Commodity Convenience Yields as an Option Profit

Robert Heinkel
Maureen E. Howe
John S. Hughes

The purpose of this article is to provide insight into the determinants of commodity convenience yields. Convenience yield refers to the inverse carrying charges or return implied when spot prices plus storage costs and interest charges exceed the futures prices, holding risk considerations aside. An analytic expression for the convenience yield is derived in a market setting where temporary shocks to demand (or supply) can arise between the time when a futures contract is written and when that contract delivers. In this setting there may be an advantage to a position in inventory in meeting unexpected demand. Such an advantage, if it exists, is reflected in the relationship between equilibrium spot and futures prices, and may partially explain a negative basis (or contango) in seasonal commodities.

The intuition that holding the commodity in inventory provides an option not available from a futures contract is well documented. Kaldor (1939) and Working (1948) discuss this intuition and it is restated in the more recent empirical work of Brennan (1986) and Fama and French (1987). These recent articles empirically examine the 'theory of storage'. Brennan (1986) tests several empirical convenience yield models and finds that convenience yield is negatively related to inventory levels. Fama and French (1987) find that seasonalities are significant in explaining agricultural commodities' futures basis. They suggest that this seasonality is due to inventory fluctuations.[1]

The model developed here formalizes the empirical fact that inventory and convenience yield are negatively related, and demonstrates two new factors that influ-

The authors would like to thank Rex Thompson, Margaret Slade, Alan Kraus, Josef Zechner, and, especially, Jim Brander for their many useful comments and suggestions.

[1]A recent article by Bresnahan and Spiller (1986) discusses the role of 'price shocks' on normal backwardation in futures markets. The model does not point out the option feature directly, nor does it draw out testable hypotheses, such as the relations of convenience yield to inventories or marginal production costs.

Robert Heinkel is an Associate Professor, Faculty of Commerce, at The University of British Columbia.

Maureen E. Howe is Senior Manager at Peat, Marwick and Thorne.

John S. Hughes is a Professor in the School of Business, University of Minnesota.

The Journal of Futures Markets, Vol. 10, No. 5, 519–533 (1990)

CCC 0270-7314/90/050519-15$04.00

ence convenience yield. These results follow from an explicit model of convenience yield based on uncertain commodity spot market demand. More specifically, the analysis uses a three-date (two-period) model in which identical price-taking firms endowed with inventory at time 0 and a production technology at time 1 seek to maximize expected profits. Spot commodities trade at all three dates while futures contracts trade at time 0 and mature at time 2.[2] A positive convenience yield results at time 0 when initial inventories are low, marginal costs of production are high and demand is not highly serially correlated. Intuitively, these conditions, accompanied by sufficiently high unexpected demand at time 1, serve to drive up the spot price at time 1 without a concurrent increase in the futures price (which in the model is equivalent to the expected spot price at time 2). An 'option' profit then results from having taken a position in the commodity rather than the futures contract.[3]

Performing comparative statics on the equilibrium model, the analysis shows that the convenience yield is decreasing in initial aggregate inventory, which is consistent with the earlier informal models. In the model used here, greater aggregate inventory implies lower equilibrium prices when it is sold. If inventory is sold at time 1, rather than carried forward to time 2, then the spot price at time 1 is reduced, while the expected spot price at time 2 is unaffected, thereby implying a smaller convenience yield.

The analysis provides additional empirical predictions. First, higher marginal costs of production imply a higher convenience yield. If marginal costs are very low, then unexpected demand at time 1 can be met from immediate production, implying no convenience yield to an inventory position in the commodity. Conversely, if marginal costs are very high, then demand can be met only from inventory, thereby implying a convenience yield. Finally, convenience yield is found to be decreasing in the serial correlation of the commodity spot price. A lower or negative correlation in spot price implies that an unusually high spot price prior to maturity of the futures contract (at time 1) is not expected to carry over to the spot price when the futures contract expires (at time 2); in this circumstance there is substantial advantage to holding the commodity rather than the futures contract, and this will be reflected in a large convenience yield.

I. THEORETICAL MODEL

Description of Economy

Consider a three-date economy (t_0, t_1, and t_2) in which the jth firm of J identical firms is endowed at t_0 with an inventory level I^j of some homogeneous commodity and a common technology for producing that commodity at t_1. Production of the

[2]A key aspect of the model is that there is a date between the creation of a futures contract and its maturity when uncertain demand is realized (time 1 in the model). Futures contracts written at time 0 and maturing at time 1 would provide no convenience yield in the model, since they would produce identical results at time 1 to holding the commodity.

[3]In this simple model, the timing of the intermediate demand shock is known, and this would be the maturity of a *European* convenience yield 'option.' If the timing of the intermediate date is uncertain, then the option would have to be *American*, and this American option would have a value exceeding a similar European option with any intermediate maturity.

commodity exhibits constant returns to scale implying linear costs in output for fixed factor prices.[4] Storage costs and interest rates are assumed to be zero.

The firm can sell the commodity at t_0, t_1, or t_2. Demand is determined exogenously and is affected by random variables η_1 and η_2 in periods one (t_0 to t_1) and two (t_1 to t_2), respectively. Demand in period two is related to that in period one by the function $\eta_2 = \rho\eta_1 + \varepsilon$, where $E(\varepsilon) = 0$, $E(\eta, \varepsilon) = 0$, $\text{Var}(\varepsilon) = \sigma^2 < \infty$ and $|\rho| < 1$. Firms have identical beliefs over $\eta_1 \varepsilon [\underline{\eta}_1, \bar{\eta}_1], 0 < \underline{\eta}_1 < \bar{\eta}_1$, and ε represented by the density functions $f(\eta_1)$ and $g(\varepsilon)$.

