Commodity Index Trading and Hedging Costs

Celso Brunetti* and David Reiffen**

April, 2012

Abstract

Trading by commodity index traders (CITs) has become an important aspect of financial markets over the past 10 years. We develop an equilibrium model of trader behavior that relates uninformed CIT trading to futures prices. The model predicts that CIT trading reduces the cost of hedging. We test the model using a unique non-public dataset which precisely identifies trader positions. We find evidence, consistent with the model, that index traders have become an important supply of price risk insurance.

JEL: G13, C32

Acknowledgements
For outstanding research assistance we thank Hannah Ropp and Mehrdad Samadi. For helpful comments we thank Pat Fishe, Jeff Harris, Steve Kane, Jim Moser, Michel Robe, Aaron Smith, Clara Vega, Hao Zhou and participants at the CFTC Conference on Commodity Markets.

Disclaimers
The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.
This paper reflects the opinions of its authors only, and not those of the CFTC, the Commissioners, or any of the authors' colleagues upon the Commission staff. All errors and omissions, if any, are the authors' sole responsibility.

*Division of Research and Statistics, Federal Reserve Board. Email: Celso.Brunetti@FRB.gov
**US Commodity Futures Trading Commission. Email: dreiffen@cftc.gov.

I. Introduction

In recent years, commodity futures contracts in have increasingly become integrated into investment portfolios (Büyüksahin, Haigh, Harris, Overdahl, and Robe, 2009). The amount of money invested globally in commodity indices has grown more than 10 fold between 2003 and 2008 (CFTC, 2008). Commodity index traders (CITs), which have been the main vehicle for investing in commodities, represent a new type of player in these markets (Stoll and Whaley 2010). Because index traders now represent a large portion of the futures trading, questions have arisen regarding the effect they have on financial markets. For example, the impact of index trading on the cost of diversifying (hedging) is an important policy issue, since diversification is a critical function of futures markets. Masters (2008), Singleton (2011) and others have expressed concerns that index trading contributes to pricing distortions, which can affect hedging costs. The goal of this paper is to study the effect of commodity index trading on financial markets. In this regard the paper provides several contributions to the existing literature.

First, we empirically analyze the effect of the increased market participation of commodity index traders on the cost of hedging. Using a unique proprietary dataset that precisely identifies the daily trading activity of commodity index traders, we find evidence that the presence of CITs reduces hedging costs.[1] While most of the literature on the role of CITs in futures markets concentrates on its price effect,[2] we believe that the cost of hedging should play a central role in the analysis.

Second, we generalize Grossman and Miller (1989) to develop an equilibrium model of trader behavior that relates uninformed CIT trading to futures prices. Theoretical models, dating back to Keynes (1930) and Hicks (1939) have focused on the role of futures markets in allowing firms to hedge their positions in the physical commodity; in particular, agents with inherent long positions in the physical product, such as raw material producers, reduce the riskiness of their portfolios by taking short positions in the futures market. The long side of these trades is taken by *speculators*, who are willing to hold these futures positions in exchange for positive expected returns. Hence, these models imply that,

[1] This is a unique dataset collected by Commodity Futures Trading Commission which tracks daily positions in each futures contract for each trader, and identifies the line of business of each trader (e.g., CIT, grain distributor, floor trader). Detailed data on commodity index positions only exists for agricultural commodities. Therefore, our empirical analysis is based on agricultural futures markets only. However, our model and our results are rather general and extend to other markets.

[2] For a comprehensive literature review on this issue see Irwin and Sanders (2010).

even absent superior information about futures price movement, speculators will earn positive returns.[3] These positive returns come about through a futures price which is below the expected spot price at contract expiration – or what is termed *backwardation*. More formal models, such as Hirshleifer (1988) or Etula (2010), show that this conclusion remains, even in a model with many risky assets and opportunities for diversification. None of these models, however, studies commodity index trading. Instead, these models are based upon the dichotomy between hedgers and speculators.[4] CITs behave differently. Like speculators in these models (and unlike hedgers), CITs have no innate position in the underlying commodity. Unlike speculators, CITs seem to follow simple rules that are unrelated to information. That is, as Stoll and Whaley (2010) note, CITs primarily buy and hold a long position in the closest-to-maturity (*nearby*) contract, which entails their *rolling* this position from one maturity to the next, as the nearby contract nears expiration. The premise that their trading is not motivated by superior information is evidenced by their trading rules which are determined and publically disseminated well prior to the trades being executed.[5] Our model analyzes how these traders interact with traditional hedgers and speculators.

The direct implication of our model is that the cost of hedging falls as CIT positions increase. The intuition behind our theoretical result comes from the fact that CITs are essentially willing to take the opposite position from hedgers at lower prices than are traditional speculators.[6] More subtle implications of our model relate to the effect of commodity index traders on inter-maturity spreads. As noted, a key characteristic of CITs is that they primarily hold positions in the nearby contract, which usually is the most liquid contract, and periodically *roll* these positions to the next maturity (*first deferred*) contract. This exogenous movement of positions between maturities provides a natural experiment with which to examine whether inter-contract spreads vary in the manner implied by the model. In particular, our model implies that the spread between the first deferred and the nearby contract depends on both the relative sizes of CIT positions in the two maturities, and on the aggregate size of CIT positions, albeit in a way that varies across the contract

[3] Empirical studies on whether the source of profits for speculators is information advantages or simply risk-taking include Hartzmark (1987), Dewally, Ederington and Fernando (2010) and Fishe and Smith (2010).
[4] See, for instance, Shleifer and Summers (1990), Lux (1995) and Shiller (2003).
[5] For example, in December of each year, the largest index fund, the Goldman Sachs Commodity Index (GSCI), announces its trading plans for the subsequent year.
[6] According to Stoll and Whaley, they are willing to do so because of the diversification benefit of commodity exposure to their portfolios.

cycle.[7] More generally, since changes in prices of contracts of different maturities are not perfectly correlated (and hence are not perfect substitutes from hedgers' perspectives) CITs' choices of which contracts to invest in will differentially impact prices along the term structure. In addition, the model implies that the size of the effect of CITs on the spread varies with the product cycle; for example, for agricultural commodities, the effect is larger just prior to the harvest, because the correlation between the return on the hedger's cash positions and the nearby futures market is higher. Similarly, in non-agricultural markets, the effect is larger for maturities that expire just prior to seasonal demand peaks, due to higher hedging demand. Finally, the model shows that the spread varies with the size of the cash market position of hedgers.

The last contribution of our paper is to test the hypotheses generated by the model using the highly disaggregate data on trader positions provided by the Commodity Futures Trading Commission. Our findings generally support the premise that commodity index traders earn returns by taking on risk that would otherwise either remain with hedgers, or be taken on by speculators at higher prices. In particular, we find that hedging costs are positive, and that the price of hedging is increasing in the cash market position of hedgers, and decreasing in the positions of index traders. In addition, as implied by the model, the inter-maturity spread increases with the percentage of CIT holdings in the first deferred contract, and this effect is particularly large later in the harvest cycle or, more generally, when seasonal demand peaks.

As Hirshleifer (1988) shows, changes in hedging demand can only have price effects when supply is less than perfectly elastic. Our finding that the relative prices of different maturities vary predictably with CIT positions is consistent with the premise that liquidity providers have less-than-perfectly-elastic supply curves.

Although our analysis concentrates on commodity futures markets, this finding can shed light on issues relevant to equities and, in particular, to equity indices. Just as commodity index funds change their portfolios as contract expiration nears, equity index funds change their portfolios in response to publicly observable events. In particular, there is a considerable literature showing that the addition of a stock to a major index increases its share price – see e.g., Shleifer (1986), Harris and Gruel (1986), Greenwood (2005). One interpretation of this phenomenon, which is consistent with the model developed here, is

[7] Examining the effect of trader behavior on inter-maturity spreads is a powerful test of our model, because taking differences mitigates the noise introduced by changes in market fundamentals.

that the redefinition of the index leads to an increase in the demand for that stock, as some mutual funds are contractually obligated to have a portfolio that is representative of a specific index. Thus, the increased share price results from a liquidity effect: i.e. the interaction of this higher demand with a less-than-perfectly elastic supply of existing shares of individual stocks.[8] An alternative explanation is that the addition of the firm to the index represents real information about the long-term prospects of the stocks (Jain, 1987) – e.g., stocks added to the index are, *ceteris paribus*, less likely to face bankruptcy, or will have lower spreads due to increased liquidity (see, e.g., Dhillon and Johnson, 1991). In the case of commodity index funds, there is no parallel to this latter interpretation. As noted above, the timing of the roll by major CITs is announced well in advance. Hence, it would seem that there is no information content in these trades, and any price change can only be attributed to a liquidity effect.

We proceed as follows. Section II describes the Large Trader Reporting System which is used to collect data on trader positions. In Section III we empirically test whether the size of CIT positions affects hedging costs, and find that larger CIT positions does reduce hedging costs. This finding motivates the development of a model of uninformed trading and price behavior in Section IV. We test the model in Section V and conclude in Section VI.

II. Trader Positions

The position data used in this study comes from the US Commodity Futures Trading Commission's (CFTC) Large Trader Reporting System (LTRS). This non-public database contains end-of-day positions for each *large* trader, where *large* is defined as having a position greater than some threshold number of contracts, with the threshold differing across contracts.[9] Large traders typically represent about 70-75 percent of the open interest in the contracts evaluated in this study. The LTRS reports the long and short positions of each large trader in each maturity futures contract, including the delta-adjusted options positions. We examine data for the period July, 2003 through December 2008.

The data used here have several advantages over the more aggregate data that is publicly-available, and has been used elsewhere (see e.g., Stoll and Whaley, 2010).

[8] The interpretation that the higher cost reflects a less-than-perfectly-elastic supply of a stock is analogous to the model developed here, in which the higher price is due to a less-than-elastic supply of price risk insurance.

[9] For wheat, for example, a large trader is defined as someone who has a position of more than 150 contracts.

Specifically, we use daily data, while the publicly-available data refers to weekly observations. Second, our data is disaggregated by maturity. In combination, these features allow us to more accurately measure movements of individual trader's (or groups of traders) positions between maturities, and hence estimate the effect of hedger and CIT positions in specific maturities on prices and inter-month spreads. Finally, our data is available at the individual trader level, which, as discussed below, allows us to measure CITs over a longer period of time.

In addition to reporting their futures and options positions, traders self-report their lines of business. Table 1 lists the nine trader categories that are important in agricultural products, along with the average number of traders and average net positions of all large traders (summed across maturities) in each category for the three most actively-traded field crops. Large traders in the first five categories are involved in some aspect of the grain industry (and are denoted *commercial* traders), and it is likely that their positions primarily reflect a desire to reduce their inherent risk, i.e. they are hedgers. For example, owners of grain storage facilities (category AD, whom we refer to as agricultural distributors), typically acquire long positions in physical grain, and therefore take short positions in futures markets to hedge their price exposure. Hence, to some extent, trader positions in these futures markets are natural reflections of their underlying business. As indicated in Table 1, short futures positions by distributors represent about 31 percent of open interest in wheat and corn, and about 23 percent in soybeans.[10] Other commercial traders likewise tend to take short futures positions, especially in soybeans and wheat.

There are other participants in futures markets who have no innate position in the physical commodity and are referred to as *non-commercials* (speculators). The non-commercial category Floor Brokers and Traders (FBT) consists of traders who have no physical presence in the industry, but instead take long or short positions in order to take advantage of what they view as favorable prices (these traders are sometimes referred to as *locals*). They typically make bids and offers on the same day, serving as market makers by effectively providing liquidity to other market participants. Two other categories of non-commercial traders, MMT and NRP, are firms that manage investment portfolios, often referred to as hedge funds.[11]

[10] Throughout the paper, we often refer to soybeans simply as soy.

