, allow no storage (e.g. Pindyck (1980), Litzenberger and Rabinowitz (1995), Carlson, Khokher, and Titman (2007), Casassus, Collin-Dufresne, and Routledge (2008), and Kogan, Livdan, and Yaron (2009)). In the context of agricultural commodities, there is an older literature that has production and storage in equilibrium, but the analysis in such models do not apply to exhaustible resources, where equilibrium profits are compatible with competitive equilibrium do to limitations in supply (e.g Scheinkman and Schechtman (1983)). With exhaustible resources, as pointed out in Litzenberger and Rabinowitz (1995), there are profits at the time of extraction that are optimized by the extraction timing decision of resource firms. In this paper we start by analyzing the E&D decision in an extraction model without storage similar to Litzenberger and Rabinowitz (1995). We show that such a model is consistent with the stylized facts noted above. We attempt to extend the analysis to the case where there is both extraction and storage, but so far have only partially solved this model. What is clear in the analysis so far is that the exploration activity of energy firms also affects the tradeoff between extraction delay and inventory in carrying the resource through time. Essentially, in the model with storage, the firm has two substitutable ways of providing resource to customers at date 1: it can either defer date 0 extraction and extract in date 1, or it can extract in date 0, and carry inventory to date 1. If for example, extraction costs are expected to increase rapidly, it will make the firm more likely extract more in date 0 and carry inventory.

The paper structure is as follows. In Section 1, we formulate a model of optimal resource extraction and the relative basis. In Section 2, we simulate the model and examine its ability in replicating the three stylized facts. Section 3 concludes. Two appendices contain the data description and sources, and some technical results.

1 A Simple Two Period Model of Resource Extraction and Exploration Activity

We build on the two period version of the model of LR, with a few differences. The most significant addition is of an E&D (investment) decision that reduces costs of future extraction of the resources. To tractably analyze the investment decision with technology spillover, we introduce multi-plant firms. Assume a continuum of price taking identical resource production multi-plant firms, each of which owns an equal share of reserves. We will focus our analysis on the representative firm.

We start with a description of the demand side of the model. The demand function for the resource at time t is given by simple function $q_t = f(S_t, \epsilon_t)$, where ϵ_t is a demand shock realization for the resource at date t. Without loss of generality, we set $\epsilon_0 = 0$, and $\epsilon_1 = \epsilon$. Conditional on a realization of ϵ, the inverse demand function is $s = f^{-1}(q_t; \epsilon_t)$.

Supply of the resource is optimally determined by the firm. The resource is of varying quality, which is parsimoniously captured by heterogeneous extraction costs. Let extraction costs across grades of resources be uniformly distributed $x \in [0, \bar{x}]$ in period 0. Plant x owned by the firm has access to technology with extraction cost in period 0 of x. Let R_0 be the reserves available at date 0. At date 0, the plant level decisions determine the cutoff reserve quality (extensive margin), x_0^e. Then at date 1, the available reserves are $R_1 = R_0 - x_0^e/\bar{x}\, R_0$.

We assume that all investments in technology are made at the firm level, and that extraction costs for resource of all qualities increases at a rate $g(K_1)$, where K_1 is the amount of capital in E&D. Note, that the E&D expenses only affect future extraction costs and hence even in this simple two-period model can be used to manage long run or future risks. The timing of capital installation is as follows: At date 0, the firm inherits capital of K_0 from past decisions. The firm can augment this capital stock by incurring E&D expenses, which we call investing. The new capital will follow the standard process

$$K_1 = (1 - \delta)\, K_0 + I_0. \tag{1}$$

The investment choice is made *before* before any extraction decisions are made. The installed capital determines the growth rate of extraction costs, $g(K_1)$ over the next period. Conditional on the investment choice at the firm level, each plant chooses its extraction decision to maximize the profits of the plant. Conditional on the firm level investment, the plant level maximization can be written as:

$$\pi_0^x = \max_{0 \le Q_0^x \le \frac{R_0}{\bar{x}}} S_0\, Q_0^x - Q_0^x x \,+\, e^{-r} E[(S_1 - x\, e^{g(K_1)})^+]\, (\frac{R_0}{\bar{x}} - Q_0^x). \tag{2}$$

The Kuhn-Tucker optimality condition for the extraction choice of firm x satisfy

$$\left[S_0 - x \,-\, C(x\, e^{g(K_1)})\right]\, Q_0^x = 0 \quad \text{or,} \quad \left[S_0 - x \,-\, C(x\, e^{g(K_1)})\right]\, (\bar{x} - Q_0^x) = 0 \tag{3}$$

In particular, for a firm with positive and interior production

$$S_0 - x = C(x\, e^{g(K_1)}), \tag{4}$$

where $C(x)$ is the value of a 1-period call option with exercise price of x. The left-hand side is the net gain to current extraction, while the right-hand side is the value of delaying extraction. It is useful to note at this point that the call option valuation in (4) is quite similar to a regular American option, with the only difference being that the price at each date of the resource is determined by the aggregate optimal extraction decision of all producers using the inverse demand function. We complete the analysis of the model by determining the investment choice at date 0 in the context of the model without and with storage in the following subsections.

1.1 Model Without Storage

We now show how the cutoff resource quality (the extensive margin is determined) x_0^e. At date 1 since there are no further options and no inventory, all plants with available resource and extraction costs smaller than the price will extract. Therefore aggregate production at date will be

$$Q_1(x^e, \epsilon) = \left(\int_{x_0 e^{g(K_1)}}^{S_1} \frac{1}{\bar{x}\, e^{g(K_1)}} dx \right) R_0 = \frac{S_1\, e^{-g(K_1)} - x_0^e}{\bar{x}}\, R_0. \tag{5}$$

Hence, the date 1 price is $\tilde{S}_1 = s(Q_1(x^e, \epsilon); \epsilon)$. Let $C(x|x_0^e, K_1)$ be the value of the extraction call option for the firm when the extensive margin is x_0^e and K_1 capital is brought into date 1. Then x_0^e satisfies the fixed-point condition:

$$S_0 - x_0^e = C(x_0^e\, e^{g(K_1)} | x_0^e, K_1), \tag{6}$$

when it lies in the interior of the interval $[0, \bar{x}]$, and with the boundary conditions:

$$\begin{aligned} x_0^e(K_1) &= 0 \text{ if } s(0|0, K_0) < C(0|0, K_1), \\ &= \bar{x} \text{ if } s(R_0|\bar{x}, K_0) - \bar{x} > C(\bar{x}|\bar{x}, K_1). \end{aligned}$$

The firm maximizes total profit at date 0

$$
\begin{aligned}
\pi_0 &= \max_{I_0>0} \left[\int_0^{\bar{x}} \pi_0^x dx - P_0\, I_0 \right] \\
&= \max_{I_0>0} S_0\, \frac{x_0^e}{\bar{x}}\, R_0 - \frac{(x_0^e)^2}{\bar{x}}\, R_0 \;+\; \left(\int_{x_0^e}^{\bar{x}} C(x\, e^{g(K_1)} | x_0^e, K_1)\, dx \right) \frac{R_0}{\bar{x}} \;-\; P_0\, I_0 \qquad (7)
\end{aligned}
$$

where P_0 is the price of capital at date 0 in consumption goods at that date. To compute expected profit we calculate the maximal investment choice numerically by choosing over a grid of values.

We now make specific assumptions on the demand function and the distribution of shocks that enable us to solve for the firm value in closed form. Specifically, similar to LR, we assume a linear demand function in each period of the form: $q_t = (a + \epsilon_t) - b\, S_t$. We assume that the shocks ϵ_t are i.i.d. distributed LN(μ, σ), where LN is the log-normal distribution. Since the resource prices at each date are dependent on the extraction choices of firms, which in turn depends on their choice of installed capital at date 0, we first formulate the value of the extraction option conditional on both these variables.

Proposition 1 *The value of the extraction call option at date 0, given installed capital K_1, and cut-off resource quality $x_0^e \in [0, \bar{x}]$ for a resource with current extraction cost of x is*

$$
\begin{aligned}
C(x|x_0^e, K_1) &= \frac{e^{-r}}{D} \left[e^{(\mu + 0.5\,\sigma^2)} N(-d_1) - k N(-d_2) \right], \\
d_1 &= \frac{\log(k) - m - \sigma^2}{\sigma}; \qquad d_2 = \frac{\log(k) - m}{\sigma}; \\
k &= D\, x\, e^{g(K_1)} - a - \frac{x_0^e}{\bar{x}} (1 - \frac{x_0^e}{\bar{x}})\, R_0; \\
D &= b + \frac{e^{-g(K_1)}}{\bar{x}} (1 - \frac{x_0^e}{\bar{x}})\, R_0.
\end{aligned}
$$

The value of a put option is

$$
P(x|x^e, K_1) = \frac{e^{-r}}{D} \left[k\, N(d_2) - e^{(\mu + 0.5\,\sigma^2)} N(d_1) \right].
$$

The proof is in the appendix.

The stock price for the linear demand case at date 1 that is derived in the proof gives us a straightforward formulation of the forward price. In particular, we have

$$F = E[s(q_t;\epsilon_t)] = \frac{a + \frac{x_0^e}{\bar{x}}\,R_0 \,+\, e^{\mu+0.5\sigma^2}}{b + \frac{1}{\bar{x}}\,e^{-g(K_1)}\,R_0}.$$

The spot price at date 0 is $S_0 = 1/b(a - \frac{x_0^e}{\bar{x}}$.