Let i_0^j denote inventory carried from t_0 to t_1, $i_1^j = i_1^j(\eta_1)$ denote inventory carried from t_1 to t_2 and $q_1^j = q_1^j(\eta_1)$ denote output at t_1, the last two being conditional on the realization of η_1. The output is produced at a cost of $C(q_1^j(\eta_1))$. All inventory carried from t_1 to t_2 must be sold at t_2. Unit sales of the commodity at t_0, t_1, and t_2 are therefore $y_0^j = I^j - i_0^j$, $y_1^j(\eta_1) = i_0^j + q_1^j(\eta_1) - i_1^j(\eta_1)$ and $y_2^j(\eta_1) = i_1^j(\eta_1)$, respectively. Prices $P_0(y_0)$, $P_1(\eta_1, y_1)$, and $P_2(\eta_2, y_2)$ are linear decreasing in aggregate units sold $y_t = \Sigma_J y_t^j$ and linear increasing in η_t, as applicable. The price functions $P_1(\cdot,\cdot)$ and $P_2(\cdot,\cdot)$ are identical.

Firms are expected profit maximizers. Each firm takes the units sold by other firms as fixed in making its decisions and acts as if its decisions have no effect on prices, i.e., they are price takers. Expected profit of firm j at t_0 parameterized by initial inventory level I^j, for a given decision i_0^j, policies of $q_1^j(\eta_1)$ and $i_1^j(\eta_1)$, and level of aggregate sales $\bar{y}_t$, t = 0,1,2, is denoted $\pi_0^j(I^j)$, where

$$\pi_0^j(I^j) = P_0(\bar{y}_0)y_0^j + E[P_1(\eta_1, \bar{y}_1)y_1^j(\eta_1) - C(q_1^j(\eta_1))] + E[P_2(\eta_2, \bar{y}_2)y_2^j(\eta_2)].$$

The decision problem of the jth firm may be stated as follows:

$$\begin{aligned} \text{Max}_{i_0^j, i_1^j(\eta_1), q_1^j(\eta_1)} \quad & \pi_0^j(I^j) \\ st: \quad & i_0^j \le I^j \\ & i_1^j(\eta_1) \le i_0^j + q_1^j(\eta_1) \\ & i_1^j(\eta_1) \ge 0 \\ & q_1^j(\eta_1) \ge 0 \qquad \text{for all } \eta_1 . \end{aligned}$$

The constraints ensure nonnegative inventories and production.

Characterization of Optimal Decisions

To characterize a solution for the economy as a whole, the analysis exploits the assumptions made earlier which allows aggregation across firms,[5] and invokes the backward induction principle of dynamic programming. Accordingly, the problem at t_1 stated in aggregates $i_t = \Sigma_J i_t^j$ and $q_1 = \Sigma_J q_1^j$ for given i_0 and η_1 is (with depen-

[4]This assumption simplifies the later analysis, but is not essential. Alternatively, one could assume that costs are locally linear, marginal costs are increasing, firms are divisible, and that there are no barriers to immediate entry. See the Appendix.

[5]It is evident that with linear costs and price taking behavior, the same characterization of an optimum in aggregates can be obtained either by solving the profit maximization problem for the representative firm and imposing market clearing conditions, or by summing objective functions and solving the industry welfare problem. See the Appendix for the analogous problem with increasing marginal costs and an endogneously determined number of firms.

dency of i_1 and q_1 on η_1 suppressed when inessential):

$$\begin{aligned} \mathrm{Max}_{i_1, q_1} \quad & \pi_1(\eta_1, i_0) \\ st: \quad & i_1 \leq i_0 + q_1, \\ & i_1 \geq 0, \\ & q_1(\eta_1) \geq 0 \qquad \text{for all } \eta_1, \end{aligned}$$

where $\pi_1(\eta_1, i_0) = P_1(\eta_1, \bar{y}_1)y_1 - C(q_1) + E(P_2(\eta_2, \bar{y}_2) \mid \eta_1)y_2$.

The Kuhn-Tucker conditions for an optimum are

$$\begin{aligned} & -P_1(\eta_1, \bar{y}_1) + E(P_2(\rho\eta_1 + \varepsilon, \bar{y}_2)) - \lambda_1 + \lambda_2 = 0, \\ & P_1(\eta_1, \bar{y}_1) - C'(q_1) + \lambda_1 + \lambda_3 = 0, \\ & \lambda_1(i_1 - i_0 - q_1) = 0, \\ & \lambda_2 i_1 = 0 \text{ and} \\ & \lambda_3 q_1 = 0, \end{aligned}$$

where λ_1, λ_2, and λ_3 are nonnegative scalars.