[11] The difference between the two groups relates to minor regulatory differences. In any case, the participation of the latter group is not very large in the markets we analyze.

Like other non-commercial traders, CITs have no physical presence in these agricultural markets. This category is different from the other categories in that it is not a self-reported category. Instead, it reflects an effort by the CFTC to develop statistics to monitor an important change in agricultural futures markets. Specifically, as commodity index traders began to hold a larger portion of open interest, the CFTC as well as many industry participants, became interested in enhanced tracking of the positions of these traders. Accordingly, in 2006, the CFTC reclassified some traders into this new category for 12 agricultural products. There was no corresponding reclassification in other futures market, such as energy or financial futures. Hence, the clearest picture of the effect of CIT trading occurs on agricultural markets, and consequently, we focus on the three largest agricultural markets. However, our model and empirical methodology are general and apply to other markets as well.

The determination as to which traders constitute CITs is based on identifying all traders with large long positions in agricultural futures contracts and evaluating whether the trades made by those firms were consistent with index trading, as well as a series of interviews with the traders (CFTC 2006). In this study, we categorize firms as CITs throughout the sample if they met the CIT criteria as of 2006 (as discussed below, the evidence suggests this treatment is appropriate). Once CITs are identified, we can track their positions back to dates prior to 2006. This allows us to have a longer time series of observations on CIT behavior. Figure 1 presents some evidence on the relative importance of CITs in the corn market. The vertical axis represents the largest end-of-day position held by CITs in each nearby contract, as a percentage of total open interest on that date. CIT's long position in the nearby corn futures contract represented about 25 to 30 percent of the total open interest in 2003. The percentage grew fairly consistently through late 2006, and fell somewhat over the last two years of the sample. One noteworthy aspect of this pattern is that it does not appear that 2006 represents a structural break in the series; instead, CIT's aggregate position in 2006 seems to be a continuation of the previous trend. The history of CIT positions in soybeans and wheat is quite similar.

III. The Price of Hedging

The traditional view of futures markets is that they allow traders with innately risky positions to hedge that risk. The canonical example is a grower who owns crops that will mature at some future date, and consequently, faces price risk until the crop is sold. By taking a short position in the futures market in that commodity, this grower is able to essentially sell the crop earlier in the season, and thereby reduce her exposure to price risk over the production horizon. Her counterparties on the futures market may also be reducing his risk. That is, by taking a long position in the futures market, a firm that plans to buy the crop after it matures (like a flour mill) can likewise reduce its exposure to price risk by buying the crop in advance. If these two kinds of hedgers are the only traders, then the futures price of the crop would reflect the relative demands of the two groups. However, the price that clears the market when only hedgers are present may be sufficiently high or low (relative to expected spot prices) that traders with no innate interest in the commodity may find it profitable to trade on one side of the market. The premise of the theory of *normal backwardation*, advanced by Keynes (1930) and Hicks (1939) is that the relative demands of long and short hedgers are such that futures price will be below the corresponding expected future spot price (i.e. short hedging demand exceeds long hedging demand). More recent work, such as Hirshleifer (1988) and Etula (2010), extends their framework to consider a broader set of portfolio options for speculators.[12] The basic conclusion remains that assets (such as futures) that have a positive correlation with the innate risk held by firms will have a positive return to long positions.

While the sign of the net position of all hedgers could conceivably vary across markets, the evidence is that for most commodities, hedgers are net short in the futures market. For agricultural commodities, the largest group of hedgers is distributors, who have innate long positions in physical agricultural products. In fact, these distributors hold a much larger absolute share of open interest than any other group, as shown in Table 1. Their physical market positions typically consist of forward agreements with growers to buy crops at set prices, as well as crop inventories. Evidence on cash market positions in agricultural commodities also suggests that traders with long physical market positions hedge, by taking a short positions in the futures contract, much more often than traders with a short physical position.[13] More generally, hedgers as a group tend to be short on net in most commodity futures markets. Then, it follows that speculators would primarily take the long side, as implied by the backwardation model.

[12] Gorton, Hayashi, and Rouwenhorst (2007) look at backwardation in a model with physical inventories. They show that the price of hedging is decreasing in inventories.
[13] See Brunetti and Reiffen (2012).

Put differently, one can think of the cost of hedging as the equilibrium discount (from expected spot prices) hedgers accept in order to avoid price risk. If the discount is positive, then the party who is short in the futures contract loses money, on average, on the futures contract. However, for an agent who has a pre-existing long position in the physical product (and therefore hedges in the futures market by taking a short position), that cost can be justified by the reduction in the variability of returns.[14] In this sense, hedging is a form of insurance.

The price of this insurance can then be modeled in terms of demand and supply. In Acharya, Lochstoer and Ramadorai (2010), demand for hedging is modeled as reflecting the risk aversion of producers (who hedge), while supply is modeled as reflecting the financing constraints on speculators. While the extent of risk aversion affects the price of hedging in the model we present in Section IV as well, we focus on changes in the hedgers' innate position in the underlying as a cause of demand shifts. The key point of departure, however, is our focus on changes in CIT positions as a source of supply shocks.[15] In either case, however, the less-than-perfectly elastic supply (*limit to arbitrage*) means that demand changes lead to price changes. As Hirshleifer (1988) emphasizes, different assumptions about the elasticity of supply of insurance yield different implications regarding the relationship between hedger characteristics and the cost of insurance. In line with Acharya *et al.* (2010) our findings suggest that hedging supply is less than completely elastic.

Our measure of the cost of hedging at time t is

$$(E_t(P_T) - P_t)/P_t$$

where P_t is the futures price on day t, and day T is the expiration date of the futures contract. Based on the usual arbitrage argument, we assume that futures and spot prices converge on the expiration date, so that $E_T(P_T)$ is equal to the expected spot price on date T. While $E_t(P_T)$ is not directly observable at every t, $E_T(P_T)$ is observable. Moreover, $E_T(P_T)$ is an unbiased estimate of $E_t(P_T)$ for each contract. Of course, this does not mean that $E_t(P_T)$ will not vary over time. In particular, the expected return on the commodity should be increasing in the risk-free rate (r) and also vary with the number of days until contract

[14] Of course, a premium over the expected spot can be consistent with hedging, if hedgers are on net, short in the physical (i.e., long in the futures market). As noted, CFTC data show that hedgers are primarily short in the futures markets.

[15] Singleton (2011) provides evidence that changes in hedging supply have had significant effects on prices in crude oil futures markets.

expiration (Λ). Thus, to test whether the price of hedging is decreasing in CIT positions and increasing in hedger cash positions, we first run a regression of the form

$$\frac{E_{T,i,j}(P_{T,i,j}) - P_{t,i,j}}{P_{t,i,j}} \equiv Y_{t,i,j} = b_{0,i,j} + b_{1,i,j}\,\ddot{v}_{\ t,i,j} + b_{2,i,j} r_{t,i,j} + b_{3,i,j} r_{t,i,j} \Lambda_{t,i,j} + \varepsilon_{t,i,j} \tag{1}$$

for each of the 27 contracts (i) for each of the three commodities (j) in our sample, during the period in which each contracts is the nearby. The left-hand side variable in equation (1) is $Y_{t,i,j}$, the cost of hedging for product j (j = wheat, corn, soybeans) in contract i ($i = 1,2, \ldots, 27$).[16] The coefficients of primary interest is $b_{0,i,j}$, which measures the average hedging cost at the median date during the period in which the contract is the nearby, and the standard deviation of the error term, $\sigma_{\varepsilon,i,j}$, which represents the volatility of the cost of hedging. Of particular interest is how $b_{0,i,j}$ and $\sigma_{\varepsilon,i,j}$, vary with the demand and supply of insurance. Hence, our second stage regressions are

$$b_{0,i,j} = a_{0,j} + a_{1,j} I_{i,j} + a_{2,j} C_{i,j}^{agg} + \nu_{i,j} \tag{2}$$

$$\sigma_{\varepsilon_{i,j}} = w_{0,j} + z_{1,j} I_{i,j} + z_{2,j} C_{i,j}^{agg} + \xi_{i,j} \tag{3}$$

where $I_{i,j}$ is our measure of CIT positions in product j in contract i, and $C_{i,j}^{agg}$ is our measure of the cash positions of agricultural distributors in the underlying commodity, computed as the median value of those positions over the period of time during which the contract is the nearby.[17] The implication of the backwardation model is that $a_{1,j}$ and $z_{1,j}$ should be negative while $a_{2,j}$ and $z_{2,j}$ should be positive.

We estimate equation (1) using OLS, with Newey-West standard errors, for the 27 maturities of each of the three commodities.[18] The results (available from the authors upon request) for the b_1, b_2 and b_3 are consistent with our expectation; the hedging discount is increasing in both r and Λ, and in the interaction term. The coefficients are statistically significant at the 5 percent level in about 95 percent of cases. Moreover, the average R^2 is slightly over 50 percent, indicating that the regressions are explaining most of the variation in the cost of hedging. Overall, standard test-statistics indicate that the regressions in the first step are well-behaved.

Table 2 presents our estimates of the regressions in equations (2) and (3). Standard errors are bootstrapped (see, e.g., Varian 2005). In our estimates for equation (2), for all three commodities, the average cost of hedging, b_0, is increasing in C^{agg}, and decreasing in

[16] We normalize the cost of hedging by the price level to facilitate comparison across commodities.
[17] We treat I as constant in each contract. In fact, as shown in figure 3, the aggregate position of CITs varies little over the course of the contract.
[18] There are 40 to 80 observations in each regression.

I, as suggested by the backwardation model. In both the corn and wheat regressions, all of the coefficients have p-values of less than 20 percent (each regression has only 27 observations). More generally, this rather parsimonious model does a reasonable job in explaining changes in average hedging costs for these two products. The lower explanatory power for the soy market could reflect the fact that futures markets exist for the two value-added soy products, soy bean meal and soy oil, so that traders may employ more complex hedging strategies (thereby reducing the accuracy of C^{agg} in measuring cash positions, among other things). To interpret the coefficients, consider the effect of a one standard deviation increase in I. For wheat, a one standard-deviation increase in CIT positions represents a change of about 50 thousand contracts, which the model implies will reduce hedging costs by 2 basis points, or about 1/3 of the mean hedging cost.

Estimation results for equation (3) suggest that the variance of hedging costs is increasing in C^{agg} and decreasing in I, with C^{agg} in the corn regression and I in the soy regression having the greatest statistical significance. The R^2s are considerably higher than in our estimations of equation (2).[19]

These results suggest that index traders provide insurance for hedgers in agricultural markets, thereby reducing hedging costs, relative to a market in which index traders are absent. Although C^{agg} and I seem to explain some portion of variation in the mean and in the variance of hedging costs, the underlying economics suggests that there is likely to be considerable measurement error associated with our hedging cost variables. In particular, $E_T(P_T)$ may be a noisy estimate of $E_t(P_T)$. Among other considerations, $E_T(P_T)$ will reflect information that is revealed between t and T (e.g., crop forecasts). To address the issue of measurement and estimation error and to better understand the role of CITs in futures markets, we formally model the market price for risk insurance, and derive implications for inter-month price spread, as well as hedging costs. The advantage of

[19] Given the potential noisiness of $E_t(P^T)$, as a robustness check, we consider an alternative measure of hedging cost: the daily hedging cost defined as $Y_{t+1,i,j} = (P_{t+1,i,j} - P_{t,i,j})/P_{t,i,j}$, and re-estimate equations (1), (2) and (3). The results (available from authors) are very similar to those reported above. As a further robustness check of our results, we also estimate the following GARCH model using daily observations:

$$Y_{t,j} = \varsigma_{0,j} + \varsigma_{1,j} r_{t,j} + \varsigma_{2,j} \Lambda_{t,j} + \varsigma_{3,j} I_{t,j} + \varsigma_{4,j} C_{t,j}^{agg} + \chi_{t,j}$$

$$\chi_{t,j} = \sqrt{h_{t,j}}\eta_{t,j};\ \eta_{t,j} \sim i.i.d.\Psi(0,1)$$

$$h_{t,j} = \omega_{0,j} + \omega_{1,j} h_{t-1,j} + \omega_{2,j} \chi_{t-1,j}^2 + \omega_{3,j} I_{t,j} + \omega_{4,j} C_{t,j}^{agg}$$

where again, $Y_{t,j} \equiv (E_{t,j}(P_{T,j}) - P_{t,j})/P_{t,j}$. Finally, we consider a version of this model in which $Y_{t+1,j} = (P_{t+1,j} - P_{t,j})/P_{t,j}$, the daily hedging cost. The results of both of these estimations unequivocally show that the cost of hedging and the volatility of the cost of hedging are decreasing in the CIT positions ($I_{t,j}$) and increasing in the cash positions of hedgers ($C_{t,j}^{agg}$). In sum, these robustness checks confirm the findings reported in the Table 2.

analyzing inter-month spreads is that changes in market fundamentals will likely be reflected in the prices of all maturities of the commodities, and hence likely not have a large impact on price differences. The next section first presents some evidence on traders' behavior in the markets we study, and then presents a formal model in which we incorporate these features to derive implications concerning the relationship between trader behavior and both hedging costs and inter-month spreads.