What does this simple two-period model imply about the relationship between investment and futures basis? While it is hard to sign this relationship in general, we can for given extraction x_0^e decisions. In this case, as seen above, the futures price is increasing in the rate of growth of extraction costs, while the spot price, conditional on x_0^e does not depend on it. Therefore, the futures basis is increasing in $g(K_1)$. An increase in the extraction costs implies a lower expected supply in the future, so that prices will be higher in the future, Under the reasonable assumption assumption that $g'(K_1) < 0$, once again for a given extraction choice we will then get a negative relationship between changes in investment and the futures basis. However, the relationship between the futures basis and investment in levels generated by this model might well be positive, since investment is triggered in periods of low capital and hence high growth of extraction costs.

1.2 Model With Storage

As mentioned in the introduction, existing models of resource extraction do not allow for storage, while models with inventory do not have optimal resource extraction. In addition, none of these models have exploration activity. Here we provide the analysis of a model with production, storage and exploration. The model will help us address the stylized facts noted in the introduction on the positive comovement of exploration activity, extraction, and inventory accumulation.

We continue to formulate the decisions of the multi-plant firm in the subsection 1.1 assuming once again that E&D investment decisions are made before extraction and inventory decisions. We assume that investment and inventory decisions are made at the firm level, while extraction decisions are made at the plant level. Essentially, in the model with storage, the firm has two substitutable ways of providing resource to customers at date 1: it can either defer date 0 extraction and extract in date 1, or it can extract in date 0, and carry inventory to date 1. Which strategy is more profitable? Each has its own advantages, and the tradeoff is to a large part determined by storage costs and the expected change in extraction costs. If the latter are expected to increase rapidly, for example, it might be worthwhile for the firm to extract in date 0 and carry inventory. In addition, the price protection offered by holding the resource in the ground (as in the case of no storage) implies that an increase in uncertainty will make the delayed extraction choice more profitable.

The plant level optimization is very similar to the case without storage, albeit with different equilibrium resource prices. The objective function of the plant still satisfies (2) and its optimal extraction policy is determined as in (4). Given this, the profit at the firm level is

$$\pi_0 = \max_{I_0 > 0} \max_{Z_0 \in [0, \frac{x_0^e}{\bar{x}} R_0], x_0^e \in [0, \bar{x}]} E\left[S_0 \frac{x_0^e}{\bar{x}} R_0 - S_0 Z_0 - 0.5 \frac{(x_0^e)^2}{\bar{x}} R_0 - P_0 I_0 + e^{-(r+u)} \tilde{S}_1 Z_0\right] + \left(\frac{1}{\bar{x}} \int_{x_0^e}^{\bar{x}} C(x\, e^{g(K_1)})\, dx\right) R_0. \tag{8}$$

We continue to assume that the extraction decision of the firm is not able to change total futures reserves, so that the first order condition for an interior choice of x_0^e still satisfies (4). However, we now assume that inventory decisions by the firm are non-negligible and have an impact on the future price of the resource, and hence the future extraction option. Given our assumption on the inverse demand function we can write the price at date 1 as

$$\tilde{S}_1 = s(Q_1 + Z_0\, e^{-u}; \tilde{\epsilon}). \tag{9}$$

The first order condition with respect to inventory, Z_0 is

$$-S_0 + e^{-(r+u)} E[S_1] + \frac{d}{dZ_0}\left(\frac{1}{\bar{x}} \int_{x_0^e}^{\bar{x}} C(x\, e^{g(K_1)})\, dx\right) R_0 = 0. \tag{10}$$

In the special case where the derivative of the inverse demand function is independent of ϵ (for example the linear demand case), we can further write,

$$-S_0 + e^{-(r+u)}\, E[S_1] + e^{-u}\, s'(Q_1 + Z_0\, e^{-u}) \left(\frac{1}{\bar{x}} \int_{x_0^e}^{\bar{x}} \Delta_C(x\, e^{g(K_1)})\, dx \right) R_0 = 0, \qquad (11)$$

where $\Delta_C(x\, e^{g(K_1)})$ is the delta of the extraction call option. In the presence of the storage technology, we also have the forward price equal the expected spot price, so that the weak basis can now be written as

$$e^{-r} F_1 - S_0 = S_0\,(e^u - 1) - s'(Q_1 + Z_0\, e^{-u}) \left(\frac{1}{\bar{x}} \int_{x_0^e}^{\bar{x}} \Delta_C(x\, e^{g(K_1)})\, dx \right) R_0, \qquad (12)$$

where $s'(\cdot) < 0$. Now, the weak relative basis is no longer bounded by $(e^u - 1)$ as in standard storage models. Since inventories are a direct substitute to the extracted resource, an increase in inventories partially destroys the value of the extraction options outstanding. For the resource producing firm then, the futures prices has to compensate the firm for providing the resource at the future date by the amount of the option value destruction in addition to the cost of storing the resource.

Specializing again to the linear demand function $q_1 = (a + \epsilon) - b\, S_t$, where ϵ is distributed LN(μ, σ), we can derive the option values similar to Proposition 1.

Proposition 2 *The value of the extraction call option at date 0 in the presence of a storage technology with proportional storage costs of u, given installed capital K_1, and cut-off resource quality $x_0^e \in [0, \bar{x}]$ for a resource with current extraction cost of x is*

$$\begin{aligned}
C(x|x_0^e, K_1) &= \frac{e^{-r}}{D^s} \left[e^{(\mu + 0.5\,\sigma^2)} N(-d_1^s) - k N(-d_2^s) \right], \\
d_1^s &= \frac{\log(k^s) - m - \sigma^2}{\sigma}; \qquad d_2^s = \frac{\log(k^s) - m}{\sigma}; \\
k^s &= D^s\, x\, e^{g(K_1)} - a - \frac{x_0^e}{\bar{x}} (1 - \frac{x_0^e}{\bar{x}})\, R_0 \,+\, Z_0\, e^{-u}; \\
D^s &= b + \frac{e^{-g(K_1)}}{\bar{x}} (1 - \frac{x_0^e}{\bar{x}})\, R_0.
\end{aligned}$$

The value of a put option is

$$P(x|x^e, K_1) \;\; = \;\; \frac{e^{-r}}{D^s}\left[k^s\, N(d_2^s) - e^{(\mu+0.5\,\sigma^2)} N(d_1^s)\right].$$

The proof is in the appendix.

For the linear demand case (12) implies that the weak basis satisfies

$$e^{-r}F_1 - S_0 = S_0\,(e^u - 1) - \frac{e^{-u}}{b}\left(\int_{x_0^e}^{\bar{x}} \Delta_x\, dx\right)\frac{R_0}{\bar{x}}, \tag{13}$$

where

$$\Delta_x = -\frac{e^{-r}}{D^s} N(-d_2^s). \tag{14}$$

We now provide a description on how the extensive margin and inventories are jointly determined in the storage version of the model. Since the firm has the extraction options, the extensive margin is still determined by the fixed point condition in (6), where all prices in this option calculation are dependent on the inventory choice. In particular $\tilde{S}_1 = s(Q_1 + e^{-u}\, Z_0; \epsilon)$. and $S_0 = s(Q_0 - Z_0; 0)$. To determine Z_0, notice that for an interior choice of inventory, we require the weak basis to satisfy both (**??**) and (12). Equating the right hand sides of the two equations now provides a second equation linking x_0^e and Z_0, so that both variables can be determined. If $F_1 - S_0 < S_0\,(e^u - 1) - \frac{e^{-u}}{b}\left(\int_{x_e^0}^{\bar{x}} \Delta_x\, dx\right)$, where $S_0 = s(Q_0; 0)$, then no inventory is carried. The full characterization of the storage and extraction equilibrium is complicated by the boundary conditions for extraction as well as inventory, and we will provide further details of this equilibrium in future versions of this paper.

2 The Infinite Horizon Model with Production, Exploration, and Storage

We preserve much of the structure of the 2-period model. The one additional assumption that we make here is that there are adjustment costs to investment, an assumption that is standard in the investment literature to reduce the volatility of the investment process. This will help us provide a more empirically realistic model relationship between investment and the futures basis.

The demand function for the resource at time t is once again given by $q_t = f(S_t, \epsilon_t)$, where ϵ_t is a demand shock realization for the resource at date t. Conditional on a realization of ϵ_t, the inverse demand function is $s = f^{-1}(q_t; \epsilon_t)$. We assume that the demand shock follows a 2-state regime switching process. In particular $\log(e_{t+1}/e_t) \sim N(\mu_i, \sigma_i)$, where the regime $i \in \{1, 2\}$, and switches between these states with transition probability:

$$\Lambda = \begin{bmatrix} 1-\lambda_{12} & \lambda_{12} \\ \lambda_{21} & 1-\lambda_{21} \end{bmatrix}$$

The plant level decisions determine the cutoff reserve quality (extensive margin), x_t^e. Then at date t, the total production equals

$$Q_t = R_0 \cdot \int_{x_{t-1}^e e^{g(K_{t-1})}}^{x_t^e} \frac{1}{\bar{x} e^{g(K_{t-1})}} dx = R_0 \, \frac{x_t^e - x_{t-1}^e e^{g(K_{t-1})}}{\bar{x} e^{g(K_{t-1})}}. \tag{15}$$

The total extraction costs incurred by the firm at date t are

$$C_t = R_0 \cdot \int_{x_{t-1}^e e^{g(K_{t-1})}}^{x_t^e} \frac{x}{\bar{x} e^{g(K_{t-1})}} dx = \frac{1}{2} R_0 \, \frac{x_t^{e2} - (x_{t-1}^e)^2 e^{g(K_{t-1})2}}{\bar{x} e^{g(K_{t-1})}}. \tag{16}$$

We assume that the firm also has a costly storage technology. It is able to place a non-negative quantity Z_{t-1} in storage at time t, We assume that the storage costs are a proportion u of the quantity stored. So, an amount Z_{t-1} placed in storage at $t-1$ will make available an amount $Z_{t-1} \, e^{-u}$ at period t. The firm behaves competitively in production markets, and we assume here that its storage decisions has no price impact either. We will extend the analysis for the case of a non-negligible storage decision in future versions of the paper. For the competitive case, we alternatively, we could assume that inventory decisions are made by a risk neutral speculator. However, with complete markets, the equilibrium will be identical with storage by either the firm or speculators. Combining production as in (15) and inventory, the total amount available for consumption in period t is

$$q_t = Q_t + Z_{t-1}e^{-u} - Z_t. \tag{17}$$

If there is a stockout, then $Z_t = 0$, that is, all available resource is consumed in period t.