The above conditions imply the following special cases (allowing that in equilibrium prices will reflect optimal decisions, i.e., $\bar{y}_t = y_t^*$, and that marginal costs are constant, i.e., $C'(\cdot) = c$, a constant):

$$\begin{aligned} & \text{A.) If } \lambda_1 > 0, \lambda_2 = 0, \lambda_3 \geq 0, \text{ then} \\ & \quad P_1(\eta_1, y_1^*) < E(P_2(\rho\eta_1 + \varepsilon, y_2^*)) \leq c, \\ & \quad y_1^* = 0 \text{ and} \\ & \quad y_2^* = i_1^* = i_0 + q_1^*. \end{aligned}$$

In this case, all firms store all inventory carried from t_0 to t_1 and any output produced at t_1 until t_2 in expectation of an equilibrium price greater than the current equilibrium price.

$$\begin{aligned} & \text{B.) If } \lambda_1 = \lambda_2 = 0, \text{ and } \lambda_3 \geq 0, \text{ then} \\ & \quad P_1(\eta_1, y_1^*) = E(P_2(\rho\eta_1 + \varepsilon, y_2^*)) \leq c, \\ & \quad y_1^* = i_0 + q_1^* - i_1^* \geq 0 \text{ and} \\ & \quad y_2^* = i_1^* \geq 0. \end{aligned}$$

In this case, all firms are indifferent between selling the commodity at t_1 and selling at t_2.

$$\begin{aligned} & \text{C.) If } \lambda_1 = 0, \lambda_2 > 0 \text{ and } \lambda_3 > 0, \text{ then} \\ & \quad E(P_2(\rho\eta_1 + \varepsilon, y_2^*)) < P_1(\eta_1, y_1^*) \leq c, \\ & \quad y_1^* = i_0 + q_1^* \text{ and} \\ & \quad y_2^* = i_1^* = 0. \end{aligned}$$

In this case, all firms sell all inventory carried from t_0 to t_1 and any output produced at t_1 in expectation of an equilibrium price at t_2 less than the current equilibrium price.

Each of the above cases can be further divided into cases where $\lambda_3 > 0$ (no production) and $\lambda_3 = 0$ (positive production). Which of these cases is feasible depends on the parameterization of cost and pricing functions. Cases in which $\lambda_1 > 0$ and $\lambda_2 > 0$ are regarded to be infeasible in equilibrium since both constraints on inventories cannot be binding unless $i_0^* = 0$, in which event they are redundant. Also note that marginal cost is an upper bound on both the equilibrium price at t_1, and the expected equilibrium price at t_2.

Case C above is of particular interest because in any other case there is no value to an option to sell at t_1. While a convenience yield has yet to be defined within the model, it is clear intuitively that the equilibrium price at t_1 must exceed the expected equilibrium price at t_2 under an optimal strategy for some realizations of η_1 for there to be an advantage to holding inventory rather than a position in a contract which enables the holder to take delivery and sell at t_2.

To identify a region within the support of η_1 for which Case C will obtain for a given i_0, the analysis identifies a critical value η_1^ℓ such that

$$E(P_2(\rho\eta_1^\ell + \varepsilon,0)) = P_1(\eta_1^\ell, i_0) < c.$$

Assuming that such a critical value exists, and using the earlier assumptions:

$$\frac{\partial E(P_2(\rho\eta_1 + \varepsilon,0))}{\partial\eta_1} < \frac{\partial P_1(\eta_1, i_0)}{\partial\eta_1}$$

it follows that

$$P_1(\eta_1, i_0 + q_1^*(\eta_1)) > E(P_2(\rho\eta_1 + \varepsilon,0))$$

for at least some η_1. Thus, η_1^ℓ is a lower bound on the region corresponding to Case C. Note that η_1^ℓ is a function of i_0. Also, define the value of η_1^i by:

$$P_1(\eta_1^i,0) = E(P_2(\rho\eta_1^i + \varepsilon, i_0)).$$

Figure 1 depicts the special cases described above under the assumption that η_1

at least some η_1. Thus, η_1^ℓ is a lower bound on the region corresponding to Case
Note that η_1^ℓ is a function of i_0. Also, define the value of η_1^i by:

$$P_1(\eta_1^i,0) = E(P_2(\rho\eta_1^i + \varepsilon, i_0)).$$

Figure 1 depicts the special cases described above under the assumption that η_1

d η_2 are (a) uncorrelated ($\rho = 0$), (b) positively correlated ($\rho > 0$), and (c) nega-
ely correlated ($\rho < 0$). Other than panel (b) of Figure 1, the regions in which
ses A, B, and/or C occur are convex. However, Case B may arise in disjoint re-
ons corresponding to $\lambda_3 > 0$, $\lambda_1 = 0$ when prices are positively correlated since it
conceivable that demand at t_1 would imply sufficiently large expected demand at
to drive expected equilibrium prices at that date as high as marginal cost. In
rn, this implies an upper bound, denoted η_1^u, on the region corresponding to
se C defined by

$$E(P_2(\rho\eta_1^u + \varepsilon,0)) = c.$$

course, such a critical value does not exist for $\rho \le 0$.

Having solved the problem at t_1 for an arbitrary decision i_0 at t_0 and period one
lization η_1, the recursion feature of dynamic programming is used to construct
problem at t_0:

$$\begin{aligned} Max_{i_0} \quad & \pi_0(I) \\ st: \quad & i_0 \le I \\ & i_0 \ge 0 \end{aligned}$$