IV. The Effect of CIT Trading on Pricing

In this section, we present a model of equilibrium in futures markets in which CITs participate. The model incorporates several salient features of commodity markets. The first important feature is that contracts of different maturities trade simultaneously. At any point in time, eight or more contracts of different maturities are trading in each product. Second, hedgers are net long in the physical product, hence their hedging consists of short futures positions. Importantly, the hedgers' risk primarily pertains to price changes between the trade date and the date at which the harvested crop can be bought or sold. As shown below, this implies that most hedgers will take large short positions in the first post-harvest futures contract each year.[20] The model enables us to characterize the impact of index trading on equilibrium prices.

A. Empirical Regularities

Figure 2 presents some evidence in support of the premise that hedgers take particularly large positions in the post-harvest contract. It shows that during the six plus years in our sample, the most important group of hedgers, distributors, have on average, established short positions of nearly 20,000 contracts in each year's December wheat contract by mid March (270 days before expiration of the December contract), and retained positions of about that magnitude until the December contract became the closest-to-maturity (nearby). In contrast, for other maturity months, their typical short positions are less than 5,000 contracts until the contract becomes the nearby. Similar patterns exist for corn and soybeans.

[20] For non-agricultural commodities, hedgers will take large short futures positions in maturities corresponding to when they are building up inventories. Consequently, the biggest short positions are in the contracts that expire just prior to demand peaks.

The other relevant institutional feature is that the leading CIT traders largely establish their trading positions independent of contemporaneous price. As Stoll and Whaley (2010) note, commodity index traders typically have simple buy-and-hold strategies, which allow them to take advantage of the diversification these assets provide. Stoll and Whaley (2010) term this *commodity index investment*, rather than trading, in order to emphasize the passive nature of their investment. The two largest index funds (the GSCI and the DJ-AIG[21]), which together represent about 1/3 of CIT positions during our sample, announce the bulk of their annual futures market trading decisions prior to the first trading day of the year.[22] In particular, they announce the percentage of their assets that will be allocated to each futures contract, which maturities of those commodity contracts they will hold, and the dates they will move positions between maturities.[23]

One common feature of CITs is that they primarily take positions in the nearby contract, which requires them to move their positions from the soon-to-expire nearby contract to the succeeding maturity contract.[24] Figure 3 shows CIT position in the nearby and the first deferred contracts for the three commodities, as functions of the number of days until the nearby contract reaches its expiration, for a typical month (May 2007).[25] The Figure shows that while CIT's overall position does not vary much over the course of a contract lifecycle, their positions in individual maturities vary dramatically over this cycle. Specifically, for all three commodities, CITs' position in the first deferred contract is small compared to their position in the nearby contract at a point two months prior to expiration. Over the succeeding month or so, they move their positions from the nearby to the first deferred contract. Most of this roll takes place between 30 and 40 days prior to contract expiration.

CITs are like open-end index mutual funds, in that they represent individual and institutional investors, who are free to change their holdings of the index at any time. Hence, while the timing of CITs' roll is pre-announced, the amount of money invested in an index can vary. For this reason, changes in CIT positions can potentially represent

---

[21] AIG sold its index fund to UBS in 2009, subsequent to our sample period. Hence, the index fund is currently known as the DJ-UBS index fund.
[22] Index funds are a major subset of index traders. See Stoll and Whaley (2010) for a thorough description of the practical aspects of commodity index trading.
[23] See, for example, the GSCI Manual 2005 Edition, dated 12/2004.
[24] In the last few years, some CITs have begun taking long positions in more distant contracts and holding them for longer, which requires less rolling between contracts.
[25] We also produced similar figures for average positions (across all 27 maturities) as a function of days until expiration. Those figures looked very similar to Figure 3.

information. We see the potential for the roll to contain information about fundamentals to be minimal, however, for several reasons. First, even if investors have private information (i.e., information not yet incorporated into futures prices), it is unlikely that the investors would gain that information at a time coincident with the roll. Moreover to the extent that investors have private information about individual commodity prices, trading on that information would be more profitable if it were directed to the individual futures market, rather than to a bundle of commodities.[26] For these reasons, we view CIT behavior during the roll as unlikely to reflect new information about fundamentals.

Of course, when CITs acquire a long position in a contract, there must be counterparties with corresponding short positions in that contract. Figure 4 displays the average positions of four groups of traders in the nearby corn contract, as it moves towards expiration. Market makers (floor brokers), hedge funds (managed money), and agricultural distributors (hedgers) hold positions that are in aggregate about the same size as CIT positions. Market makers' overall position looks quite different, however. Figure 5 shows the pattern of market maker positions in the wheat market. Market makers hold long positions in deferred wheat contracts that are nearly equal in size to their short positions in the nearby contract. As we show below, this pattern is consistent with rational behavior by market makers, given the strategies of CITs and agricultural distributors (hedgers).

To summarize, as an empirical regularity, we see that CITs primarily take long positions in the nearby contract, and their counter-parties in that contract consist of distributors, market makers and hedge funds. These latter two groups appear to hedge their short positions in the nearby by taking long positions in deferred contracts. CITs move positions from the nearby to the first deferred contract in a predictable manner, as the nearby moves towards expiration. Finally, hedgers take especially large short positions in the post-harvest contract each year.

B. Modeling Trader Behavior

To reflect these empirical regularities, we consider a model with two maturities of contracts in a single commodity each year, and three trader types; short hedgers (distributors), index traders (CIT), and speculators (specifically market makers and

[26] Put somewhat differently, informed traders are likely to be firms involved in commercial activities in a commodity, a group that generally trades in futures markets. In contrast, investors who trade through CITs are generally institutions and individuals with no direct involvement in commodity markets.

investment managers).[27] Each contract is the nearby for a *T*-day period, and we refer to period i as the T days in which maturity $i+1$ is the nearby. We characterize each kind of agent in a way that is broadly consistent with their observed trading patterns.

Index Traders: To reflect the buy-and-hold strategy of their investments, we model index trader positions as exogenous; during the period in which contract i is the nearby, CITs have an initial long position of size I_i in contract i, and roll into maturity $i+1$ over the course of the T-day period (consistent with the pattern shown in Figure 3). We let γ denote the share of the CIT position remaining in the nearby contract (so that $\gamma = 1$ at $t = 0$).

Other Traders: Other trader groups optimally allocate their portfolios, anticipating CIT behavior, and viewing the trading activity of CITs as not being information-based. Hence, we assume that hedgers and speculators take utility-maximizing positions in the various maturity futures contracts that are traded each day, and have identical knowledge of market fundamentals. Traders in these groups differ only in regard to their endowments; hedgers have positions in the underlying that essentially result in their being long in the physical commodity. Importantly, these "physical" positions cannot be sold at $t=0$ (very much in the spirit of Grossman and Miller, 1989). Specifically, we assume that hedgers (distributors) have cash market positions of size C in the maturity $i=2$ contract (e.g., the current-year crop will be harvested sometime between the expiration of contract 1 and the expiration of contract 2).[28] This leads to seasonality in hedging demand. A similar seasonality in hedging demand exists in many non-agricultural markets. For example, due to seasonality in product demand, primary metal producers and fabricators hold larger physical inventories in anticipation of increased sales (e.g., in late winter), and seek to hedge those inventories through futures markets positions.

C. Equilibrium Pricing

In contrast to CITs, hedgers (distributors) and speculators choose utility-maximizing positions, given prices. In this section, we solve for the positions of these traders, and the resultant prices. We simplify the analysis by assuming that consumption only takes place

[27] Of course, CITs are speculators in a fundamental sense. We use the term speculator in the context of our model to refer to traders that both have no position in the underlying physical product, and take positions based on contemporaneous prices.

[28] In the two-period model presented here, it is appropriate to think of period 1 as a post-harvest contract (in which hedgers realize the value of their endowment at the end of the period), and period 0 as a pre-harvest period. In the Appendix, we generalize the model to have multiple pre-harvest periods each year, and multiple years. This allows us to examine how spreads vary over the harvest cycle.

in period 2. We consider a two-period model; trading occurs at two dates, 0 and 1, in two tradable assets (contracts). The two assets are futures contracts on the same commodity, maturing at dates 1 and 2, respectively. The maturity 1 contract only trades on date 0, while the maturity 2 contract trades on both date 0 and date 1. Figure 6 portrays the timing of events.

We can write an agent's utility as

$$U[W_0 + X_1^2(P_1^2 - P_0^2) + X_2^2(P_2^2 - P_1^2) + X_1^1(P_1^1 - P_0^1) + (P_2^2)C_k] \qquad (4)$$

where P_j^i is the price of futures contract i *(i=[1,2])* at time j *(j= [0,1,2]),* X_j^i is the trader's position at time j in the futures contract that matures at time i, W_0 is initial wealth, and C_k is the trader's cash position in the underlying product (= C for hedgers, and 0 for speculators). That is, the agent consumes his entire period 2 wealth, which is equal to his initial wealth, plus the value of his position in the underlying as of t = 2, plus or minus the gain/loss he or she makes on his or her futures trades.

To make the analysis tractable, we make the standard assumptions that the distribution of price changes is normal, and that each distributor and speculator has the same exponential utility functions of the form

$$U(W) = A - exp(-\alpha W_2).$$

In combination, these assumptions mean that each agent's utility function depends only on the mean and variance/covariances of the price change distributions.

To determine the optimal futures positions of these agents in periods 0 and 1, we use backward induction. Solving for the period 1 optimum, we note that both speculators and hedgers choose $X_2^2$to maximize (4), given the realizations of previous prices and previous choices of X_j^i, and their expectations about the mean and variance of the distribution of P_2^2. This leads to demand for futures positions equal to

$$X_{2,k}^2 = \frac{E_1[P_2^2 - P_1^2]}{\alpha(\sigma_2^2)^2} - C_k \qquad (5)$$

where σ_{jd}^{if} is the covariance between the time j and d changes in prices of maturities i and f (e.g., $\sigma_{12}^{12}$is the covariance between changes in the period 1 price of maturity 1 and the period 2 price of maturity 2; when j = d or i = f, we indicate the time/maturity with a single subscript/superscript; e.g., $(\sigma_2^2)^2 \equiv (\sigma_{22}^{22})^2$ is the variance of changes in P_2^2).[29] In order for the market for maturity 2 to clear on date 1, it must be the case that

$$N_H X_{2,H}^2 + N_S X_{2,S}^2 = -I_2$$

where I_2 is index trader position at t = 1, all of which is invested in contract 2, by construction; underscores H and S refer to hedgers and speculators, respectively. Since $X_{2,S}^2 = X_{2,H}^2 + C$, we have $X_{2,H}^2 = \frac{-(I_2+N_S C)}{N_S+N_H}$.