To solve for equilibrium prices and quantities, we solve the related problem of a social planner who maximizes the discounted expected consumer plus producer surplus (see e.g. Weinstein and Zeckhauser (1975) and Carlson, Khokher, and Titman (2007)). The social surplus at time t is therefore,

$$\text{SS}_t = \int_0^{q_t} s(x;\epsilon_t)dx - C_t - P_t\, I_t, \tag{18}$$

where total production, costs of production, and consumption, are given in (15), (16), and (17), respectively, and P_t is the price of capital goods in units of consumption goods at date t. We hold $P_t = 1$ for all t.

The social planning problem can be solved by standard dynamic programming methods. The Hamilton-Jacobi-Bellman equation is

$$J(x^e_{t-1}, Z_{t-1}, K_{t-1}) \;=\; \max_{(x^e_t \in [x^e_{t-1}\, e^{g(K_{t-1})}, \bar{x}], Z_t \geq 0, 0 \leq I_t \leq \kappa\, K_{t-1})} SS_t \;+\; \frac{1}{1+r}E[J(x^e_t, Z_t, K_t)]. \tag{19}$$

The first order conditions for this problem are:

$$\frac{R_0\;(s(q_t\;\epsilon_t) - x^e_t)}{\bar{x}\; e^{g(K_{t-1})}} \;+\; \frac{1}{1+r}E[J_x] \;\leq\; 0; = 0 \quad \text{if} \quad x^e_t > x^e_{t-1}e^{g(K_{t-1})} \tag{20}$$

$$\frac{R_0\;(s(q_t\;\epsilon_t) - x^e_t)}{\bar{x}\; e^{g(K_{t-1})}} \;+\; \frac{1}{1+r}E[J_x] \;\geq\; 0 \quad \text{if} \quad x^e_t = \bar{x}e^{g(K_{t-1})} \tag{21}$$

$$-e^{-u}s(q_t;\epsilon_t) \;+\; \frac{1}{1+r}E[J_Z] \;\leq\; 0; = 0 \quad \text{if} \quad Z_t > 0, \tag{22}$$

$$-P_I + \frac{1}{1+r}E[J_K] \;\leq\; 0; = 0 \quad \text{if} \quad 0 < I_t < \kappa\, K_t, \tag{23}$$

$$-P_I + \frac{1}{1+r}E[J_K] \;\geq\; 0 \quad \text{if} \quad I_t = \kappa\, K_t. \tag{24}$$

It is worth noting that the optimality of the extensive margin and investment must be checked at both lower and upper boundaries. However, we only write the optimality condition for inventory at the lower boundary (zero). since the optimality condition for inventory at the upper boundary (sum of production output and inventory carried over) will never be chosen given the Inada condition on the inverse demand function. In particular we use an inverse demand function of the form: $s(q_t, \epsilon_t) = e_t/q_t^{\alpha}$, where $\alpha < 1$. We solve the HJB equation using projection methods as described in Judd (1999). Using the policy functions written in polynomial form, we can calculate expected future production in each state, and hence using the inverse demand function and the Markovian shocks, we can compute the forward prices as the expected value of the future spot price.

As a final comment to this section, it is useful to note that the two period welfare maximizing problem here has the same solution as the 2-period problem in the previous section, where we found the optimal policy of the firm using option pricing logic. This happens because, with only two periods, the firm's extraction decision in the second period is to simply extract for all plants where the extraction cost is lower than price, so that the future value of any plant is the value of a call option.

3 Explaining the Stylized Facts

In this section we examine the ability of the model in explaining the three stylized facts. At the outset, it is important to note that we are using a two period model to study the investment decisions in a long lived economy. Since call options have more value in a multi period setting, our model correlation between investment and the basis are understated. In future versions of the paper we will generalize the analysis in Section 1 to a multi-period problem. The advantage of using the two-period model is that its futures price is solved in closed-form and the intuition between investment, the basis, and the put value is simple and quite clear.

To study the model predictions we need to make several choices. First, we need to assume the functional form for demand shocks, and we use the linear specification in the previous section to facilitate calculations. Second, we must specify a functional form for the function

$g(K)$, which determines the growth rate of extraction costs. We use:

$$g(K) = \frac{\gamma}{K},$$

which implies that the growth of extraction costs explode as capital tends to zero so that positive capital is required to ensure the supply of the resource.

Finally, we need to specify the distribution of demand shocks to the economy specified. In part to obtain the low and high frequency components, we specify a regime structure for the state of the demand shock. We assume that the state of demand switches between two states with the transition matrix: Prob$(s_{t+1} = j | s_t = i) = p$, for $i \neq j$. The mean and volatility of shocks each shifts with the regime, so that $\mu \in \{\mu_1, \mu_2\}$ and $\sigma \in \{\sigma_1, \sigma_2\}$. In addition, we assume that he expected growth rate at each period has an additional component unrelated to the macro regime. In particular we assume that the growth rate at time t is $m(t) = \mu(t) + \sigma_u \tilde{u}$, where $\tilde{u}$ is an uniformly distributed random variable, which represents idiosyncratic and uncorrelated reasons for growth to diverge within an economic regime.

We next discuss a set or results for some parameter values assumed for the model. In future work, we will attempt to provide a more careful calibration of these parameters, in particular, using the properties of empirically observed demand shocks.

3.1 Parameter Values for Model

First we look at the quantity of reserves and the annual demand function. We set the parameter $a = 1$, which is the amount of oil to be consumed in a year if the price of oil would be zero. The meaning of this parameter is only relevant when comparing to reserves. We set $Q_0 = 10$, and $Q_1 = 3$, so that total reserves are 13. This implies that in our model the resource would be exhausted in 13 years if its price was zero. Of course the price is determined endogenously in the model, and in the equilibrium there is substantial probability that the resource will never be fully exhausted (see the proof of Proposition 1 for the optimal decision rule). The parameter $b = 0.1$, and this must be carefully compared to estimates of elasticity of demand for oil.

We assume that the riskless rate is 5%, and the parameter $\gamma = 0.1$. This implies that extraction costs grow at 10% if capital equals 1. We note that the second channel that determines the average basis in our model is determined by the difference between the growth of extraction costs and the riskless rate, and so we can control the average basis in the model by changing γ in a reasonable range. The only other parameter governing the capital process is the rate of capital depreciation, which we set at 10 percent a year.

Finally we need to assume some parameters for the regimes of demand. We assume that the probability or returning to the same regime in the following year is 0.9. We assume that the annual growth rate of consumption is 6 and -3 percent in the two macroeconomic regimes. Finally, we assume that volatility in the second regime is 3 times higher than in the second regime. For this reason we will call the two regimes 'stable' and 'volatile'. Since volatility increases call option value, there will be more investment in the more volatile regime. We note that in a multiperiod model of investment, the difference required would not be as large, and we will verify this in future versions of this paper.

3.2 Slow Asymmetric Capital Cycles in Model and Stylized Facts

As we will highlight in this section this model displays slow capital cycles that help us understand the low frequency variation in the relative basis and the other stylized facts. We start by showing a typical sample path of the model's relative basis and optimal capital in Figure 7. As the top panel shows, the relative basis is mostly in negative (in backwardation) and jumps for short spurts to being positive (in contango). These spikes give the model relative basis the low frequency variation that we see in the data.

The capital process in the bottom panel of Figure 7 shows why this model has this feature. The model has two macroeconomic regimes, stable and volatile, and optimal capital in the volatile regime is larger. In addition, the model has a built in irreversibility in capital accumulation, which leads to an interesting asymmetry. Capital can be increased fairly fast with new investment, but can only be reduced at the rate of capital depreciation. When the economy shifts from the volatile to the stable regime, investment actually stops. Capital slowly reduces to its optimal level in the stable regime and then it jumps again, when the economy shifts to the volatile regime again.

We now discuss the relation between investment and the relative basis. In the stable regime, the economy inherits a large capital stock, and there is no new investment for several periods.

However, the high capital stock implies that the futures market is in backwardation (see the discussion below Proposition 1. When the economy shifts back to the volatile regime, the immediate first impact is an increase in volatility that increases the value of put options and hence temporarily increases backwardation. As can be seen in the top panel of Figure 7, there is a dip in the futures basis, before it spikes. In the following periods, investment sharply increases and lowers extraction costs, thus lowering the value of puts (see the discussion on the relation between the basis and investment in the previous section), and causing a sharp contango. Initially, the price increase is not big, so that the relative contango is large as well. However, as more capital is installed so that future extraction costs are lowered, less of the resource is extracted currently, so that current prices increase, and the relative basis falls. Therefore, there is a temporary spike in the contango. With further capital accumulation, we return to a backwardation. Overall, the model implies a positive relation between investment and the relative basis, particularly in the period of the spike. The variables are not perfectly correlated, which is consistent with the first stylized fact in the introduction.

What does this model imply about roll returns and the relationship between roll returns and inventory? As seen in Table 4, returns on both the long futures and unconditional roll are almost zero. However, as in the data, investment is a useful conditioning variable for the roll strategy, and using it we obtain an average return of 6.5%, which is smaller than in the data, but still quite large. It is important to note that in this model agents are assumed to be risk neutral, and therefore, the returns cannot be meaningfully called a "risk premium". In fact, the simulations show that the model is able to produce a roll return after conditioning as described by Gorton and Rouwenhourst (2006) and Erb and Harvey (2006).