ere $\pi_0(I) = P_0(\bar{y}_0)(I - i_0) + E(\pi_1^*(\eta_1, i_0))$

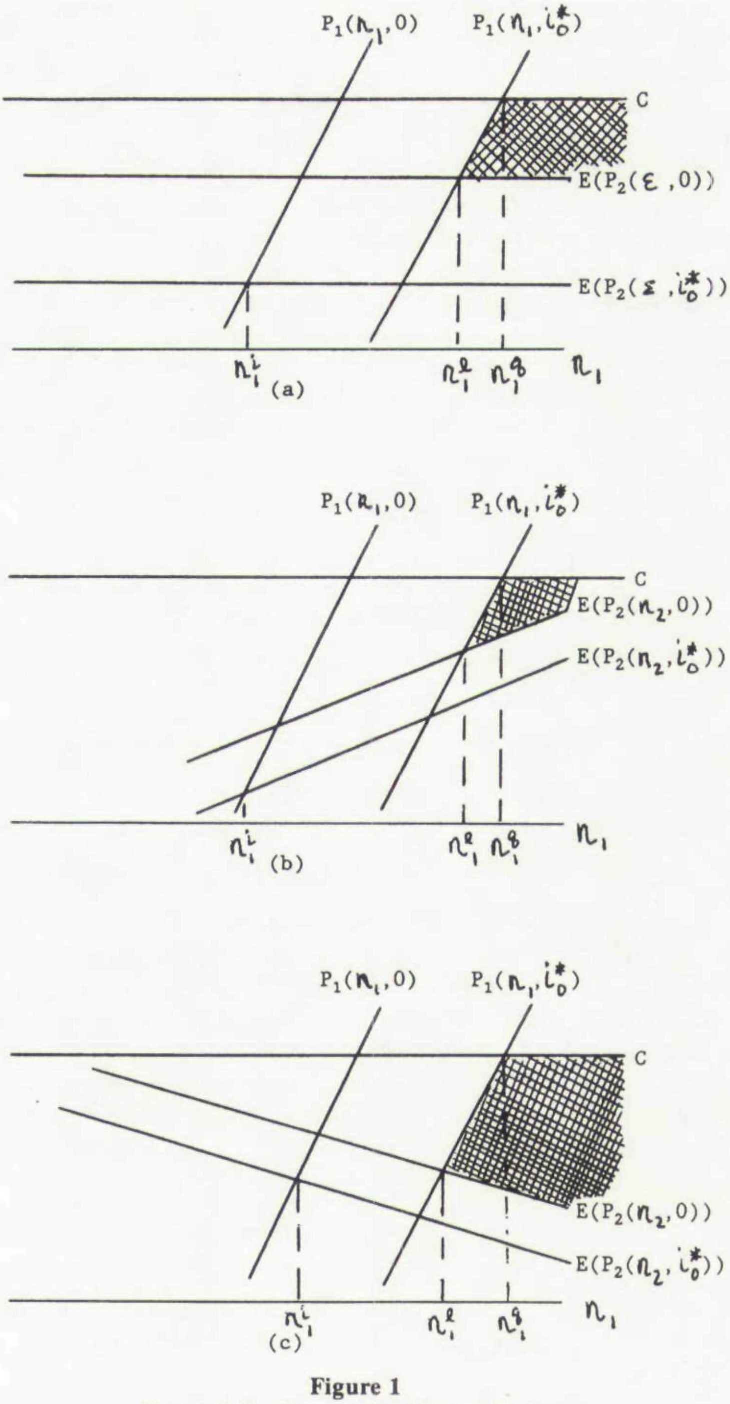

Figure 1
Demand Shocks and Equilibrium Prices.

Let η_1^ℓ denote the lower bound on the region containing Case C, conditional on i_0, and $\bar{y}_t$ denote the (fixed) units to be sold at t conditional on η_t:

$$E(\pi_1^*(\eta_1, i_0)) = \int_{\underline{\eta}_1}^{\eta_1^\ell} [E(P_2(\rho\eta_1 + \varepsilon, \bar{y}_2))(i_0 + q_1^*) - C(q_1^*)]f(\eta_1)d\eta_1$$
$$+ \int_{\eta_1^\ell}^{\eta_1^u} [P_1(\eta_1, \bar{y}_1)(i_0 + q_1^*) - C(q_1^*)]f(\eta_1)d\eta_1$$
$$+ \int_{\eta_1^u}^{\bar{\eta}_1} [E(P_2(\rho\eta_1 + \varepsilon, \bar{y}_2))(i_0 + q_1^*) - C(q_1^*)]f(\eta_1)d\eta_1,$$

exploiting the equilibrium conditions at t_1 which require that the price at time t_1 equal the expected price at t_2 when η_1 is in $[\eta_1^i, \eta_1^\ell]$.[6]

Forming the Lagrangean function and differentiating with respect to i_0 leads to the following condition:

$$-P_0(\bar{y}_0) + \frac{\partial\eta_1^\ell}{\partial i_0}[E(P_2(\rho\eta_1^\ell + \varepsilon, \bar{y}_2)) - (P_1(\eta_1^\ell, \bar{y}_1))]f(\eta_1^l)$$
$$+ \left\{\int_{\underline{\eta}_1}^{\eta_1^\ell} E(P_2(\rho\eta_1 + \varepsilon, \bar{y}_2))f(\eta_1)d\eta_1\right.$$
$$+ \int_{\eta_1^\ell}^{\eta_1^u} P_1(\eta_1, \bar{y}_1)f(\eta_1)d\eta_1$$
$$\left. + \int_{\eta_1^u}^{\bar{\eta}_1} E(P_2(\rho\eta_1 + \varepsilon, \bar{y}_2))f(\eta_1)d\eta_1\right\} - \lambda_1 + \lambda_2 = 0$$

where λ_1 and λ_2 are the Lagrangean multipliers that correspond to the first and second constraints, respectively.

By definition of η_1^ℓ, in equilibrium

$$E(P_2(\rho\eta_1^\ell + \varepsilon, y_2^*)) - P_1(\eta_1^\ell, y_1^*) = E(P_2(\rho\eta_1^\ell + \varepsilon, 0)) - P_1(\eta_1^\ell, i_0 + q_1^*) = 0,$$

and hence the second term in the above expression disappears. The terms in braces $\{\cdot\}$ can be viewed as an expected future price, given optimal decisions at t_1. Denoting this price by $\bar{P}(i_0^*)$, then, similar to before, there are three cases to consider:

A'.) $P_0(y_0^*) < \bar{P}(i_0^*)$,
$i_0^* = I$ so that $y_0^* = 0$.