Finally, using equation (5), we find that

$$P_1^2 = E_1(P_2^2) + \frac{\alpha(\sigma_2^2)^2}{N_S+N_H}(I_2 - N_H C) \qquad (6).$$

Turning to the optimization at t = 0, given the choice of $X_{2,k}^2$, agents' choose levels of X_1^2 and X_1^1in order to maximize (4), given their expectations as of t = 0, and their optimal decision at t = 1 (i.e., equation (5)).

The two first-order conditions at t = 0 respect to positions are

$$(\sigma_1^1)^2 X_{1,k}^1 + \sigma_1^{12} X_{1,k}^2 = \frac{E_0[P_1^1 - P_0^1]}{\alpha} - \sigma_{12}^{12}(X_{2,k}^2 + C_k) + \sigma_1^{12}(X_{2k}^2) \qquad (7)$$

$$\sigma_1^{12} X_{1,k}^1 + (\sigma_1^2)^2 X_{1,k}^2 = \frac{E_0[P_1^2 - P_0^2]}{\alpha} - \sigma_{12}^2(X_{2,k}^2 + C_k) + (\sigma_1^2)^2 X_{2,k}^2 \qquad (8)$$

As above, market clearing implies that

$$N_H X_{1,H}^1 + N_S X_{1,S}^1 = -\gamma I_1$$

and

$$N_H X_{1,H}^2 + N_S X_{1,S}^2 = -(1-\gamma) I_1.$$

Using the relationship $X_{2,H}^2 = X_{2,S}^2 - C$, it follows from (7) and(8) that $X_{1,H}^1 = X_{1,S}^1$, and $X_{1,H}^2 = X_{1,S}^2 - C$. Hence, we can express $X_{1,S}^1$ and $X_{1,S}^2$ in terms of the exogenous variables: C (the hedgers' cash positions), commodity index futures positions and the variance/covariance matrix. We then use equations (7) and (8) to solve for the two equilibrium time 0 prices

$$\frac{E_0[P_1^1 - P_0^1]}{\alpha} = \sigma_{12}^{12}(X_2^2 + C) + (\sigma_1^1)^2 X_1^1 + \sigma_1^{12} X_1^2 - \sigma_1^{12} X_2^2$$

which yields the following price for the maturity 1 contract

$$P_0^1 = E_0(P_1^1) - \alpha\left[\sigma_{12}^{12}\left(\frac{N_H C - I_2}{N_S+N_H}\right)R - (\sigma_1^1)^2\left(\frac{(\gamma I_1)}{N_S+N_H}\right) - \sigma_1^{12}\left(\frac{(1-\gamma)I_1 + N_S C}{N_S+N_H}\right) + \sigma_1^{12}\left(\frac{I_2+N_S C}{N_S+N_H}\right)R\right], \qquad (9)$$

and

[29] We model each trader as a price-taker with respect to market prices, so that each trader views the variances and covariances of price changes as exogenous to his or her trading decisions.

$$\frac{E_0[P_1^2 - P_0^2]}{\alpha} = \sigma_{12}^2(X_2^2 + C) + \sigma_1^{12}\, X_1^1 + (\sigma_1^2)^2\, X_1^2 - (\sigma_1^2)^2\, X_2^2$$

which yields the following price for the maturity 2 contract

$$P_0^2 = E_0(P_1^2) - \alpha\left[\sigma_{12}^2\left(\frac{N_H C - I_2}{N_S + N_H}\right) R - \sigma_1^{12}\left(\frac{(\gamma I_1)}{N_S + N_H}\right) - (\sigma_1^2)^2\left(\frac{(1-\gamma)I_1 + N_S C}{N_S + N_H}\right) + (\sigma_1^2)^2\left(\frac{I_2 + N_S C}{N_S + N_H}\right) R\right] \quad (10)$$

where $R = \left[N_S^1 / \left(N_S^1 + N_H^1\right)\right] / \left[N_S^0 / \left(N_S^0 + N_H^0\right)\right]$.

This analysis characterizes equilibrium behavior and the implications for pricing at two trade dates; t = 0 and t = 1. More generally, if one were to imagine a series of dates during period 0 (i.e., prior to the expiration of contract 1), equations (7) and (8) represent equilibrium behavior on each of those dates. Similarly equation (2) characterizes behavior at every time between date 1 and date 2 (period 1). Of particular interest is the effect of CIT rolling during each period. That is, as depicted in Figure 3, during each period, CITs generally move their positions from the nearby contract (i.e., contract 1 in period 0) to the first-deferred contract (i.e., contract 2 in period 0). As shown below, our model yields testable implications for prices and spreads associated with those changes in positions. In this sense, equations (9) and (10) characterize the cost of hedging using contracts 1 and 2, respectively, for every t in period 0. These equations formalize the premise that hedging costs are increasing in C and decreasing in I. It also suggests that, roughly speaking, whether traders can make positive expected returns by holding contract 1 to maturity (normal backwardation) depends on the relative sizes of C and I. That is, the price of insurance is likely to be positive if the size of the cash-market position of traders seeking to hedge (C) is large relative to positions of traders seeking exposure to futures price variability (I). Equations (9) and (10) also formalize the proposition that the price of hedging is increasing in the covariance between the returns of hedgers' endowed position and the futures contract.[30]

Note that the model predicts that speculators will take calendar spread positions. Specifically, it predicts that they will be short in the nearby and long in other maturity contracts, whenever CITs are primarily invested in the nearby (i.e., when γ is greater than 0.5).[31] As CITs move their positions into the first deferred contract, speculators will likewise take a larger short (or smaller long) position in the first deferred contract. It also predicts

---

[30] Hirshleifer (1988) finds this same result in a model in which there is only one futures contract, but speculators can hold assets in other asset classes.

[31] For example, if $\gamma = 0$ in period 1, then $X_{1,S}^2 > 0$, and $X_{1,S}^1 < 0$.

that hedgers will have a larger position in the nearby in the post-harvest contract than the pre-harvest contract (i.e.,$|X_2^2| > |X_1^1|$).

These predictions are consistent with observed trading patterns (see Figures 2 and 5). As shown in Figure 4, the position of speculators is the mirror image of CIT positions in the nearby contract, especially between 60 and 15 days prior to expiration. However, their positions aggregated across all maturities are quite different; the net positions of speculators, aggregated across all maturities, is much closer to zero, and changes very little as the nearby reaches maturity. These patterns suggest that speculators are serving as counter-parties to CITs in the nearby, and to hedgers in more distant maturities.

In the Appendix, we generalize the model to develop additional testable implications. First, we consider a model with multiple pre-harvest periods. This allows us to evaluate how prices change over each production cycle. Second, we consider a model in which hedgers have an endowment in both periods, and consume in multiple periods, which corresponds to an environment in which hedgers have periodic (e.g., annual) harvests. This enables us to generate implications for prices in the post-harvest periods.

Overall, we find that the model explains many of the patterns of traders' behavior. It also yields implications for futures prices. In the next section, we develop some testable implications for price differences between maturities. We focus on price differences, because price levels will likely largely reflect demand and supply fundamentals, but changes in these fundamentals are likely to affect all prices similarly, so that changes in differences are more likely to reflect the factors in our model.

D. Comparative Statics

Because CITs primarily take positions in only one or two maturities at any time, to the extent that CIT trading affects futures prices, it will affect different maturities differentially. Thus, one might expect the intermonth spread to vary predictably with CIT behavior. Specifically, we next consider how S_0 (the intermonth spread) $\equiv P_0^2 - P_0^1$ varies over the course of a period. Using equations (9) and (10), we have the following comparative static result

$$\frac{\partial S_0}{\partial \gamma} = \frac{\alpha I}{N_H + N_S}[2(\sigma_1^{12} - (\sigma_1^1)^2 - (\sigma_1^2)^2] < 0.$$

That is, as the index traders (in aggregate) roll their positions from maturity 1 contract to maturity 2 contract (γ falls), the spread between the futures prices of contract 2 and contract 1 rises. This is to be expected, since there is a selling pressure on the maturity 1 contract and a buying pressure on the maturity 2 contract. Spreads will also vary with the aggregate size of CIT positions at $t = 0$

$$\frac{\partial S_0}{\partial I_1} = \frac{\alpha}{N_H + N_S}[(2\gamma - 1)\sigma_1^{12} - \gamma(\sigma_1^1)^2 + (1-\gamma)(\sigma_1^2)^2].$$

The sign of this expression varies with γ; it is positive for γ = 0, and negative for γ = 1. The logic is that when γ = 0, CITs only have positions in maturity 2, and the larger their positions, the higher is P_0^2, while P_0^1 is unaffected (and conversely when γ = 1). More generally,

$$\frac{\partial S_0^2}{\partial\gamma\partial I} = \frac{\alpha}{N_H + N_S}[2\sigma_{12}^2 - \sigma_1^2 - \sigma_2^2] < 0$$

so that the larger the size of the position being rolled, the more rapidly the spread increases with the percentage of their holdings in the first deferred contract.

Finally, changes in the hedgers' cash positions, C, will tend to have a negative effect on the spread

$$\frac{\partial S_0}{\partial C} = \frac{1}{N_H + N_S}[N_S(R-1)(\sigma_1^{12} - (\sigma_1^2)^2 + N_H(\sigma_{12}^{12} - \sigma_{12}^2)].$$

This is negative as long as R ≥ 1 and $\sigma_{12}^2 > \sigma_{12}^{12}$, i.e., P_0^2 is more closely correlated with P_1^2 than is P_0^1. The latter condition seems reasonable, since we would expect the correlation between movements in period 1 and period 2 futures prices to be higher between the same maturity than adjacent maturities.

As shown in the Appendix, this conclusion regarding the relationship between C and the spread depends on where in the harvest cycle one evaluates the spread. The analysis in the Appendix shows that the effect of C on the spread is more likely to be negative later in the harvest cycle (because the correlation between the price change for the nearby maturity and the endowment is higher), and will be unambiguously positive during the post-harvest period.

Another implication of the model is that the price of hedging – which is really the reciprocal of the return to holding a long position – should be correlated across commodities, at least those commodities within the typical fund's holdings. That is, since index funds tend to hold a fixed portion of their portfolios in each of many commodities, changes in CIT positions will be highly correlated across the commodities CITs buy and hold. Since changes in CIT positions will change futures returns in the same direction for all of these commodities, we would anticipate that the presence of CITs should increase the correlation of futures returns across contracts in which CIT take positions, even those unrelated in demand and supply. Indeed, Tang and Xiong (2010) find that the correlations of returns for contracts for which CITs take positions are higher than those for contracts in which CIT do not invest.

V. Empirical Implementation and Results

The model outlined in Section 3 yields predictions about the relationship between prices and trader positions (specifically, *I,* and *C*). Because we have daily observations on CIT and hedger positions in each maturity, we can directly test these hypotheses. We measure prices by the daily closing (settlement) prices on the Chicago Mercantile Exchange. As discussed above, price levels are more likely affected by changes in fundamentals than are price differences. Hence, the primary variable of interest in testing our model is the difference between the daily settlement prices of the first deferred contract and the nearby contract, which we refer to as the spread.[32]

The trader position variables are constructed from the daily position data in the CFTC Large Traders Reporting System database. The empirical counterpart of I is the maximum observed end-of-day position of CITs in each maturity contract. The maximum is typically reached 50-60 days before contract expiration. γ is the ratio of the end of day CIT position in the nearby to I. The empirical counterpart of N_k, the number of traders in category k for each maturity, is the maximum number of traders in category k for each maturity.[33]

In the model, C represents the physical quantity that hedgers will possess at some future date in the current year. As such, in the agricultural context, it is most appropriate to think of C as the post-harvest, cash-market long positions of these traders. Although the commercial traders in several of the LTRS categories are net short hedgers, we focus on the futures positions of the largest such category, agricultural distributors. These traders are

---

[32] For soy, the definition of the first deferred and nearby is somewhat ambiguous, in that while there are 7 contract maturities each year, only 5 of these have significant volume. In particular, CITs rarely trade the August and September soybean contracts; generally they roll their positions from the July contract to the November contract. In the results below, we consider the November contract as the first deferred when the July is the nearby, and treat it as the nearby from mid-July through its expiration. We have, however, checked the robustness of our results to defining the spread as the difference between the August and July maturities when July is the nearby; our results are unaffected.