We next look at Figure 8, which shows the variance frequency decomposition of the model's relative basis. As can be seen, the decomposition is U-shaped, just as for the model in Figure 5. The slow capital adjustment process discussed above is consistent with this pattern. Indeed, there is only a rapid increase in investment when the economy has been in the volatile regime for a substantial period of time, and then switches back to the stable regime. In some cycles where the downturn is of a shorter duration, the investment boom is much smaller. In the stable regime, the small investment and movement in the basis due to the idiosyncratic shocks, gives the high frequency variation.

We finally look at the last stylized fact, which is the U-shaped relation between the relative basis and volatility. In the simulation, volatility is calculated using a simple GARCH (1,1) model, from the path of model spot prices. As can be seen in Figure 9, the model displays a similar pattern as in the data in Figure 6. As discussed above, there is large volatility in prices as the regime shifts from weak to stable, accompanied with a large relative contango. We also get some high volatility during stable states, which have backwardated futures markets. These occur in periods when there are high idiosyncratic (unrelated to the macroeconomic) shocks to demand. These shocks matter more in the stable regime, because they add to already higher macroeconomic volatility in this regime.

It is important to compare the reason for the U-shape relation between the basis and volatility in our model and that in Carlson, Khokher, and Titman (2007) and Kogan, Livdan, and Yaron (2009). In these papers, the U-shape arises because there are additional constraints on firms' production or capital adjustment process, in addition to the depreciation constraint that we have here. Why is our model able to get the relation without these additional constraints? In our model, the variation in the basis arises from the variation in the put option variation that we described above. These above papers, price the futures price of the resource above the ground without any price protection services for the resource owner. They hence do not make use of this additional important channel.

4 Conclusion

In this paper we present some evidence that the slope of the oil futures curve is more strongly related to firms' long term E&D expenditures on exploration and development rather than their inventory decisions. In addition we show that roll strategies in futures contracts conditioned on E&D expenditures rather than inventories have a stronger performance over the past 25 years. The relative performance improvement is line with the better predictive performance of the basis on the futures curve. We think that these result are novel and important as the information in inventories is strongly tied to the theory of storage, which has been very influential in academics and the financial community. We also see that historical data displays significant components of very low (less than once in six years) and very high (more than once every two months) frequency variation in prices, which supports the presence of both long and short term risk.

Building on the work of Litzenberger and Rabinowitz (1995), we develop a theoretical model where firms change E&D expenses over the business cycle to manage the value of their extraction options. Firms optimally invest in short bursts when aggregate resource demand shifts from a stable to a volatile regime, and when their capital stock is far from the new optimum level. Such adjustments happen infrequently only when the stable regime has persisted for a fairly long period. Overall, the model generates the positive covariation between the futures basis and E&D, the positive roll return, the low frequency variation in futures similar to the historical evidence, and the U-shaped relation between the futures basis and spot volatility. It is important to note that the roll return in our model arises despite not having a risk premium. Thus our model provides a rationale for roll trade returns not based on risk.

Data Appendix

We obtain historical crude oil futures contracts prices from July 1986 to December 2010 from the *Chicago Mercantile Exchange* (CME). The data series provided summarize the prices from all public traded exchanges. We obtain FOB WTI Cushing, Oklahoma spot prices of oil from 1986 - 2010 from *International Energy Agency* (IEA). We obtain the series of constant maturity Treasury yields from the *Federal Reserve Board*, which are required for calculating weak backwardation and contango. We define the weak basis on 1-year contracts as $e^{-r_t}F_t - S_t$, which is the discounted value of the one year forward price less the spot price.

We obtain data on oil and gas companies expenses on exploration and development from the *International Energy Agency* (IEA). Since these data are only available at an annual frequency, we also obtain the total capital raised in debt and equity markets at the quarterly frequency by US and Canadian companies in oil and gas field exploration services, SIC code 1382 from *Securities Data Company* (SDC). The time series of this series are in Figure 1

The annualized capital raised series from SDC have a 90 percent correlation with expenses series from the IEA. To ensure stationarity, we normalize the new capital series by US Gross Domestic Product obtained from the *National Income and Product Accounts*. We call this series new capital share (NCS), which we use in regressions. We obtain oil inventories (excluding strategic petroleum reserves) from the *United States Department of Energy*. In regressions we use the detrended inventorty to GDP ratio.

Appendix

Proof of Proposition 1. Given x^e, production at date 0 is $Q_0 = x_0^e/\bar{x} \cdot R_0$. Therefore the total amount of resource available for production at date 1 is $R_1 = R_0 - Q_0 = R_0\,(1 - \frac{x_0^e}{\bar{x}})$. At date 1, no further options to extract are available. Hence, the firm will extract the resource if $S_1 > x\,e^{g(K_1)}$. Since the extraction costs of all qualities increase at the same rate, $g(K_1)$, total production of the resource in period 1 equals

$$Q_1 = \int_{x_0^e\,e^{g(K_1)}}^{S_1} \frac{1}{\bar{x}\,e^{g(K_1)}}dx = \frac{1}{\bar{x}}\left(S_1\,e^{-g(K_1)} - x_0^e\right)\,R_0.$$

Now using the demand function at date 1, equilibrium entails that:

$$\frac{1}{\bar{x}}\left(S_1\,e^{-g(K_1)} - x_0^e\right)\,R_0 = a + \epsilon - b\,S_1.$$

Solving for S_1 we have

$$S_1 = \frac{a + \frac{x_0^e}{\bar{x}}\,R_0\, +\, \epsilon}{b + \frac{1}{\bar{x}}\,e^{-g(K_1)}\,R_0}.$$

The call option value is simply

$$\begin{aligned} C(x|x_0^e, K_1) &= e^{-r}\,E[\left(\frac{a + \frac{x_0^e}{\bar{x}}\,R_0\, +\, \epsilon}{b + \frac{1}{\bar{x}}\,e^{-g(K_1)}\,R_0} - x\,e^{g(K_1)}\right)^+] \\ &= \frac{e^{-r}}{D}E[\left(\epsilon - (D\,x\,e^{g(K_1)} - a - \frac{x_0^e}{\bar{x}}\,R_0)\right)^+] \\ &= \frac{e^{-r}}{D}\,\left[E[e^{\log(\tilde{u}_1)}|\tilde{u}_1 > \log(k)] - k\,\text{Prob}[\tilde{u}_1 > \log(k)]\right] \\ &= \frac{e^{-r}}{D}\,\left[e^{(\mu+0.5\,\sigma^2)}N(-d_1) - kN(-d_2)\right], \end{aligned}$$

as stated. We note that $\tilde{u}_1$ on the third line is a normal distribution variable with mean μ and volatility σ, while the fourth line uses the conditional expectation for log normal variables (see e.g. Proposition 2.29 in Nielsen (1999)). The proof for the put is similar. ■

Proof of Proposition 2. The proof is similar to the proof of Proposition 1, so we shall be brief. Total production of the resource in period 1 again equals

$$Q_1 = \int_{x_0^e}^{S_1} \frac{1}{\bar{x}\,e^{g(K_1)}}dx = \frac{1}{\bar{x}}\left(S_1\,e^{-g(K_1)} - x_0^e\right)\,R_0.$$

Now using the demand function at date 1, equilibrium entails that:

$$\frac{1}{\bar{x}}\left(S_1\, e^{-g(K_1)} - x_0^e\right)\, R_0 + Z_0\, e^{-u} = a + \epsilon_1 - b\, S_1.$$

Solving for S_1 we have

$$S_1 = \frac{a + \frac{x_0^e}{\bar{x}}\, R_0 - Z_0\, e^{-u} \, + \, \epsilon}{b + \frac{1}{\bar{x}}\, e^{-g(K_1)}\, R_0}.$$

Subsituting for R_0 and using steps similar to the proof of Proposition 1 completes the proof.

■

Table 1: What Explains the Futures Relative Basis For Oil and Natural Gas?

No.	α	β_1	β_2	R^2
Oil (1986:Q3 - 2010:Q4)				
One Quarter Lag:				
1.	-5.961	0.610		0.102
	[-4.0439]*	[1.864]*		
2.	-10.568		0.303	0.168
	[-5.713]*		[4.666]*	
3.	-9.862	0.433	0.258	0.226
	[7.519]*	[2.392]*	[3.703]*	
Four Quarter Lag:				
4.	-5.9212	0.144		0.005
	[-3.596]*	[0.395]		
5.	-10.437		0.311	0.155
	[-6.826]*		[4.674]*	
6.	-10.439	-0.000	0.310	0.155
	[7.454]*	[0.006]	[4.027]*	
Natural Gas (1994:Q1 - 2010:Q4)				
One Quarter Lag:				
7.	3.676	2.432		0.135
	[1.101]	[3.105]*		
8.	-12.067		1.246	0.222
	[-2.622]*		[3.563]*	
9.	-10.969	2.136	1.159	0.325
	[2.803]*	[3.124]*	[4.247]*	
Four Quarter Lag:				
10.	3.542	-1.749		0.067
	[0.939]	[-2.343]*		
11.	-8.285		0.996	0.117
	[-1.978]*		[2.754]*	
12.	-8.636	-1.786	1.008	0.187
	[1.883]	[2.271]*	[2.999]*	

We report the coefficients of the fitted regression:

$$\text{Relative Basis}(t) = \alpha + \beta_1 \text{ Inventory}(t{-}k) + \beta_2 \text{ New Capital Share}(t{-}k) + \epsilon(t),$$

for $k = 1, 4$. The weak relative basis (in basis points) on 1-year contracts in quarter t is $[e^{-r(t)}F(t) - S(t)]/S(t)$, where $F(t)$ is the 1-year futures prices at the beginning of each quarter and $S(t)$ is the spot price of WTI oil in Cushing, Oklahoma. The explanatory variables are the the inventory to GDP ratio or the New Capital Share (NCS), which is the ratio of new capital raised (debt and equity) by oil and gas firms in exploration and development to US GDP (in percent). Inventories of oil (ex-SPR) and natural gas obtained from the EIA and are time detrended. T-statistics are in parenthesis and are adjusted for heteroskedasticity and autocorrelation.