B'.) $P_0(y_0^*) = \bar{P}(i_0^*)$,
$0 < i_0^* < I$ so that $0 < y_0^* < I$.

C'.) $P_0(y_0^*) > \bar{P}(i_0^*)$,
$i_0^* = 0$ so that $y_0^* = I$.

[6] While not crucial to the analysis, the first two integrals above generally contain nondifferentiable points corresponding to the critical values which delineate the special cases where inventory and/or production is sold at time t_1. However, given the way these values (limits of integration) are defined, it is not important to the analysis which follows to embellish the expression provided. Only in the Appendix where it becomes necessary to differentiate under the integral sign is this considered.

For exposition the analysis is restricted to Case B', although this does not rule out the other two. Accordingly, it is assumed that i_0^* is defined by

$$P_0(y_0^*) = \int_{\underline{\eta}_1}^{\eta_1^\ell} E(P_2(\rho\eta_1 + \varepsilon, y_2^*))f(\eta_1)d\eta_1 + \int_{\eta_1^\ell}^{\eta_1^u} P_1(\eta_1, y_1^*)f(\eta_1)d\eta_1 + \int_{\eta_1^u}^{\bar{\eta}_1} E(P_2(\rho\eta_1 + \varepsilon, y_2^*))f(\eta_1)d\eta_1 \quad (1)$$

Analysis of Convenience Yield

It is assumed that forward contracts can be written at time t_0 wherein the parties agree to buy or sell the commodity at time t_2, at a price $F_{0,2}(i_0^*)$, referred to as the futures price. The analysis uses a forward contract rather than a futures contract in order to eliminate the added complexity of "marking to market" at t_1. This does not impair the results in a qualitative sense since even with a settling up at t_1 there will be an expected value to a position in inventories which cannot be duplicated by a position in either a forward or futures contract.[7]

Recalling that storage costs and interest rates are zero, the convenience yield, CY, can be expressed simply as the spread between the spot price at t_0 and the futures price:

$$CY = P_0(y_0^*) - F_{0,2}(i_0^*)$$

Given risk neutrality, the futures price at t_0 will equal the expected spot price at t_2:

$$F_{0,2}(i_0^*) = \int_{\underline{\eta}_1}^{\bar{\eta}_1} E(P_2(\rho\eta_1 + \varepsilon, y_2^*))f(\eta_1)d\eta_1$$

By substitution from (1),

$$\begin{aligned} CY &= \int_{\eta_1^\ell}^{\eta_1^u} [P_1(\eta_1, y_1^*) - E(P_2(\rho\eta_1 + \varepsilon, y_2^*))]f(\eta_1)d\eta_1 \\ &= \int_{\eta_1^\ell}^{\eta_1^u} [P_1(\eta_1, i_0^* + q_1^*) - E(P_2(\rho\eta_1 + \varepsilon, 0))]f(\eta_1)d\eta_1 > 0. \end{aligned}$$

The inequality follows from the observation that the analysis involves integrating over the region in η_1 for which $P_1(\eta_1, y_1^*) > E(P_2(\rho\eta_1 + \varepsilon, y_2^*))$, i.e., Case C. The relevant regions are shaded in Figure 1.

Intuitively, it can be seen that the value of holding inventory (and rights to production) over a long position in a forward contract is analogous to the value of a call option written at t_0 and maturing at t_1 with a stochastic exercise price, i.e., opportunity cost of not selling at t_2 which depends upon the demand in period one. More specifically, this option can be identified by the boundary condition $\text{Max}\{0, P_1(\eta_1, i_0^* + q_1^*) - E(P_2(\rho\eta_1 + \varepsilon, 0))\}$.

Modelling convenience yield in this way allows a prediction consistent with that of earlier models (Proposition 1), but also allows more specific implications that are subject to cross-sectional analysis. Namely, besides predicting that convenience yield and aggregate inventory are negatively related, the convenience yield can be related to marginal production costs (Proposition 2) and to the degree of serial correlation in spot prices (Proposition 3).

[7] The necessary assumption is that there be some point in time prior to maturity of the forward contract (time t_1 here) such that there might exist excess demand for the commodity at that time.

Proposition 1: Convenience yield is decreasing in the level of aggregate inventory, I.
Proof: See the Appendix.
This result obtains since a larger aggregate inventory reduces the likelihood of excess demand for the commodity prior to maturity of the forward contract, thereby reducing the profit advantage to holding the commodity.

As indicated by the model, allowing costly production at the intermediate date can reduce the advantage of holding the commodity over the forward contract, and Proposition 2 states that higher marginal production costs raise the convenience yield.

Proposition 2: Under reasonable assumptions on the pricing function, P_t, and the range of η_1, and assuming that $\bar{\eta}_1 < \eta_1^u$, convenience yield is increasing in marginal production costs.
Proof: See the Appendix.

Finally, the three-date nature of the model allows a prediction on a third cross-sectional determinant of convenience yield that has not been pointed out previously: autocorrelation in commodity demand (and price) shocks, (ρ). First, the lemma states that reducing the correlation of demand shocks (lowering ρ) implies a less autocorrelated equilibrium commodity price.

Lemma: Reducing the autocorrelation of the demand shock, ρ, lowers the autocorrelation of the equilibrium commodity price.
Proof: See the Appendix.

Thus, given optimal inventory and production decisions, commodity price autocorrelation is reduced (or made more negative) when demand shocks are more negatively correlated.