[33] We take the maximum number under the logic that all of those traders could potentially trade on any given day, which corresponds to the notion of N_k in the model.

particularly relevant to our analysis, not only because they represent the largest category of commercial trader, but also because as a group, they are consistently short in the futures market – both over time and between commodities.[34]

As a result, our estimates of C are based on the observable futures positions of agricultural distributors, which by (10), bears a relationship to C. Specifically, from equation (10), we know that during period 0 (i.e., the pre-harvest period)

$$C = -\frac{(N_H + N_S)X_{1H}^2 + (1-\gamma)I_1}{N_H}$$

As Figure 2 indicates, hedgers begin to establish a futures position in the following year's post-harvest contract during the period in which the current year's post-harvest contract is the nearby. This suggests that hedgers begin accumulating positions in the following year's physical product (i.e., through planting and forward contracts) during the period in which this year's physical product is obtained. In the Appendix, we evaluate a more general model in which hedgers have physical positions in multiple periods. This enables us to derive expressions for both current and next year physical positions. The model implies that during post-harvest period, t, the two relevant cash positions are

$$C_t^{curr} = -\frac{\gamma I_t + (N_S + N_H)X_{t,H}^t}{N_S},$$

while

$$C_t^{next} = -\frac{(N_S + N_H)X_{t,H}^{t+\tau}}{N_S},$$

where τ is the number of periods each year.

It is important to note that we do not observe C^{curr} or C^{next} directly. The above relationships are derived from the model and we use them to approximate the hedgers' cash positions. Hence, our tests of the effect of C on spreads derive directly from the model.

In addition to CIT and hedger (distributor) positions, we would anticipate that spreads would also be affected by the period of time until expiration of the nearby contract, Λ, which also accounts for seasonality in the data. Finally, we would also like to test whether, holding the level of positions constant, the rate of change in CIT positions in the nearby (the *roll*) affects spreads. We measure the roll as the absolute value of the daily change in commodity index trader positions in the nearby contract.

---

[34] To be sure, there are many traders in other categories who behave similarly to the distributors. However, we choose not to reclassify traders into categories based on our perception of that trading; preferring instead to use the existing classifications established by the CFTC.

Before describing the details of how we test the model's predictions, we provide some discussion of the data.

A. Summary Statistics

Our data cover the period July 2003 until December 2008 and refer to daily observations. Table 4 reports descriptive statistics of the data; corn in panel I, soy[35] in panel II and wheat in panel III. The three products are similar in most respects. For example, our tests show that none of the variables have Gaussian distributions, although all are stationary. Average calendar spreads are more than 10 cents for all three products, indicating that the term structure of futures prices is typically upward sloping in our sample. All three spreads are also highly autocorrelated (the least autocorrelated is wheat, with a first-order autocorrelation of 0.83). Soybean calendar spreads are much more volatile than the other two products.

As discussed above, a key determinant of hedging costs is the relative size of I, CITs' long futures market position, and $C_t^{agg} = C_t^{curr} + C_t^{next}$, the physical (cash) positions held by hedgers. These two variables are of similar magnitude for corn, but I is about 3 times larger than C_t^{agg} for soybeans and wheat.[36] We also note that C_t^{agg} is more volatile than I for corn, but the reverse holds for the other two products. I and C_t^{agg} are highly autocorrelated.

The mean of γ is about 0.5 for all three products, indicating that, on average, CITs hold roughly half their positions in the nearby contract. This in turn suggests that the roll occurs roughly symmetrically around the middle of the period in which each maturity of each contract is the nearby. Finally, the last column of Table 4 reports summary statistics for the ratio of C_t^{curr} to C_t^{agg} which represents the percentage of total hedger cash position in the current year crop. It averages between 0.4 and 0.5 for the three products. That is, more than half of hedger cash positions are in the following year's harvest well before the current year's post-harvest contract reaches expiration.

The model provides additional testable predictions, which we examine in the following section.

---

[35] In the second quarter of 2004, the soybean spread was abnormally large and negative. We exclude this time period from our sample.

[36] Note that the estimate of C_t^{curr} here is based on distributor (hedger) positions only. As shown in Table 1, there is short hedging by traders in other categories, especially for soybeans and wheat. As such, total C_t^{curr} may be significantly larger than suggested by Table 3. Nevertheless, since distributors constitute the largest portion of these traders, the estimation should reflect the bulk of the changes in C_t^{curr}.

B. Testing the Model's Predictions

The spread exhibits serial correlation and heteroskedasticity. To mitigate the effect of these factors and to fully capture the dynamics of both the conditional mean and the conditional variance, we adopt the GARCH(1,1) specification which is very flexible and widely used for describing the evolution of financial variables.[37] More specifically, we estimate GARCH models with variance targeting (where the unconditional variance of the GARCH model is restricted to be equal to the sample unconditional variance). Francq, Horvath and Zakoïan (2009) show that when the model is misspecified, GARCH estimates with variance targeting are superior to unrestricted GARCH estimates.[38] We estimate the following model for the spread for each commodity j

$$S_{t,j} = \theta_{0,j} + \theta_{1,j} I_{t,j} + \theta_{2,j}\gamma_{t,j} + \theta_{3,j}\gamma_{t,j} I_{t,j} + \theta_{4,j} C_{t,j}^{curr} + \theta_{5,j} C_{t,j}^{curr} D_{t,j}^{pre-harv}$$

$$+ \ \theta_{6,j} C_{t,j}^{curr} D_{t,j}^{post-harv} + \theta_{7,j}\ddot{v}_{\ t,j} + \ \theta_{8,j} Roll_{t,j} + \theta_{9,j} C_{t,j}^{next} D_{t,j}^{post-harv} + \ \varepsilon_{t,j}$$

$$\varepsilon_{t,j} = \sqrt{h_{t,j}}\, u_{tji}$$

$$h_{t,j} = \omega_{0,j} + \omega_{1,j} h_{t-1,j} + \omega_{2,j}\varepsilon_{t-1,j}^{2} + \omega_{3,j}\ddot{v}_{\ t,j}$$

where $u_{t,j}$ is a sequence of independent and identically distributed (*i.i.d.*) random variables such that $E(u_{t,j}^2) = 1$; $D_{t,j}^{pre-harv}$ is a dummy variable which is equal to 1 when the pre-harvest is the nearby contract (September for corn and wheat, and July for soy); $D_{t,j}^{post-harv}$ is a dummy variable which is equal to 1 when the post-harvest is the nearby contract (December for corn and wheat, and November for soy). In the conditional variance equation ($h_{t,j}$), we add an additional term equal to the number of days until expiration ($\ddot{v}_{\ t,j}$) to account for the time pattern of prices as contracts move towards expiration.[39]

The estimation technique requires us to choose a distribution for $u_{t,j}$. Most GARCH models are estimated using a normal distribution. Unfortunately, the spreads here are highly non-normal with negative skewness and high kurtosis. The markets we analyze are typically characterized by spreads that are positive almost all the time (the spread is

---

[37] Hansen and Lunde (2005) compare over 300 volatility models and show that the GARCH(1,1) model well describes and well predicts the conditional variance of financial assets.

[38] In our empirical application we employ both the unrestricted GARCH model and the variance targeting GARCH and found the latter better describes the data in terms of likelihood ratio tests, Akaike and Schwartz information criteria.

[39] $D_{t,j}^{pre-harv}$, $D_{t,j}^{post-harv}$ and $\Lambda_{t,j}$ also account for seasonality in the data.

negative only 1.5 percent of the time for corn and wheat and 6.2 percent for soy). We, therefore, chose the generalized error distribution, which was introduced in the GARCH literature by Nelson (1991), since it accommodates the behavior of the spread in the tails.[40]

The theoretical model in the previous section makes a number of predictions about the parameter values and our goal is to test these predictions using the (reduced form) representation in the above equation. One prediction is that $\frac{\partial S}{\partial I} > 0$ for $\gamma = 0$ and $\frac{\partial S}{\partial I} < 0$ for $\gamma = 1$. This implies that $\theta_1 > 0$, and $\theta_3 < 0$, such that $\theta_1 + \theta_3 < 0$. Moreover, our model predicts a negative value on the coefficient of hedger cash positions (θ_4) for months in which the post-harvest contract is neither the nearby or the first deferred. When the last pre-harvest contract is the nearby, we would expect a larger (in absolute value) negative value for the coefficient on C_t^{curr}, so that $\theta_5 < 0$. For the post-harvest contract, we would expect $\frac{\partial S}{\partial C^{curr}} > 0$, so that $\theta_6 > 0$ and $\theta_6 + \theta_4 > 0$. Conversely, in the post-harvest period, we expect $\frac{\partial S}{\partial C^{next}} < 0$, so that $\theta_9 < 0$. These predictions are summarized in Table 3, while Table 5 reports estimation results.

For all commodities the signs of the estimated parameters in the spread equation are generally in line with the model's predictions, and are statistically significant. For example, the negative signs on θ_2 in the three regressions mean that the spread increases as CITs move their positions from the nearby to the first deferred.[41] Similarly, the negative signs on θ_4 and θ_9 (which are essentially our estimates of the effect of physical and forward cash positions) and the positive sign on θ_6 mean that the greater the extent to which hedgers seek to buy insurance (hedge their risks), the higher is the price they have to pay. The negative sign on θ_5 suggests that, consistent with the model, this effect is (absolutely) larger when the last pre-harvest contract is the nearby, since the correlations between the nearby futures contract and the hedgers' underlying positions are higher. In combination,

[40] We also employ a student t-distribution where we estimate the degrees of freedom. However, standard test statistics show that the generalized error distribution fits the data better than the t-distribution. The results are nevertheless quite similar.

[41] Stoll and Whaley calculate the change in spread between the beginning and end of the primary rolling period for the largest CITs. They find that the elasticity of the spread with respect to the change in CIT position in the nearby is about 0.0033 for wheat, 0.0066 for corn and -0.009 for soy. Our findings suggest larger elasticities; in the range of .05 – .13 for the January-May contracts. A related finding, due to Mou (2010), is that spreads are higher during the period in which the largest CIT rolls its position than prior to the roll. Unlike these authors, we have daily data on CIT and hedger positions, which enables us to estimate specific relationships; for example, the relationship between day-to-day changes in CIT positions and the associated price changes. It also allows us to test specific predictions of the model, such as the prediction that the effect of CIT positions on the spread will change with the harvest cycle.

these findings indicate that futures prices are determined by the sizes of hedger and CIT positions, as implied by the model and the preliminary results in Table 2. To get an idea of the magnitude of these effects, we note that a one standard deviation increase in the current-year cash position of hedgers (distributors) in soybeans leads to a decrease in the spread of nearly 1 cent if the nearby is the January, March, or May contracts, and close to a 3 cent decrease if the nearby is the July contract. Overall, the pattern of coefficients suggests that the model correctly interprets market behaviors.