Table 2: Predicting Oil Roll Excess Returns with Economic Variables

No.	α	β_1	β_2	R^2
Oil (1986:Q3 - 2010:Q4)				
1.	2.629	-0.007		0.067
	[0.542]	[-2.189]*		
2.	15.787		-0.287	0.070
	[2.410]*		[-2.857]*	
3.	13.253	-0.006	-0.243	0.116
	[2.107]*	[-1.863]*	[-2.487]*	
Natural Gas (1994:Q1 - 2010:Q4)				
4.	1.914	-0.368		0.001
	[0.409]	[-0.220]		
5.	-11.740		-1.134	0.040
	[-1.810]		[-2.685]*	
6.	-11.821	-0.408	-1.137	0.041
	[-1.799]	[-0.265]	[-2.673]*	

We report the coefficients of the fitted regression:

$$\text{Roll Return}(t) = \alpha + \beta_1(t)\,\text{Inventory}(t-1) + \beta_2\,\text{New Capital Share}(t-1) + \epsilon(t)$$

The roll excess return is defined as:

$$\begin{aligned} \text{Roll Return}(t) &= -\left(\frac{S(t+4) - F(t)}{F(t)}\right) && \text{If} \quad F(t) > S(t) \\ &= \left(\frac{S(t+4) - F(t)}{F(t)}\right) && \text{If} \quad F(t) < S(t), \end{aligned}$$

where $F(t)$ is the 1-year futures prices at the beginning of each quarter and $S(t)$ is the spot price of WTI oil in Cushing, Oklahoma. The explanatory variables are the the inventory to GDP ratio or the New Capital Share (NCS), which is the ratio of new capital raised (debt and equity) by oil and gas firms in exploration and development to US GDP (in percent). Inventories of oil (ex-SPR) and natural gas obtained from the EIA and are time detrended. It is assumed that the futures holder has a fully collateralized position invested at the riskless rate as in Gorton and Rouwenhourst (2006). T-statistics are in parenthesis and are re:gorton, adjusted for heteroskedasticity and autocorrelation.

Table 3: Statistics of Alternative Rolling Strategies on Oil and Natural Gas Futures

Strategy	Mean	Sharpe Ratio	Skewness
Oil (1986:Q3 - 2010:Q4)			
Long Futures	0.068	0.074	-1.061
Unconditional Roll	0.034	0.095	0.026
Roll Conditioned on Inventory	0.041	0.433	0.059
Roll Conditioned on E&D Capital	0.151	0.461	0.434
Natural Gas (1994:Q1 - 2010:Q4)			
Long Futures	-0.004	-0.011	-0.678
Unconditional Roll	0.043	0.104	-1.116
Roll Conditioned on Inventory Capital	-0.103	0.213	-0.903
Roll Conditioned on E&D Capital	0.117	0.337	1.367

The "Unconditional Roll" strategy is the roll strategy in the footnote to Table 2. The roll return conditioned on a variable x is defined as:

$$\begin{aligned} \text{Roll Return}(t) &= -\left(\frac{S(t+4)-F(t)}{F(t)}\right) && \text{If} \quad x > \bar{x} \\ &= \left(\frac{S(t+4)-F(t)}{F(t)}\right) && \text{If} \quad x < \bar{x}, \end{aligned}$$

where $F(t)$ is the 1-year futures prices at the beginning of each quarter and $S(t)$ is the spot price of WTI oil in Cushing, Oklahoma. x is either the inventory to GDP ratio or the New Capital Share (NCS), which is the ratio of new capital raised (debt and equity) by oil and gas firms in exploration and development to US GDP (in percent). Inventories of oil (ex-SPR) and natural gas are obtained from the EIA. We deasonalize the series and time detrend the ratios. It is assumed that the futures holder has a fully collateralized position invested at the riskless rate as in Gorton and Rouwenhourst (2006).

Table 4: Statistics of Alternative Rolling Strategies for Model

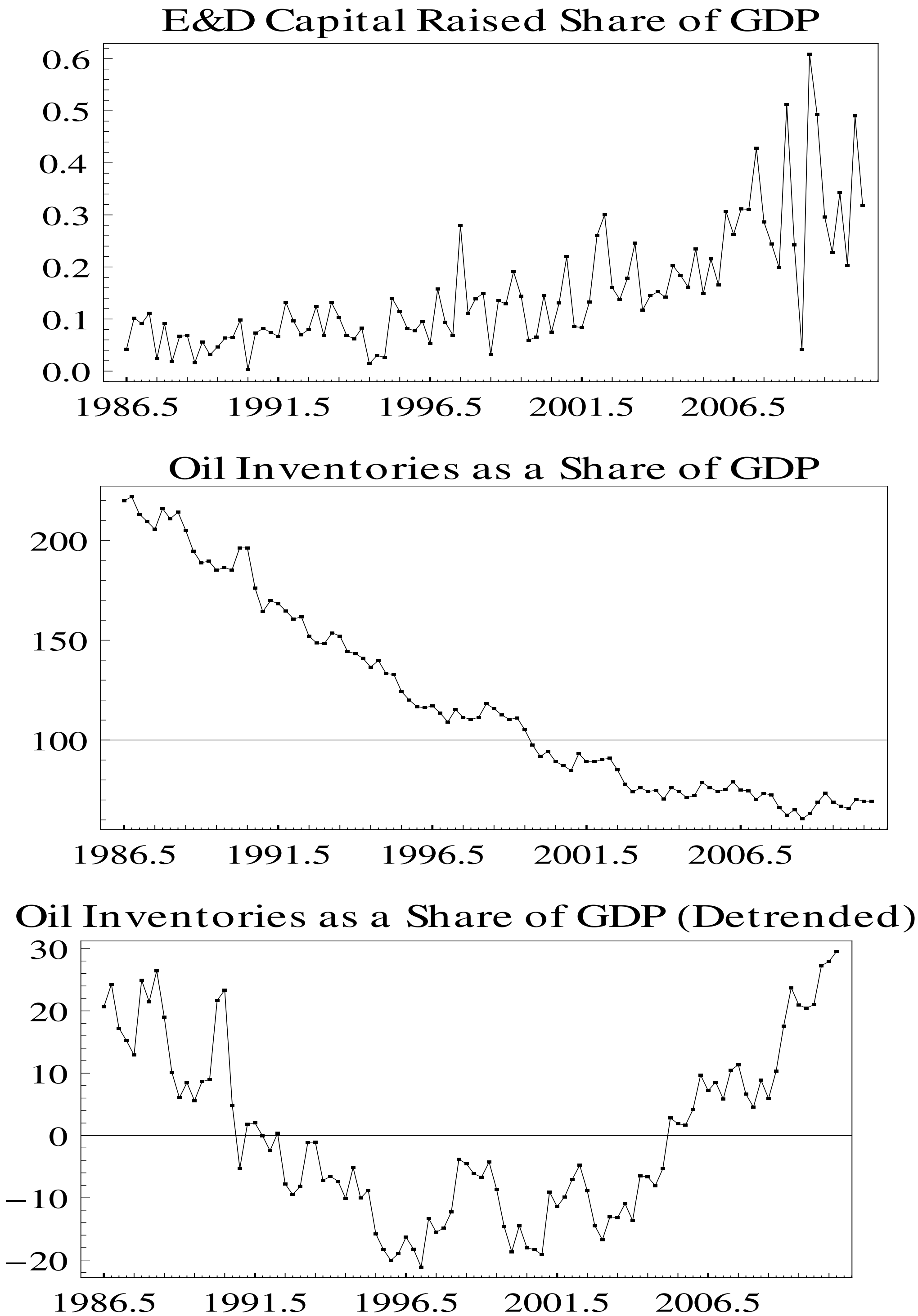
Strategy	Mean	Sharpe Ratio	Skewness
Long Futures	0.001	0.000	-0.55
Unconditional Roll	0.000	0.000	-0.67
Roll Conditioned on Investment	0.065	0.213	0.051

The "Unconditional Roll" strategy is the roll strategy in the footnote to Table 2. The roll return conditioned on a variable x is defined as:

$$\begin{aligned} \text{Roll Return}(t) &= -\left(\frac{S(t+4)-F(t)}{F(t)} - r(t)\right) && \text{If} \quad x > \bar{x} \\ &= \left(\frac{S(t+4)-F(t)}{F(t)} - r(t)\right) && \text{If} \quad x < \bar{x}, \end{aligned}$$