The lemma establishes that changing ρ changes equilibrium commodity price autocorrelation. Proposition 3 can then be stated.

Proposition 3: Under reasonable assumptions on the pricing function, P_t, and the range of η_1, convenience yield is decreasing in the autocorrelation of equilibrium commodity price.
Proof: See the Appendix.

The intuition for this result is apparent in Figure 1, where the shaded area in panel (c), with $\rho < 0$ is greater than the area in panel (b), with $\rho > 0$. The result is apparent also in the characterization of the convenience yield payoff as an option with a stochastic exercise price; if spot price is negatively serially correlated, then when P_1 is high, $E(P_2)$ is low, yielding a larger option payoff.

II. CONCLUSIONS

A simple three-date model is used to formalize the intuition that there may exist a value to holding inventory that cannot be duplicated by a long position in a forward contract.[8] Given a positive probability of sufficiently high demand in the first period and less than perfectly positively autocorrelated spot prices, it is possible for the price at the intermediate date to exceed the expected price at the ending date thereby providing a profit to intermediate date sales. Holding the commodity rather than the futures contract provides an option maturing at the intermediate date with a payoff of $\max\{0, P_1 - E(P_2)\}$.

[8] The model is kept simple for analytic tractability and ease of exposition. Extending the model to nonlinear costs, shown in the Appendix, makes the comparative statics more complex to no obvious advantage, and extending it to allow production at time 2 does not affect the results.

The model confirms the intuition that higher inventory levels imply a smaller convenience yield. This has been empirically verified, for example, by Brennan (1986), who finds high convenience yields for low inventories and near zero convenience yields for high inventories of precious metals. Fama and French (1987) find seasonalities in convenience yields, which can also be linked to inventory fluctuations.

More importantly, the model provides two additional, and as yet untested, hypotheses, namely that the convenience yield is increasing in marginal production costs and decreasing in spot price serial correlation. To test the inverse association between convenience yield and spot price serial correlation one might compare convenience yields across commodities which vary in estimates of spot price serial correlation. For example, higher serial correlations for precious metals than for agriculturals, all else held constant, would be consistent with lower convenience yields in the precious metals. Similar cross-commodity comparisons might be made with respect to the positive association predicted between convenience yields and marginal costs of production. Moreover, within seasonal agriculturals, higher preharvest quarter convenience yields would be consistent with higher (essentially infinite) marginal costs of production in that quarter.

A simplifying assumption used in the model is to assume zero interest rates and zero storage charges. Adding these features to the model would, of course, reduce the convenience yield by raising the cost of holding the commodity rather than the futures contract. However, the comparative statics results that relate convenience yield to inventory, production costs and spot price serial correlation are unaffected.

An important limitation of the model is the restriction to a single cycle from when inventory is acquired, to when it is entirely sold. Other limitations include the use of forward contracts in place of futures contracts, and the lack of attention given to risk in characterizing the relationship between forward prices and expected future spot prices. Nonetheless, the intuition made precise by the model seems quite strong and likely to be robust across refinements to incorporate more realistic assumptions. Additionally, the role of spot price serial correlation in determining convenience yield is certainly empirically testable.

Bibliography

Brennan, M. J. (1958): "The Supply of Storage," *American Economic Review,* 48:50–72.

Brennan, M. J. (1986): "The Cost of Convenience and the Pricing of Commodity Contingent Claims," University of British Columbia Working Paper, March.

Bresnahan, T. F., and P.T. Spiller, (1986): "Futures Market Backwardation Under Risk Neutrality," *Economic Inquiry,* 24:429–441.

Fama, E., and K. French, (1987): "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage," *Journal of Business,* 60:55–73.

Kaldor, N. (1939): "Speculation and Economic Stability," *Reveiw of Economic Studies,* 7:1–27.

Telser, L. G. (1958): "Futures Trading and the Storage of Cotton and Wheat" in *Selected Writings on Futures Markets,* A. E. Peck, Ed., Chicago, 1977.

Working, H. (1948): "Theory of Inverse Carrying Charges in Futures Markets," *Journal of Farm Economics,* 30:1–28.

Appendix

Model with Nonlinear Costs. Suppose marginal costs are increasing and that for each η_1 an optimal solution at t_1 always involves positive production. Assume further that the number of firms J is no longer a constant, but is determined (endogenously) by dividing aggregate industry output by the efficient output of the representative firm; i.e., $J = q_1/q_1^j$. The problem at t_1 is then

$$Max_{i_1,q_1} P_1(\eta_1, \bar{y}_1)(i_0 + q_1 - i_1) - q_1 \frac{C(\bar{q}_1^j)}{\bar{q}_1^j} + E(P_2(\rho\eta_1 + \varepsilon, \bar{y}_2))i_1$$

$$st \qquad i_1 < i_0 + q_1,$$

$$i_1 > 0.$$

The first-order conditions imply three special cases analogous to those with constant marginal costs:

Case A.)

$$P_1(\eta_1,0) < E(P_2(\rho\eta_1 + \varepsilon, i_0 + q_1^*(\eta_1)) = \frac{C(q_1^{j*}(\eta_1))}{q_1^{j*}(\eta_1)} = C'(q_1^{j*}(\eta_1)).$$

Case B.)

$$E(P_2(\rho\eta_1 + \varepsilon, i_1^*(\eta_1))) = P_1(\eta_1, i_0 + q_1^*(\eta_1) - i_1^*(\eta_1)) = C'(q_1^{j*}1(\eta_1)).$$

Case C.)

$$E(P_2(\rho\eta_1 + \varepsilon,0)) < P_1(\eta_1, i_0 + q_1^*(\eta_1)) = C'(q_1^{j*}(\eta_1)).$$

Again the existence of a critical value η_1^ℓ such that

$$E(P_2(\rho\eta_1^\ell + \varepsilon,0)) = P_1(\eta_1^\ell, i_0 + q_1^*(\eta_1^\ell)) = C'(q_1^{j*}(\eta_1^\ell))$$

is sufficient to ensure the existence of a nontrivial region within which Case C will obtain.