Our model implies that the *level* of CIT positions in individual maturities should affect the price of that maturity, and that trading activity on a trading day (i.e., the change in position) would only affect prices if they introduced new information. As noted above, the roll-over strategies of CITs are announced well in advance and are unlikely to introduce any additional information to the market. Hence, we would expect θ_8, the coefficient of $Roll_t$, not to be statistically significant. For both soy and wheat, however, the rolling activity appears to increase the spread; the greater the daily increase in CITs' positions in the first deferred contract, the bigger is the spread. As such, the data suggests that trading activity does affect prices in these two markets. One possible explanation of this finding is that some of the roll may not be completely predictable (i.e., the rolling by CITs other than the major funds). Moreover, rolling strategies may reflect information about the desired CIT position in the first deferred contract (which may also change between maturities). These effects may be exacerbated in less liquid markets. This might explain why the effects are more significant in wheat and soybeans than in corn (which is the most liquid of the three).

There are also two cross-parameter restrictions implied by the theory: $\theta_1 + \theta_3 < 0$ and $\theta_4 + \theta_6 > 0$. We fail to reject the latter restriction at the 5 percent significance level for wheat and soy, but reject for corn. That is, for wheat and soy, the evidence supports the model's premise that when the first post-harvest contract is the nearby, the spread (that is, the price of the second post-harvest contract minus the price of the first post-harvest contract) increases with the size of hedger positions in the current-year crop. The evidence is less favorable for the first restriction.

The conditional variance equation is well-specified and stable with the sum of ω_1 and ω_2 less than unity. The parameter ω_3 is significant, indicating that there is seasonality in the second moment due to the life cycle of futures contracts. Although ω_3 is negative, the conditional variance is always positive. In line with the summary statistics in Table 4, the GED parameter is less than 2 for all commodities, implying that the spread has fat tails. Finally the R^2 indicates that the model well describes the evolution of the spread. This is particularly true for corn. Perhaps the lower R^2 for soybeans reflects the fact that soy traders have a broader set of instruments to use for hedging their risk, since futures and options markets also exist for soy meal and soy oil.[42]

VI. Conclusion

This paper analyzes the role of index traders in financial markets. Our perspective is that CITs fill the gap between short-hedge and long-hedge demand. That is, the prices that would have resulted from the trading of hedgers and traditional speculators alone allowed index traders to profitably take long futures positions. Consistent with this premise, we find that hedging costs fell as CITs positions grew.

Within this overall framework, there appeared to be additional opportunities for profitable trading due to the temporal mismatch between the contract maturities in which CITs are taking long positions, and the maturities in which hedgers take short positions. The evidence suggests that other market participants are able to profitably accommodate both of these groups by taking spread positions (short in the nearby, long in deferred contracts).

We show that a sizable portion of the inter-month spread can be explained by the sizes of the positions of CITs and hedgers. In particular, consistent with our theoretical model, increases in the size of hedgers' cash positions lead to lower spreads (at least in the early portions of the harvest cycle). This reflects the idea that the price of assets that are highly correlated with hedgers' cash market positions more closely track changes in those positions. In addition, we find that CITs relative positions in different maturities affect the relative prices of those maturities in predictable ways. As such, the findings suggest that our model explains important aspects of the trading behavior of various agents in the market, and how their trading has reacted to changes in the size of CIT positions in futures markets.

[42] This may explain why agricultural distributors represent a much smaller percentage of open interest for soybeans than for the other two products. Some evidence of the use of such cross-hedging can be found in Brunetti and Reiffen (2012).

Underlying the premise of our model is the more general notion that traders are willing to take on additional risk only in exchange for higher compensation. By tracking the behavior of groups of similarly-situated traders, we document that traders behave consistent with these models of finite liquidity. That is, it appears that, observed price effects from changes in the demand and supply for insurance against price risk can be explained by the higher cost (in terms of portfolio risk) incurred by speculators. This in turn implies that observed changes in spreads are not necessarily opportunities for arbitrage profits.

References:

Acharya, V., L. Lochstoer, and T. Ramadorai, (2010). "Limits to Arbitrage and Hedging: Evidence from Commodity Markets." Working Paper.

Brunetti, C., and D. Reiffen, (2012). "Does Hedging Activity Reflect Cash Market Positions? Evidence from Agricultural Markets." Working Paper.

Büyükşahin, B., M. S. Haigh, J. H. Harris, J. A. Overdahl, and M. A. Robe, (2009). "Fundamentals, Trading Activity and Derivative Pricing." Working Paper.

Commodity Futures Trading Commission, (2006). "Comprehensive Review of the Commitment of Traders Reporting Program." Federal Register 71, FR 35267.

——— (2008). "Staff Report on Commodity Swap Dealers & Index Traders with Commission Recommendations." Washington, DC.

Dewally, M., L. Ederington, and C. Fernando, (2010). "Determinants of Trading Profits of Individual Traders: Risk Premia or Information." Working Paper.

Dhillon, U., and H. Johnson, (1991). "Changes in the Standard and Poor's 500" *Journal of Business* **64**: 75-85.

Etula, E. (2010) "Broker-Dealer Risk Appetite and Commodity Returns." Working Paper.

Fishe, R. P. H., and A. D. Smith, (2010). "Identifying Informed Traders in Futures Markets." Working Paper.

Francq, C., L. Horvath, and J-M. Zakoïan, (2009). "Merits and Drawbacks of Variance Targeting in GARCH Models." Working Paper.

Gorton, G., F. Hayashi, and K. G. Rouwenhorst, (2007). "The Fundamental of Commodity Futures Returns." Working Paper.

Greenwood, R., (2005). "Short- and Long-term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage." *Journal of Financial Economics* **75**: 607-649.

Grossman, S., and M. Miller, (1989). "Liquidity and Market Structure." *Journal of Finance* **43**: 617-633.

Hansen, P.R., and A. Lunde, (2005). "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" *Journal of Applied Econometrics* **20:** 873-889.

Harris, L., and E. Gurel, (1986). "Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures." *Journal of Finance* **41**: 815-29.

Hartzmark, M. L., (1987). "Returns to Individual Traders of Futures: Aggregate Results." *Journal of Political Economy* **95**: 1292-1306.

Hicks, J. R., (1939). *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory*. London: Oxford University Press.

Hirshleifer, D. A., (1988). "Residual Risk, Trading Costs, and Commodity Futures Risk Premia." *Review of Financial Studies* **1**: 173-193.

——— (1990). "Hedging Pressure and Futures Price Movements in a General Equilibrium Model." *Econometrica* **58**: 411-428.

Jain, P. C., (1987). "The Effect on stock price of inclusion in or exclusion from the S & P 500." *Financial Analysts Journal* 43:58-65.

Irwin, S. H., and D. R. Sanders, (2010). "The Financialization of Commodity Futures Markets or: How I Learned to Stop Worrying and Love Index Funds." Working Paper.

Keynes, J. M., (1930). *Treatise on Money*. Vol. 2, London: MacMillan & Co.

Masters, M., (2008). "Testimony Before Senate Committee on Homeland Security and Governmental Affairs." May 20th, 2008.

Mou, Y., (2010). " Limits to Arbitrage and Commodity Index Investment: Front-Running the Goldman Roll." Working Paper.

Lux, T., (1995). "Herd Behaviour, Bubbles and Crashes." *The Economic Journal* **105**: 881-896.

Nelson, D., (1991). "Conditional Heteroscedasticity in Asset Returns: a New Approach." *Econometrica* **59**: 347-370.

Shiller, R. J., (2003). "From Efficient Markets Theory to Behavioral Finance." *Journal of Economic Perspectives* **17**: 83-104.

Shleifer, A., (1986). "Do Demand Curves for Stocks Slope Down?" *Journal of Finance* **41**:579- 90.

Singleton, K., (2011). "Investor Flows and the 2008 Boom/Bust in Oil Prices." Working Paper.

Shleifer, A., and L. H. Summers, (1990). "The Noise Trader Approach to Finance." *The Journal of Economic Perspectives* **4**: 19-33.

Stoll, H. and R. Whaley, (2010). "Commodity Index Investing and Commodity Futures Prices." *Journal of Applied Finance* **20**: 7-46.

Tang, K., and W. Xiong, (2010). "Index Investment and Financialization of Commodities." Working Paper.

Varian, H.. (2005). "Bootstrap Tutorial." *Mathematica Journal* **9**: 768-775.

Table 1

Large Trader Reporting System – Average Participation Rate by Category

	Categories of Traders	Corn		Soy		Wheat	
		Ave. # of Traders	*Ave. Share of Open Interest (%)*	*Ave. # of Traders*	*Ave. Share of Open Interest (%)*	*Ave. # of Traders*	*Ave. Share of Open Interest (%)*
Commercial	Agri. Distributors (AD)	244.5	-31.59	111.3	-22.82	60.43	-31.06
	Agri. Manufacturers (AM)	50.93	1.052	25.77	-4.905	17.55	-3.881
	Agri. – Other (AO)	39.41	2.114	14.21	-0.144	5.383	-0.351
	Agri. Producers (AP)	34.24	2.433	13.97	-0.731	7.501	-0.517
	Swap Dealers (AS)	7.613	0.016	5.921	0.603	7.862	-1.532
Non-Commercial	Floor Traders (FBT)	90.92	-0.336	78.40	-0.311	55.06	-0.531
	Regis. Managed Money (MMT)	117.3	5.287	112.6	7.983	108.5	-2.738
	Other Managed Money (NRP)	132.7	0.553	115.9	-0.102	68.68	-0.721
	Not classified (NC)	2.952	0.034	2.402	0.098	1.654	0.361
	Index Traders (CIT)	19.57	27.11	18.39	26.52	19.34	45.81

Table 2
Hedging Costs

	Corn		Soy		Wheat	
	Hedging Cost	Volatility of Hedging Cost	Hedging Cost	Volatility of Hedging Cost	Hedging Cost	Volatility of Hedging Cost
Dependent Variable Mean (basis points)	2.5700	3.1351	7.5012	2.6832	5.7921	3.3372
Constant	0.0139 (0.0203)	0.0253*** (0.0023)	-0.0141 (0.0300)	0.0306*** (0.0067)	0.0339 (0.0290)	0.0301*** (0.0056)
I_t	-2.35e-7* (1.66e-7)	-3.01e-8*** (1.94e-8)	-1.01e-8 (2.91e-7)	-4.84e-6*** (6.66e-8)	-3.99e-7* (2.80e-7)	-3.47e-8* (1.99e-8)
C_t	1.74e-7* (9.16e-8)	5.63e-8* (1.45e-8)	2.95e-7 (5.39e-7)	2.96e-8 (1.16e-7)	6.09e-7* (4.57e-7)	1.51e-8 (1.12e-7)
R^2	0.1351	0.3109	0.0133	0.0217	0.0931	0.1761

Bootstrapped standard errors in parenthesis. Asterisks indicate significance at 20% (*), 5% (**) and 1% (***). The number of obs. in each regression is 27.