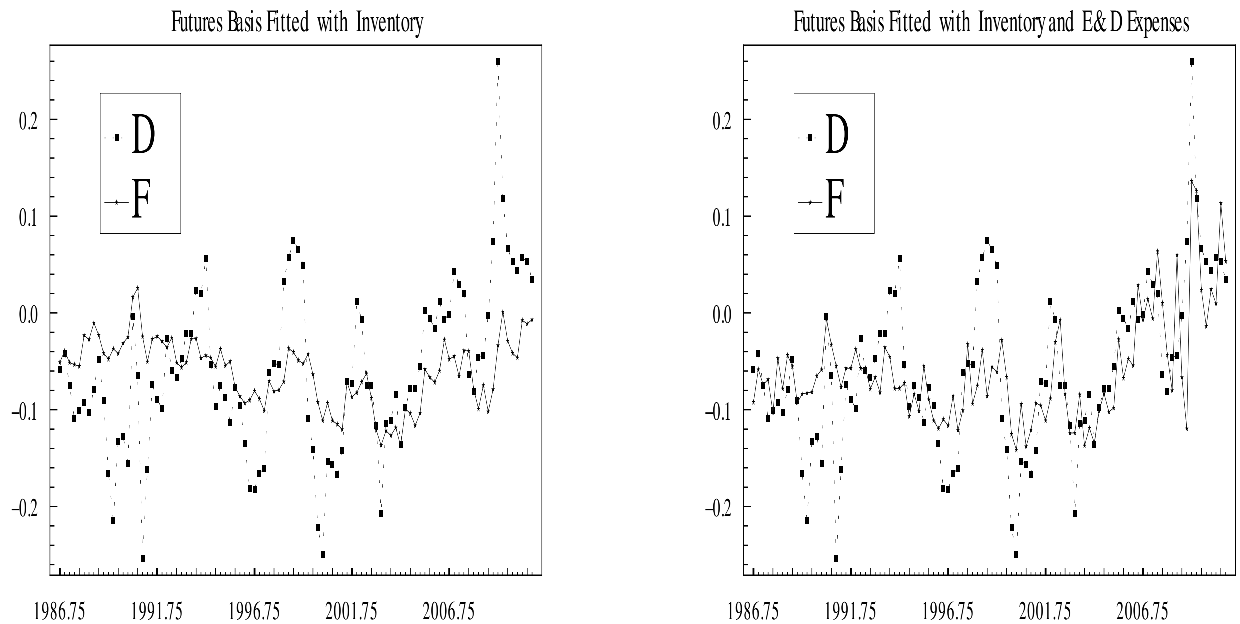
where x is the level of investment in the model.

Figure 1: New Capital of Oil and Gas Field Exploration Services as a share of GDP (1986:Q2-2010:Q4)



New capital share is defined as total capital raised in debt and equity markets at the quarterly frequency by US and Canadian companies in oil and gas field exploration services (SIC 1382) divided by US GDP. New capital is proceeds from debt and equity raised and are obtained from *Securities Data Company* (SDC).

Figure 2: One-Year Futures Market Weak Relative Basis Fitted with Inventory and Capital Raised By E&D Firms (1986:Q3-2010)

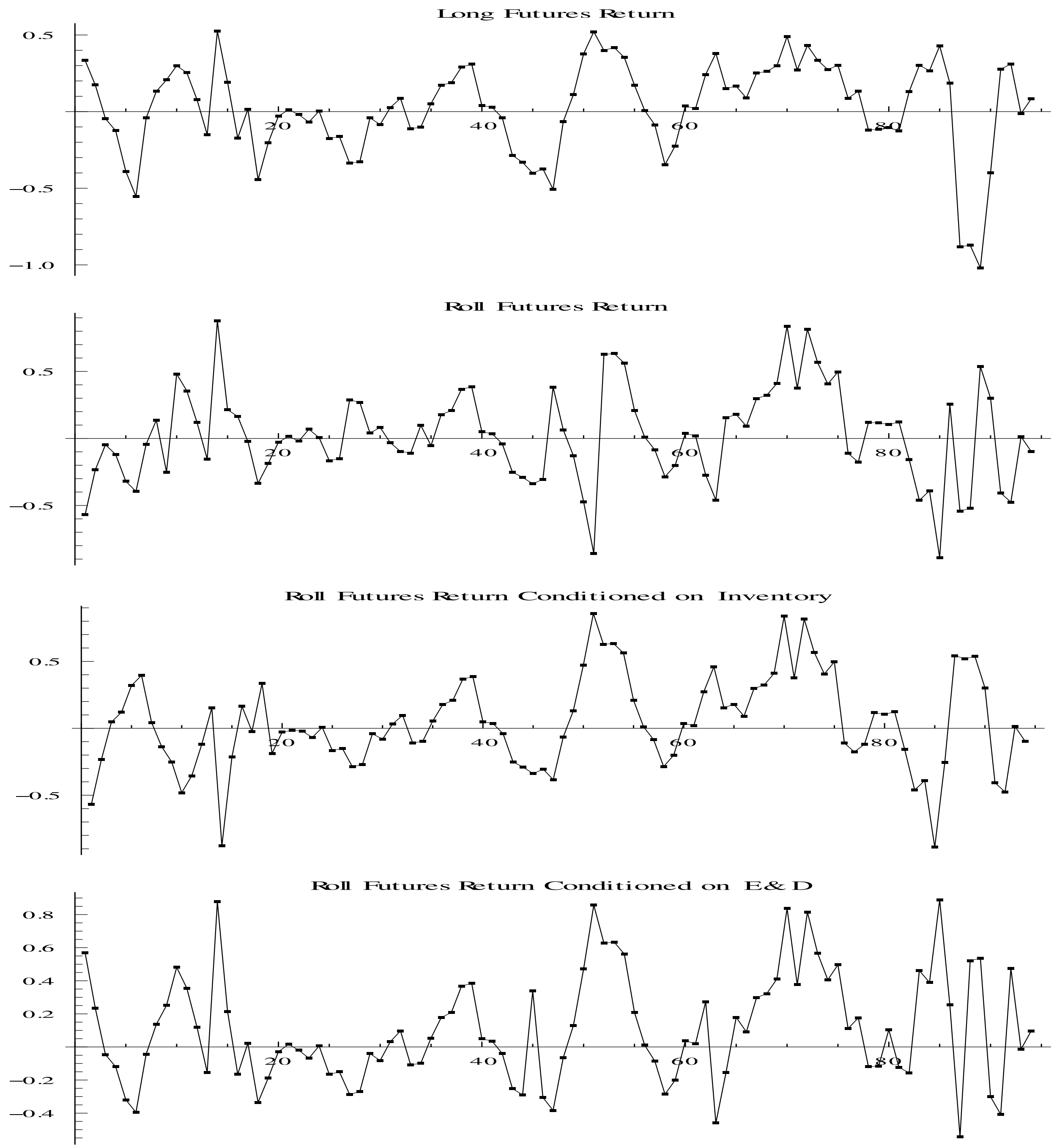


We report the actual weak relative basis and its fitted value from a simple linear regression:

$$\text{Relative Basis}(t) = \alpha + \beta_1 \text{Inventory}(t-1) + \beta_2 \text{New Capital Share}(t-1) + e(t)$$

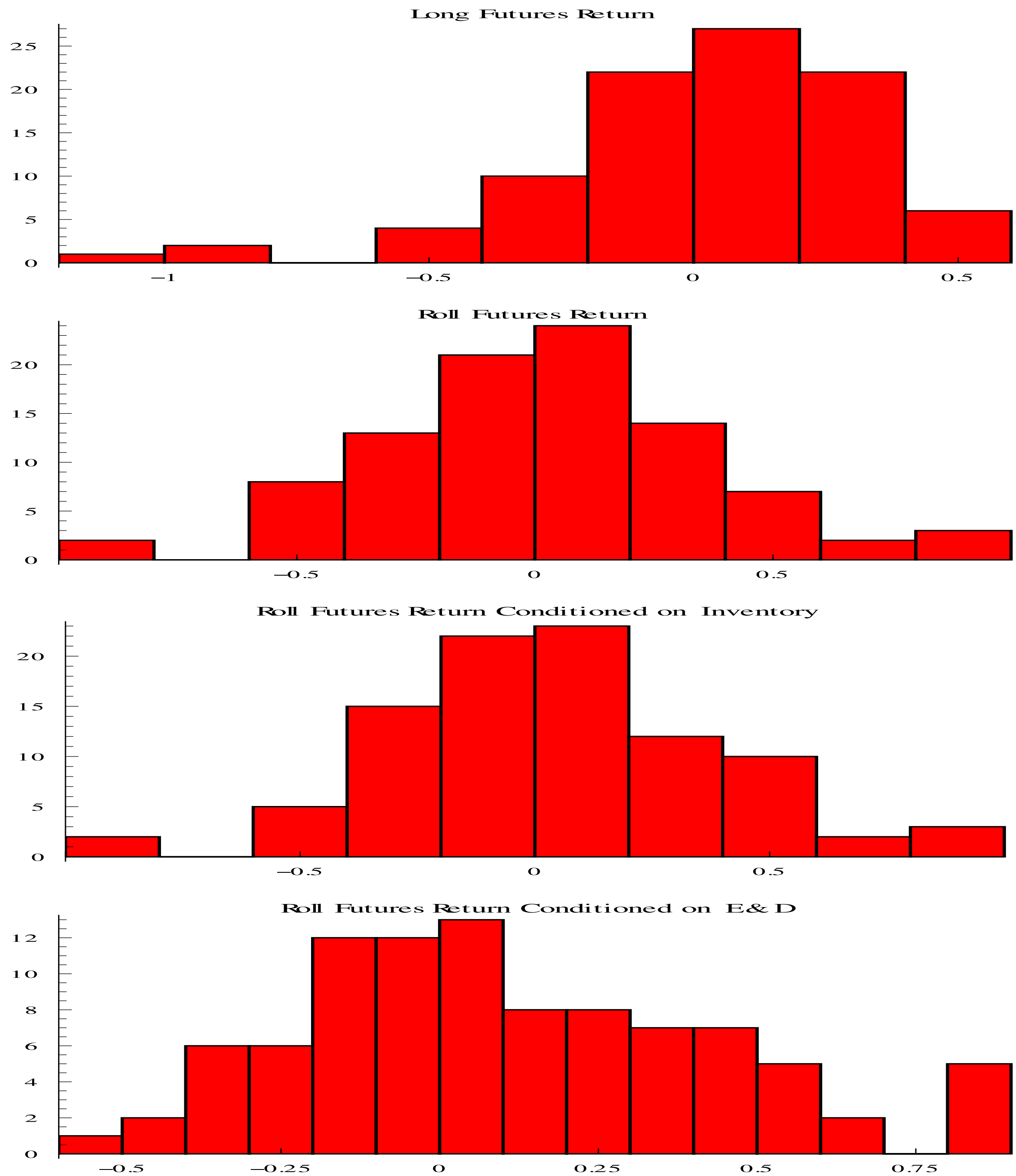
In the left panel, $\beta_2 = 0$. The results of these regressions are in Table 1. The weak relative basis on 1-year contracts in quarter t is $[e^{-r(t)}F(t) - S(t)]/S(t)$, where $F(t)$ is the 1-year futures prices at the beginning of each quarter and $S(t)$ is the spot price of WTI oil in Cushing, Oklahoma. Inventory is the total inventory of oil in the U.S. ex-SPR, and New Capital Share (NCS) is the ratio of new capital raised (debt and equity) by oil and gas firms in exploration and development to US GDP (in percent).