Proof of Proposition 1

It is shown that

$$\frac{dCY}{dI} < 0,$$

where

$$CY = \int_{\eta_1^\ell}^{\eta_1^u} [P_1(\eta_1, i_0^* + q_1^*) - E(P_2(\rho\eta_1 + \varepsilon,0))]f(\eta_1)d\eta_1.$$

Since CY depends on I only through i_0^*, then

$$\frac{dCY}{dI} = \frac{dCY}{di_0^*}\frac{di_0^*}{dI}.$$

The proof is completed in two parts by showing that

$$\frac{di_0^*}{dI} > 0 \textit{ and } \frac{dCY}{di_0^*} < 0.$$

By assumption, in equilibrium, $y_0^* = I - i_0^*$ and

$$P_0(y_0^*) = \overline{P}(i_0^*)$$

where, with equilibrium values:

$$\begin{aligned}\overline{P}(i_0^*)) = &\int_{\underline{\eta}_1}^{\eta_1^\ell} E(P_2(\rho\eta_1 + \varepsilon, i_0^*))f(\eta_1)d\eta_1 \\ &+ \int_{\eta_1^\ell}^{\eta_1^q} P_1(\eta_1, i_0^*)f(\eta_1)d\eta_1 + \int_{\eta_1^q}^{\eta_1^u} cf(\eta_1)d\eta_1 \\ &+ \int_{\eta_1^u}^{\overline{\eta}_1} E(P_2(\rho\eta_1 + \varepsilon, i_0^* + q_0^*))f(\eta_1)d\eta_1.\end{aligned}$$

Totally differentiating, one obtains

$$\frac{\partial P_0(y_0^*)}{\partial y_0^*}dI - \frac{\partial P_0(y_0^*)}{\partial i_0^*}di_0^* = \frac{d\overline{P}(i_0^*)}{di_0^*}di_0^*.$$

The above implies

$$\frac{di_0^*}{dI} = \frac{\partial P_0(y_0^*)/\partial y_0^*}{\partial P_0(y_0^*)/\partial y_0^* + d\overline{P}(i_0^*)/di_0^*}.$$

Since by assumption prices are decreasing in units sold, it is apparent that the following condition is sufficient to imply $di_0^*/dI > 0$.

$$\begin{aligned}\frac{d\overline{P}(i_0^*)}{di_0^*} = &\int_{\underline{\eta}_1}^{\eta_1^i} \frac{\partial E(P_2(\rho\eta_1 + \varepsilon, i_0^*))}{\partial y_2^*} f(\eta_1)d\eta_1 \\ &+ \int_{\eta_1^i}^{\eta_1^\ell} \left(1 - \frac{di_1^*}{di_0^*}\right) \frac{\partial P_1(\eta_1, i_0^* - i_1^*)}{\partial y_1^*} f(\eta_1)d\eta_1 \\ &+ \int_{\eta_1^\ell}^{\eta_1^q} \frac{\partial P_1(\eta_1, i_0^*)}{\partial y_1} f(\eta_1)d\eta_1 + \int_{\eta_1^u}^{\overline{\eta}_1} \frac{\partial E(P_2(\rho\eta_1 + \varepsilon, i_0^* + q_1^*))}{\partial y_2^*} f(\eta_1)d\eta_1 < 0.\end{aligned}$$

where η_1^i and η_1^q are determined by:

$$P_1(\eta_1^i, 0) = E(P_2(\rho\eta_1^i + \varepsilon, i_0^*)), P_1(\eta_1^q, i_0^*) = c$$

i.e., values of η_1 for which inventory and production, respectively, commence to be sold at t_1, and substitutions have been made as appropriate from equilibrium conditions which must hold at t_1. Note that by the definitions of limits of integration which depend on i_0^*, the derivatives at these limits all disappear. Since prices and expected prices are decreasing in sales, all four terms in $d\overline{P}(i_0^*)/di_0^*$ are nonpositive provided that

$$\frac{di_1^*}{di_0^*} < 1.$$

Since totally differentiating the equilibrium condition

$$P_1(\eta_1, i_0^* - i_1^*) = E(P_2(\rho\eta_1 + \varepsilon, i_1^*)$$

implies this inequality, then the first part of the proof is complete.

For the second part of the proof, it must be shown that

$$\frac{dCY}{di_0^*} < 0.$$

Similar to the above, the derivatives at the limits of integration disappear leaving

$$\frac{dCY}{di_0^*} = \int_{\eta_1^\ell}^{\eta_1^q} \frac{\partial P_1(\eta_1, i_0^*)}{\partial y_1^*} f(\eta_1) d\eta_1 < 0,$$

which completes the proof. Q.E.D.