Table 3
Predicted Sign From the Model and Restrictions

Coefficient	Predicted Sign
θ_1	+
θ_2	–
θ_3	–
θ_4	–
θ_5	–
θ_6	+
θ_7	+
θ_8	0
θ_9	–
Restrictions	
$\theta_1 + \theta_3$	–
$\theta_4 + \theta_6$	+

Table 4

Summary Statistics

Panel I: Corn	*Spread*	I	γ	C^{agg}	C^{curr}/C^{agg}
Mean	0.1083	212652	0.4649	215051	0.4037
Median	0.1100	241178	0.5171	211058	0.4295
Std. Dev.	0.0436	92853	0.4086	154297	0.2853
Skewness	-0.3988	-0.1710	-0.0058	0.1987	0.0592
Kurtosis	3.6454	2.1653	1.1995	1.6692	2.0305
Jarque-Bera	0.0000	0.0000	0.0000	0.0000	0.0022
ADF	0.0001	0.0789	0.0000	0.0654	0.0025
AC(1)	0.9677	0.9967	0.9445	0.9942	0.9193
AC(10)	0.8144	0.9669	0.2463	0.9068	0.0732
AC(50)	0.4783	0.8418	0.2314	0.6715	-0.2692

Panel II: Soy	*Spread*	I	γ	C^{agg}	C^{curr}/C^{agg}
Mean	0.1041	93759	0.4812	29207	0.4961
Median	0.1200	107309	0.5656	23861	0.6222
Std. Dev.	0.0916	38808	0.4185	26666	0.4238
Skewness	-0.6602	-0.5013	-0.0516	0.8185	-0.1167
Kurtosis	5.9485	1.7862	1.1748	2.8292	1.2295
Jarque-Bera	0.0000	0.0000	0.0000	0.0000	0.0000
ADF	0.0035	0.1148	0.0000	0.0295	0.0000
AC(1)	0.9468	0.9973	0.9443	0.9912	0.8794
AC(10)	0.7390	0.9726	0.2399	0.8550	0.4509
AC(50)	0.2535	0.8729	-0.1735	0.2451	-0.0815

Panel III: Wheat	*Spread*	I	γ	C^{agg}	C^{curr}/C^{agg}
Mean	0.1296	118198	0.4603	31935	0.4489
Median	0.1300	131696	0.5186	22308	0.4903
Std. Dev.	0.0552	50808	0.4087	30290	0.4222
Skewness	-0.6284	-0.5054	0.0089	1.5463	0.0682
Kurtosis	8.6866	1.8854	1.1993	6.0809	1.2252
Jarque-Bera	0.0000	0.0000	0.0000	0.0000	0.0000
ADF	0.0001	0.0776	0.0000	0.0566	0.0001
AC(1)	0.8277	0.9954	0.9409	0.9948	0.8766
AC(10)	0.5858	0.9772	0.2097	0.8931	0.2746
AC(50)	0.3198	0.8913	0.2129	0.4508	0.0453

Jarque-Bera refers to the probability that the distribution of the variable is normal, using the Jarque-Bera normality test (i.e., the null hypothesis is that of normality). ADF refers to the probability that the variable is non-stationary, using the Augmented Dickey-Fuller test, (i.e., where the null hypothesis is that of non-stationarity). AC(w) refers to the autocorrelation at lag w. I refers to the CIT positions; γ denotes the percentage of the CIT positions remaining in the nearby contract; C^{agg} is the aggregate cash position in the underlying product held by hedgers in both the this year's and next year's crop; C^{curr}/C^{agg} is the percentage of the hedgers' cash position in the current year's crop during the post-harvest period; *Roll* is the absolute value of the daily change in commodity index trader positions in the nearby contract.

Table 5

Estimation Results – Main Model GARCH(1,1)

Conditional Mean	Corn	Soy	Wheat
θ_0	0.0457***	-0.0026	0.0781***
	(0.0010)	(0.0017)	(0.0018)
$\theta_1(I_t)$	2.74e-7***	1.26e-6***	3.15e-7***
	(5.05e-9)	(1.92e-8)	(1.91e-8)
$\theta_2(\gamma_t)$	-0.0271***	-0.0152***	-0.0097**
	(0.0018)	(0.0028)	(0.0040)
$\theta_3(\gamma_t I_t)$	-2.83e-8***	-3.25e-8	-4.76e-8**
	(7.21e-9)	(2.95e-8)	(2.21e-8)
$\theta_4(C_t^{curr})$	-5.25e-8***	-3.23e-7***	-1.35e-7***
	(2.95e-9)	(4.17e-8)	(5.20e-8)
$\theta_5(C_t^{curr} D_t^{pre-harvest})$	-1.67e-7***	-6.83e-7***	-7.92e-7***
	(3.04e-9)	(5.48e-8)	(4.31e-8)
$\theta_6(C_t^{curr} D_t^{post-harv})$	5.41e-8***	3.81e-7***	3.12e-7***
	(1.14e-8)	(4.44e-8)	(4.82e-8)
$\theta_7(\Lambda_t)$	2.58e-4***	-5.92e-5	-2.29e-5
	(2.82e-5)	(3.12e-5)	(5.75e-5)
$\theta_8(Roll_t)$	-3.73e-5	1.29e-6***	1.17e-6***
	(4.30e-5)	(1.13e-7)	(1.11e-7)
$\theta_9(C_t^{next} D_t^{post-harv})$	-2.02e-7***	-2.00e-7	-1.84e-6***
	(1.10e-8)	(1.36e-7)	(1.47e-7)
Conditional Variance			
ω_0	1.67e-5	8.93e-5	2.10e-5
$\omega_1(h_{t-1})$	0.6175***	0.6560***	0.8082***
	(0.0195)	(0.0242)	(0.0081)
$\omega_2(\varepsilon_{t-1}^2)$	0.3619***	0.3301***	0.1818***
	(0.0192)	(0.0247)	(0.0075)
$\omega_3(\Lambda_t)$	-2.16e-7***	-7.60Ee-7***	-3.07e-7***
	(3.99e-8)	(1.35e-7)	(5.87e-8)
GED	1.7173***	1.3381***	1.2790***
	(0.0970)	(0.0906)	(0.0342)
Restrictions:			
$\theta_1 + \theta_3 < 0$	2.36e-7	1.22e-6	2.68e-7
	(7.03e-9)	(1.83e-8)	(1.48e-8)
$\theta_4 + \theta_6 > 0$	1.57e-9	5.76e-8†	1.78e-7†
	(1.11e-8)	(2.00e-8)	(2.79e-8)
R^2	0.5743	0.2338	0.3137
Log-Lik	3568.4	2620.5	2766.8
# of Obs.	1338	1203	1304

Standard errors in parenthesis. Asterisks indicate significance at 5% (**) and 1% (***), respectively. † indicates fail to reject the restriction. The estimated model is

$$S_t = \theta_0 + \theta_1 I_t + \theta_2\gamma_t + \theta_3\gamma_t I_t + \theta_4 C_t^{curr} + \theta_5 C_t^{curr} D_t^{pre-harv} + \theta_6 C_t^{curr} D_t^{post-harv} + \theta_7 \ddot{v}_t$$
$$+ \theta_8 Roll_t + \theta_9 C_t^{next} D_t^{post-harv} + \varepsilon_t$$
$$\varepsilon_t = \sqrt{h_t}u_t$$
$$h_t = \omega_0 + \omega_1 h_{t-1} + \omega_2\varepsilon_{t-1}^2 + \omega_3 \ddot{v}_t$$

I_t refers to the CIT positions; γ_t denotes the percentage of the CIT position in the nearby contract; C_t^{curr} is this year (current) cash position in the underlying product held by hedgers; C_t^{next} is next year cash position in the underlying product held by hedgers; $Roll_t$ indicates the amount of roll-over by CIT and is computed as the absolute value of the daily change in commodity index trader positions in the nearby contract; Λ_t is the number of days until contract expiration; $D_t^{pre-harv.}$ and $D_t^{post-harv.}$ are dummy variables indicating the pre-harvest and the post-harvest contracts, respectively.

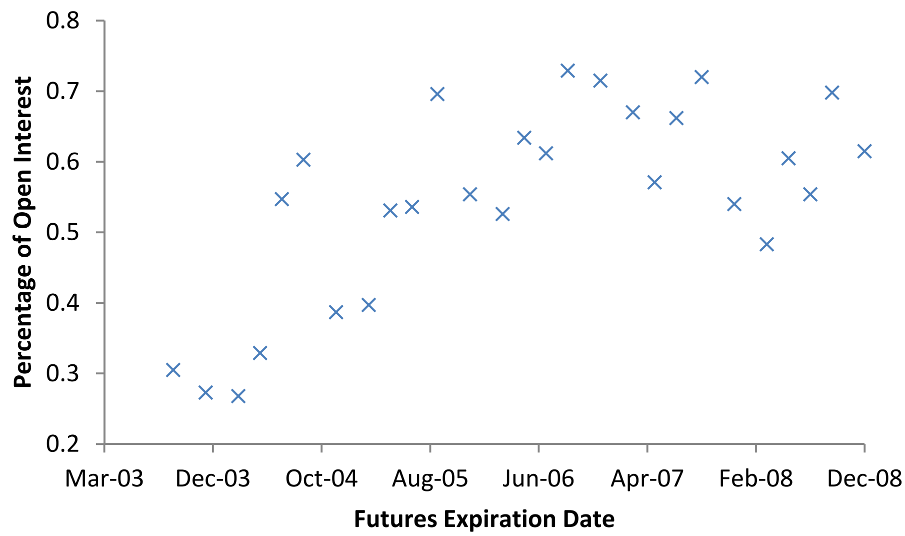

Figure 1
Corn: Aggregate CIT positions in the nearby contract as a percentage of open interest.

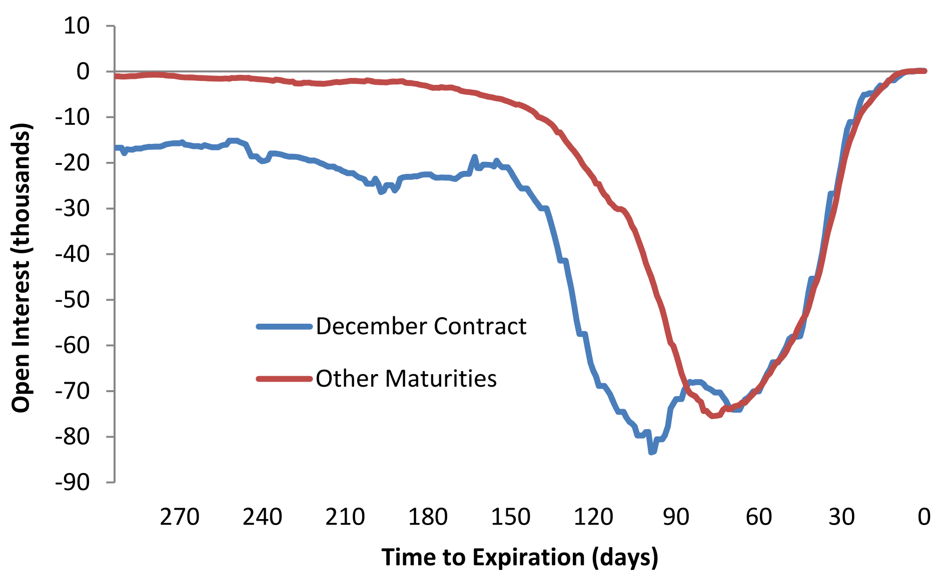

Figure 2
Wheat: Distributor average positions in the December contracts and in the other maturities (excluding December). On the horizontal axis, 0 is the expiration of the contract.

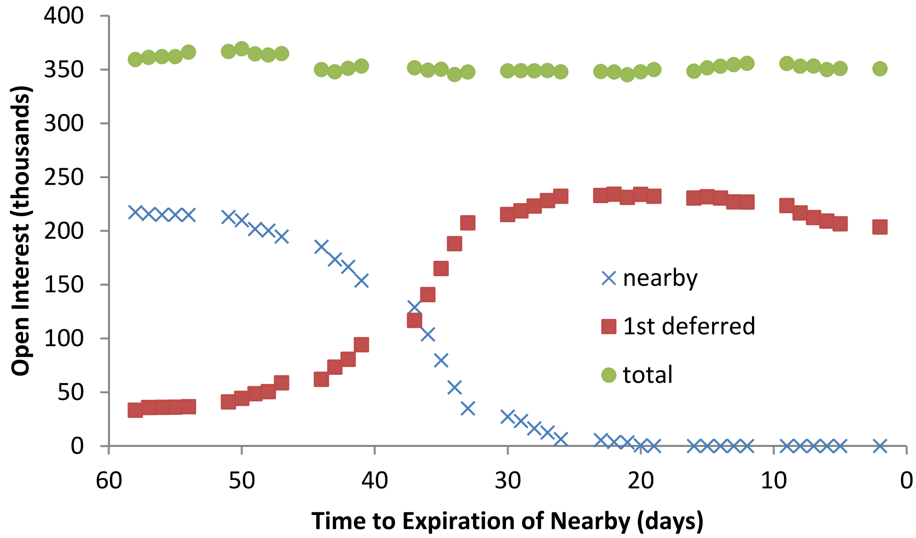

Figure 3A
Corn: CIT positions in the May 2007contract (nearby), July 2007 contract (1st-deferred) and sum of the two (total). On the horizontal axis, 0 is the expiration of the contract.