Figure 3: Roll Excess Returns on Oil Futures Under Alternative Assumptions (1986:Q3-2010)



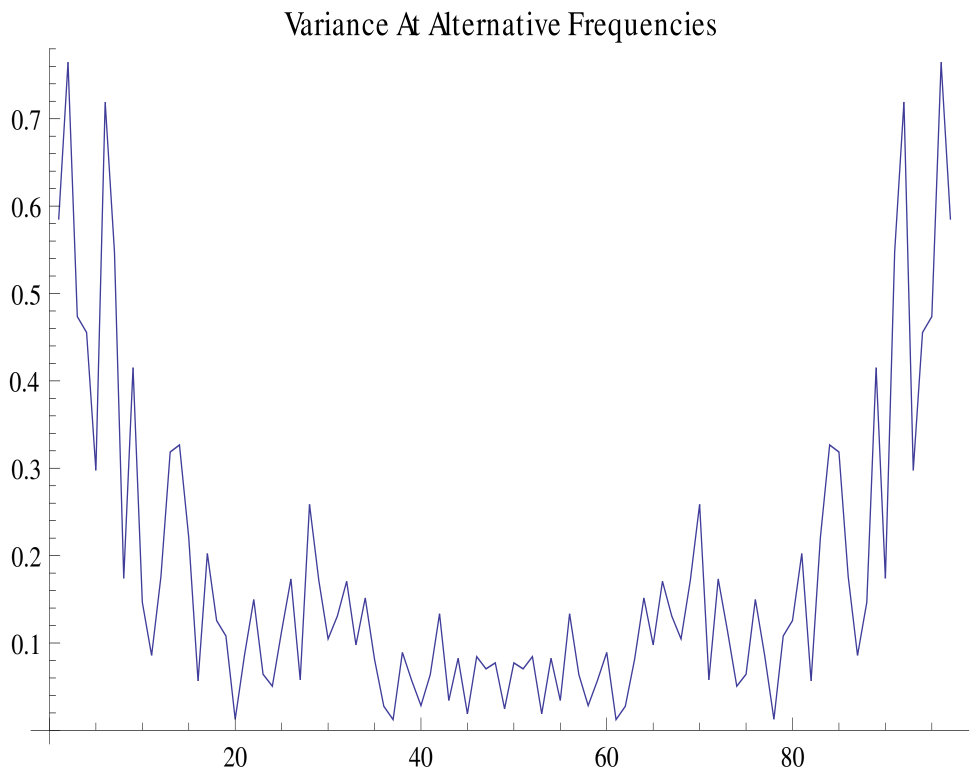
The excess returns are for the different series whose statistics and description are in Table 3.

Figure 4: Distribution of Roll Excess Returns on Oil Futures Under Alternative Assumptions (1986:Q3-2010)



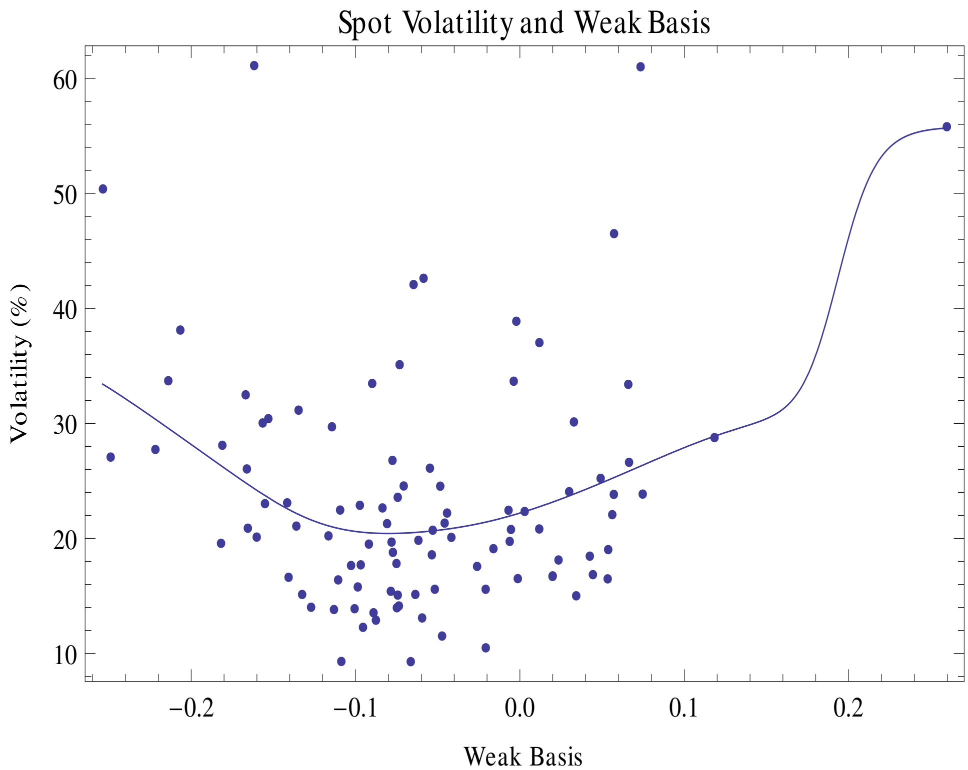
The excess returns are for the different series whose statistics and description are in Table 3.

Figure 5: Variance Frequency Decomposition of the WTI Weak Relative Basis (1986 -2010)



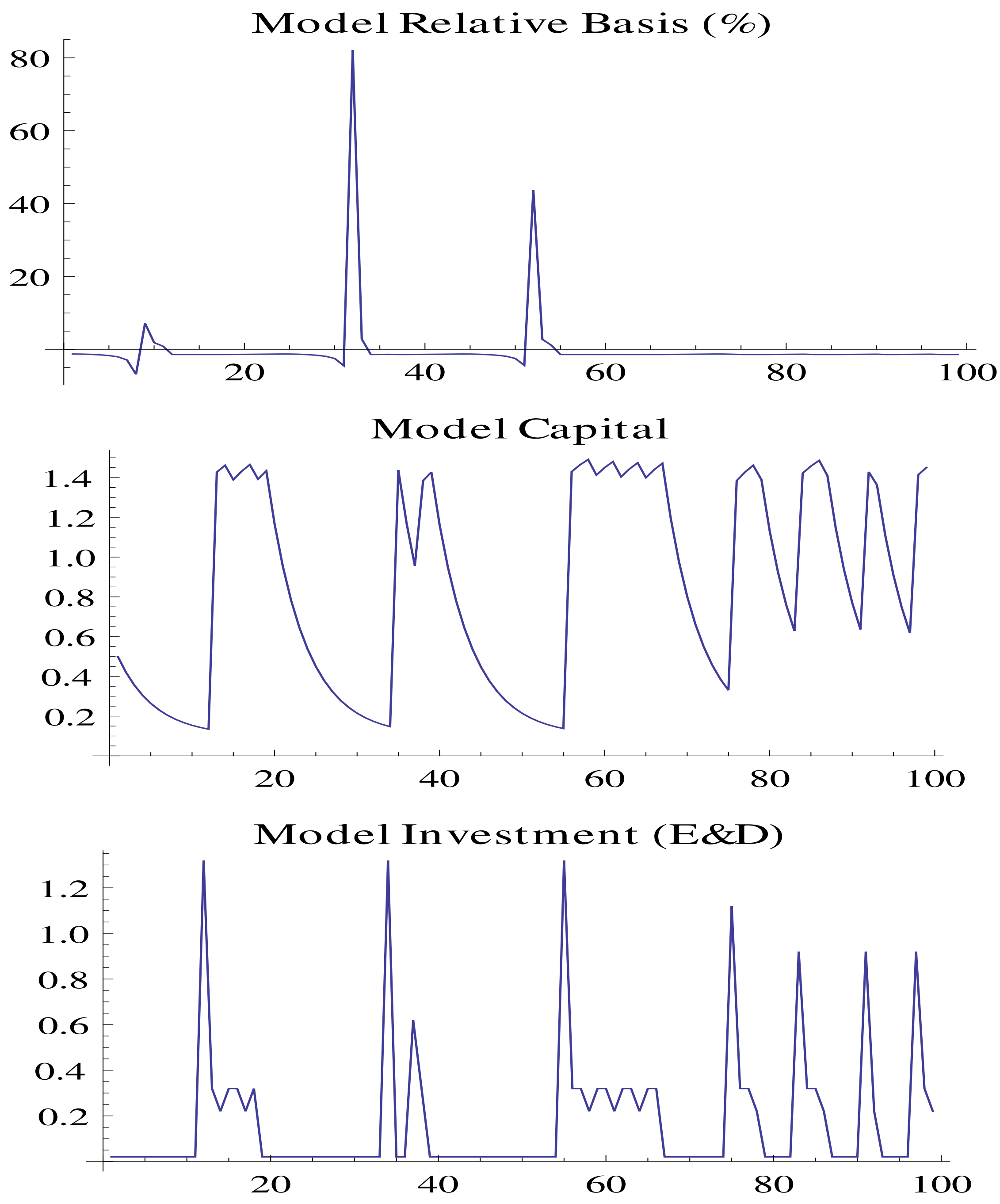
We report the variance frequency decomposition (Fourier Transform or spectrum) of the weak WTI relative basis $(e^{-r(t)}F_(t) - S(t)$, where $r(t)$ is the one year Treasury Bill rate, $F(t)$ is the one year futures price, and $S(t)$ is the spot price.