Proof of Proposition 2

It is shown that

$$\frac{dCY}{dc} = \frac{\partial CY}{\partial i_0^*} \frac{di_0^*}{dc} + \frac{\partial CY}{\partial c} > 0.$$

As before, one can totally differentiate the equilibrium condition

$$P_0(I - i_0^*) = \bar{P}(i_0^*)$$

to obtain

$$\frac{di_0^*}{dc} = \frac{-\partial \bar{P}(i_0^*)/\partial c}{\partial P_0(I - i_0^*)/\partial y_0^* + \partial \bar{P}(i_0^*)/\partial i_0^*}$$

where

$$\frac{\partial \bar{P}(i_0^*)}{\partial c} = \int_{\eta_1^q}^{\bar{\eta}_1} f(\eta_1) d\eta_1 > 0$$

and

$$\frac{\partial CY}{\partial c} = \int_{\eta_1^q}^{\eta_1^u} f(\eta_1) d\eta_1 > 0.$$

Using the above yields

$$\frac{dCY}{dc} = -\left[\int_{\eta_1^\ell}^{\eta_1^q} \frac{\partial P_1(\eta_1, i_0^*)}{\partial y_1^*} f(\eta_1) d\eta_1\right]\left[\frac{\int_{\eta_1^q}^{\bar{\eta}_1} f(\eta_1) d\eta_1}{\partial P_0(y_0^*)/\partial y_0^* + \partial \bar{P}(i_0^*)/\partial i_0^*}\right] + \int_{\eta_1^q}^{\eta_1^u} f(\eta_1) d\eta_1$$

or

$$\frac{dCY}{dc} = -\left[\frac{\int_{\eta_1^\ell}^{\eta_1^q} \frac{\partial P_1(\eta_1, i_0^*)}{\partial y_1^*} f(\eta_1) d\eta_1}{\partial P_0(y_0^*)/\partial y_0^* + \partial \bar{P}(i_0^*)/\partial i_0^*}\right] \int_{\eta_1^q}^{\bar{\eta}_1} f(\eta_1) d\eta_1 + \int_{\eta_1^q}^{\eta_1^u} f(\eta_1) d\eta_1$$

The first term is negative, with the bracketed term a positive fraction. dCY/dc cannot be signed without further restriction because of the possibility that $E(P_2(\eta_2,0)) > c$, i.e., $\eta_1 > \eta_1^u$; this can only occur if ρ is very positive. A sufficient condition to yield $dCY/dc > 0$ is to assume $\eta_1^u > \bar{\eta}_1$, so that $E(P_2(\eta_2,0))$ does not intersect c in the relevant η_1 range. In this case $\bar{\eta}_1$ replaces η_1^u in the last term, and $dCY/dc > 0$.

Proof of Lemma. The result is obvious in Figure 1. In all three panels, for $\eta_1 \leq \eta_1^\ell$, $P_1(\eta_1, y_1^*) = E(P_2(\rho\eta_1 + \varepsilon, y_2^*))$ so that P_1 and P_2 are perfectly positively correlated. However, for $\eta_1 \geq \eta_1^q$, P_1 and P_2 are uncorrelated since $P_1 = c$ in this region. Only in $\eta_1^\ell < \eta_1 < \eta_1^q$ is the correlation of P_1 and P_2 related to ρ, and it is clear that reducing ρ reduces equilibrium price autocorrelation.

Proof of Proposition 3

It is shown that

$$\frac{dCY}{d\rho} < 0.$$

Since CY depends on ρ directly as well as through i_0^*, then

$$\frac{dCY}{d\rho} = \frac{\partial CY}{\partial i_0^*}\frac{di_0^*}{d\rho} + \frac{\partial CY}{\partial \rho}.$$

First, it is easy to show that:

$$\frac{\partial CY}{\partial \rho} = -\frac{\partial E(P_2(\eta_2,0))}{\partial \eta_2}\int_{\eta_1^\ell}^{\eta_1^*} \eta_1 f(\eta_1)d\eta_1 < 0.$$

Then, from the proof of Proposition 1, one knows that

$$\frac{\partial CY}{\partial i_0^*} = \int_{\eta f}^{\eta_1^q} \frac{\partial P_1(\eta_1, i_0^*)}{\partial y_1^*} f(\eta_1)d\eta_1 < 0.$$

Similar to before, one can totally differentiate the equilibrium condition

$$P_0(y_0^*) = \bar{P}(i_0^*)$$

to show that

$$\frac{di_0^*}{d\rho} = \frac{-\partial \bar{P}(i_0^*)/\partial \rho}{\partial P_0(y_0^*)/\partial y_0^* + \partial \bar{P}(i_0^*)/\partial i_0^*}.$$

From the proof of Proposition 1, $\partial \bar{P}(i_0^*)/\partial i_0^* < 0$, implying that the denominator is negative. Thus, the sign of $\partial \bar{P}(i_0^*)/\partial \rho$ determines the sign of $di_0^*/d\rho$.

Given the definition of $\bar{P}(i_0^*)$, then

$$\frac{\partial \bar{P}(i_0^*)}{\partial \rho} = \int_{\underline{\eta}_1}^{\eta_1^i} \frac{\partial E(P_2(\rho\eta_1 + \varepsilon, 0))}{\partial \eta_2} \eta_1 f(\eta_1) d\eta_1$$
$$= \frac{\partial E(P_2(\eta_2, 0))}{\partial \eta_2} \int_{\underline{\eta}_1}^{\eta_1^i} \eta_1 f(\eta_1) d\eta_1 .$$

Thus, a sufficient condition for $dCY/d\rho > 0$ is $di_0^*/d\rho > 0$, i.e., $\partial\bar{P}(i_0^*)/\partial\rho > 0$. By assumption, $\partial E(P_2)/\partial\eta_2 > 0$, and $\underline{\eta}_1 > 0$, so that $\partial\bar{P}(i_0^*)/\partial\rho > 0$, completing the proof. Q.E.D.