Figure 6: Relation Between Weak Relative Basis and WTI Spot Volatility (1986:Q3-2010)



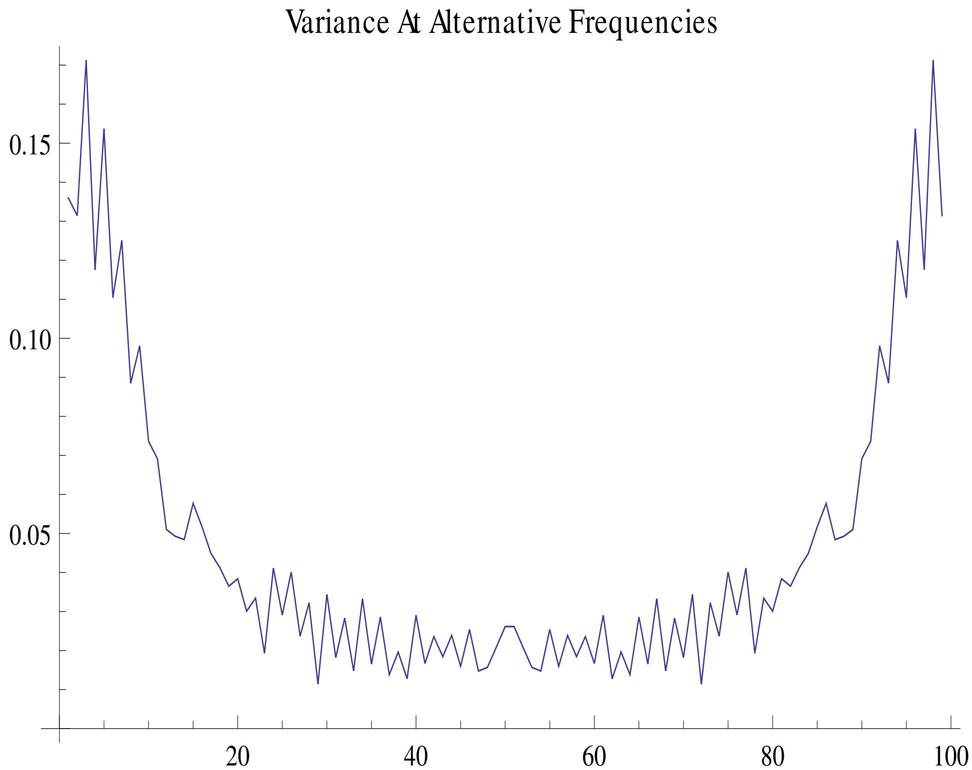
We report the fitted values from the nonparametric regressions of the historical monthly WTI spot realized volatility (constructed from daily returns) on the weak relative basis estimated with a Gaussian kernel (see e.g. Hardle (1990)). The weak relative basis is defined as $e^{-r(t)}F_(t) - S(t)$, where $r(t)$ is the one year Treasury Bill rate, $F(t)$ is the one year futures price, and $S(t)$ is the spot price.

Figure 7: Model Relative Basis, Capital, and Investment from simulation



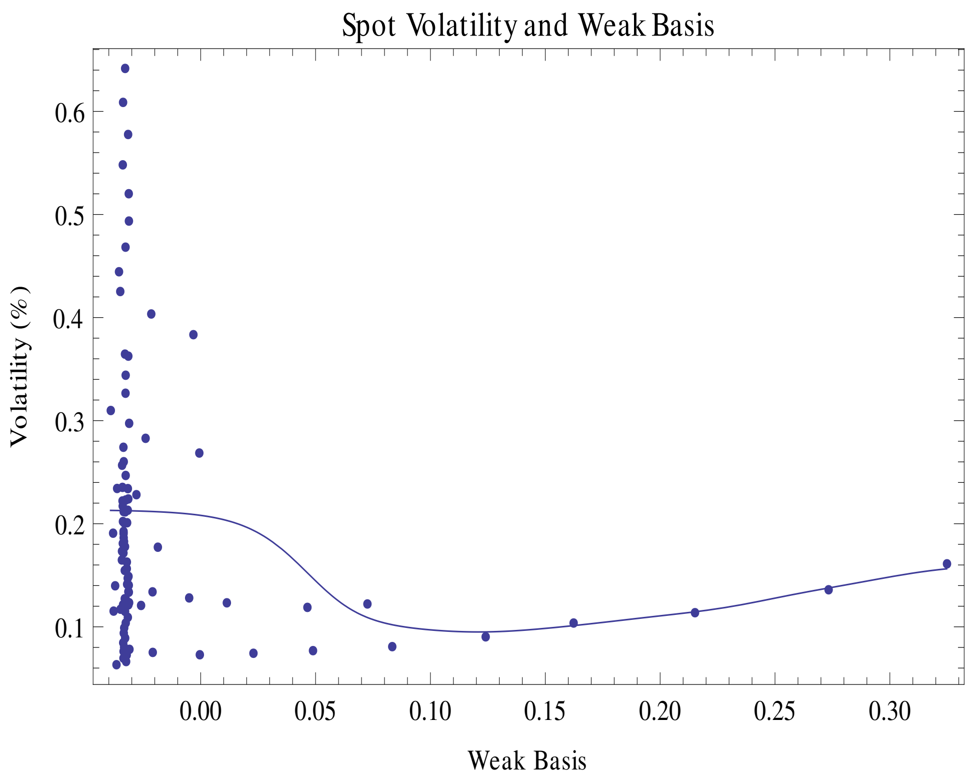
We plot the model relative basis and optimally chosen capital using Monte Carlo simulation using the parameters reported in Section 3.1.

Figure 8: Variance Frequency Decomposition of the Model Volatile Relative Basis (1986 - 2010)



We report the variance frequency decomposition (Fourier Transform or spectrum) of the weak WTI relative basis $(e^{-r(t)}F_(t) - S(t)$, where $r(t)$ is the one year Treasury Bill rate, $F(t)$ is the one year futures price, and $S(t)$ is the spot price.

Figure 9: Relation Between Volatile Relative Basis and Spot Volatility For Model



We report the fitted values from the nonparametric regressions of the model volatility (constructed using a simple GARCH(1,1) model) on the volatile relative basis estimated with a Gaussian kernel (see e.g. Hardle (1990)). The weak relative basis is defined as $\frac{e^{-r(t)}F(t)-S(t)}{S(t)}$, where $r(t)$ is the one year Treasury Bill rate, $F(t)$ is the one year futures price, and $S(t)$ is the spot price.

References

Baker, S. D., and B. R. Routledge, 2012, "The Price of Oil Risk," Working Paper, Carnegie Mellon University.

Brennan, M. J., and E. Schwartz, 1985, "Evaluating Natural Resource Investments," *Journal of Business*, 58, 135–157.

Carlson, M., Z. Khokher, and S. Titman, 2007, "Equilibrium Exhaustible Resource Price Dynamics," *Journal of Finance*, 62, 1663–1703.

Casassus, J., P. Collin-Dufresne, and B. R. Routledge, 2008, "Equilibrium Commodity Prices with Irreversible Investment and Non-Linear Technologies," Working Paper, Carnegie-Mellon University.

Deaton, A., and G. Laroque, 1992, "On the Behaviour of Commodity Prices," *Review of Economic Studies*, 59, 1–23.

Erb, C. B., and C. R. Harvey, 2006, *The Strategic and Tactical Value of Commodity Futures*62, 69–88.

Gorton, G., and K. G. Rouwenhourst, 2006, "Facts and Fantasies About Commodity Futures," *Financial Analysts Journal*, 62, 47–68.

Hamilton, J. D., 1994, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

Hardle, W., 1990, *Applied Nonparametric Regression*, Cambridge University Press.

Judd, K. L., 1999, *Numerical Methods in Economics*, The MIT Press,Cambridge, Massachusetts.

Kaldor, N., 1939, "Speculation and Economic Stability," *Review of Economic Studies*, 7, 1–27.

Kogan, L., D. Livdan, and A. Yaron, 2009, "Oil Futures Prices in a Production Economy with Investment Constraints," *Journal of Finance*, 64, 1345–1375.

Litzenberger, R. H., and N. Rabinowitz, 1995, "Backwardation in Oil Futures Markets: Theory and Empirical Evidence," *Journal of Finance*, 50, 1517–1545.

Nielsen, L. T., 1999, *Pricing and Hedging of Derivative Securities*, Oxford University Press.

Pindyck, R. S., 1980, "Uncertainty and Exhaustible Resource Markets," *Journal of Political Economy*, 88, 1203–1225.

Routledge, B. R., D. Seppi, and C. Spatt, 2000, "Equilibrium Forward Curves," *Journal of Finance*, 55, 1297–1338.

Scheinkman, J., and J. Schechtman, 1983, "A Simple Competitive Model With Production and Storage," *Review of Economic Studies*, 50, 427–441.

Weinstein, M. C., and R. J. Zeckhauser, 1975, "The Optimal Consumption of Depletable Natural Resources," *Quarterly Journal of Economics*, 92, 371–392.

Working, H., 1948, "Theory of an Inverse Carrying Charge in Futures Markets," *Journal of Farm Economics*, 30, 1–28.