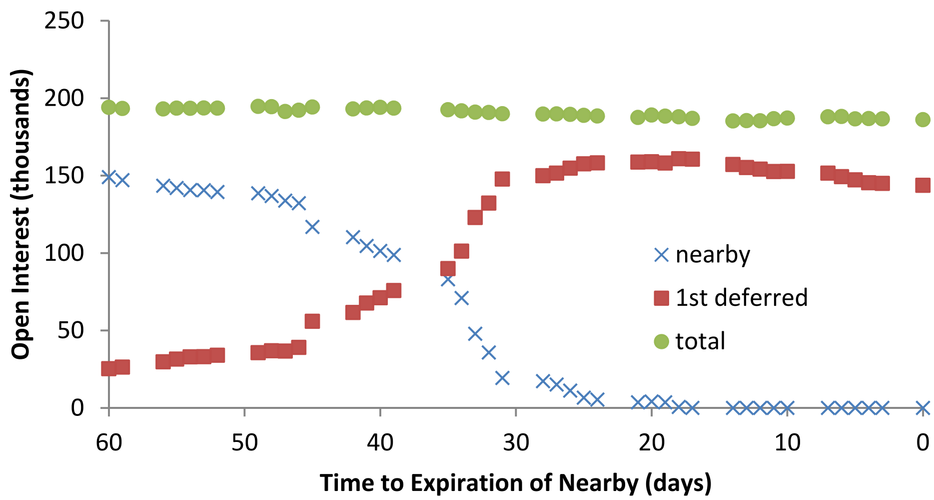

Figure 3B
Wheat: CIT positions in the May 2007contract (nearby), July 2007 contract (1st-deferred) and sum of the two (total). On the horizontal axis, 0 is the expiration of the contract.

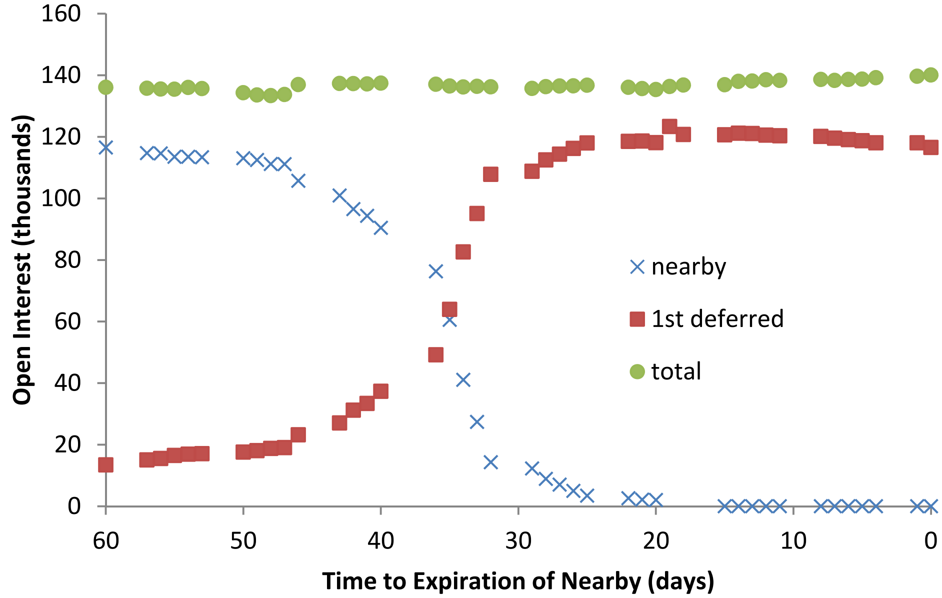

Figure 3C
Soy: CIT positions in the May 2007contract (nearby), July 2007 contract (1st-deferred) and sum of the two (total). On the horizontal axis, 0 is the expiration of the contract.

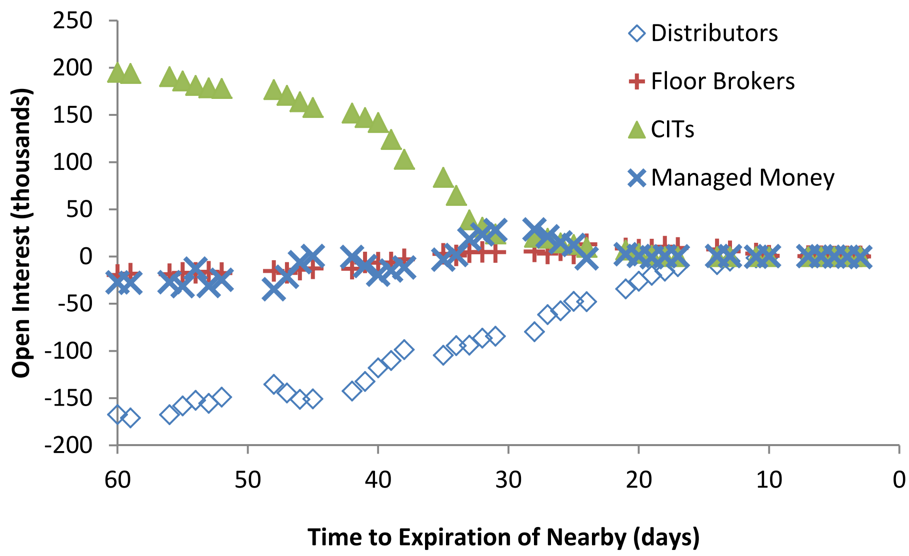

Figure 4
Corn: Average positions of distributors, floor brokers, commodity index traders (CITs) and managed money traders. On the horizontal axis, 0 is the expiration of the contract.

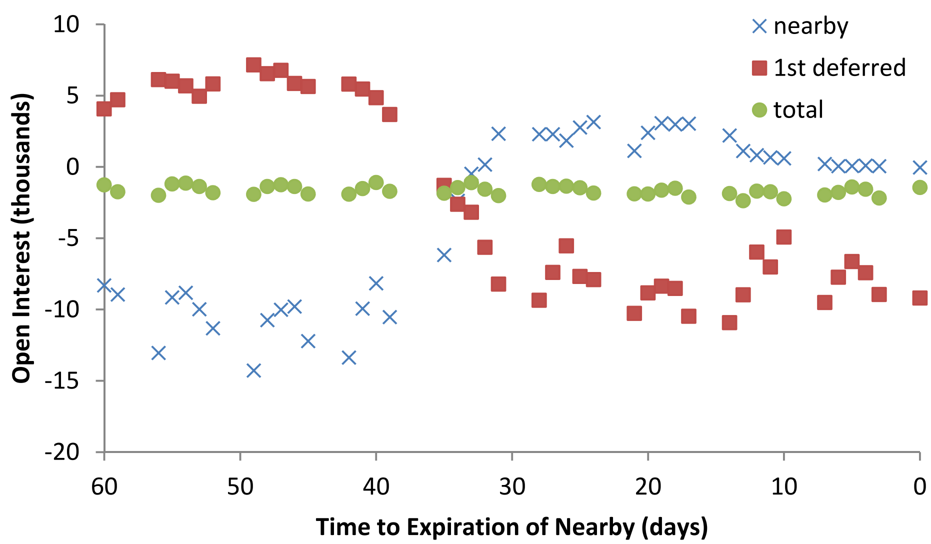

Figure 5
Wheat: Average floor broker (market maker/locals) positions in the nearby contract, 1st-deferred contract, and across all maturities (total). On the horizontal axis, 0 is the expiration of the contract.

Timing of Decisions

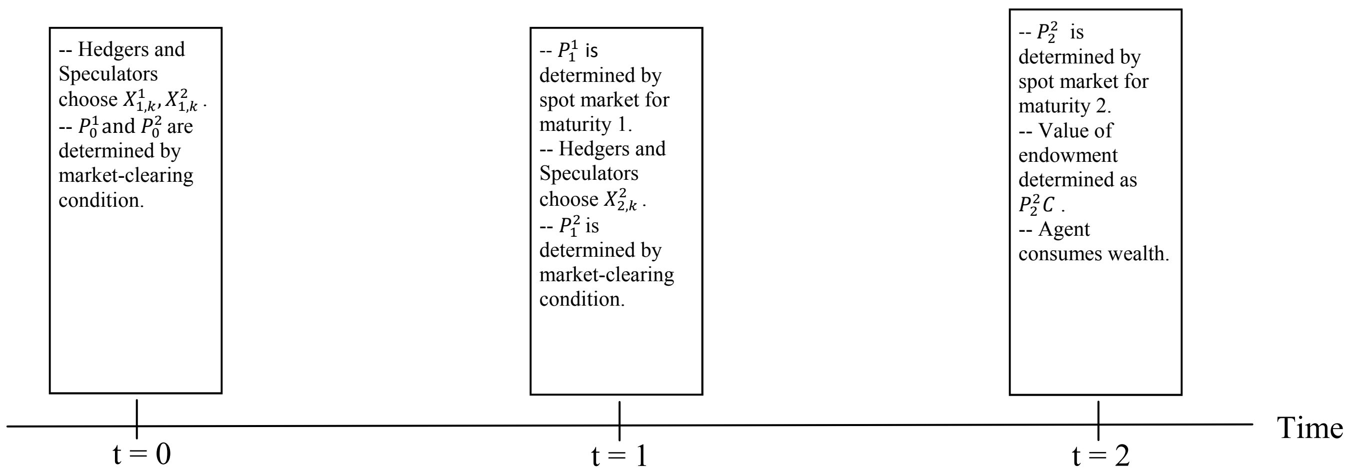

Figure 6

Appendix:

In this Appendix, we generalize the model through two extensions, in order to derive a richer set of implications for prices. One artifact of the two-period model in the text is that index traders are active in all traded contracts in all periods. Generalizing the model to three-periods allows us to examine what happens when the maturities that hedgers wish to trade differ from those for which index traders take positions. As in the two period model, we assume hedgers each have an endowment of the physical product of size C that they will receive in the final period, and the price they will receive for the physical will be determined at that time. The optimization at $t = 0$ is now

$$U[W_0 + X_2^3(P_2^3 - P_1^3) + X_3^3(P_3^3 - P_2^3) + X_2^2(P_2^2 - P_1^2) + (P_3^3)C_k + X_1^2(P_1^2 - P_0^2) + X_1^3(P_1^3 - P_0^3) + X_1^1(P_1^1 - P_0^1)]$$

Using the same kind of backward induction as in the text, the equilibrium at $t = 2$ and $t = 3$ are the same as those in equations (6), (9) and (10) (recognizing that the terminal period is now $t = 3$ rather than $t = 2$, and the period preceding the final period is now $t = 2$, rather than $t = 1$). There are now 3 contracts traded at $t = 0$, and some additional notation is required to adjust to this change. Specifically, let I_i^j be index trader positions in contract j at time i (so that $I_i = \sum_{j=1}^{3} I_i^j$). Each trader's first-order conditions with respect to their positions in the three contracts (X_1^1, X_1^2 and X_1^3) are now

$$(\sigma_1^1)^2 X_{1,k}^1 + \sigma_1^{12} X_{1,k}^2 + \sigma_1^{13} X_{1,k}^3 = \frac{E_0[P_1^1 - P_0^1]}{\alpha} - \sigma_{13}^{13}(X_{3,k}^3 + C_k) - \sigma_{12}^{13}(X_{2k}^3 - X_{3k}^3) - \sigma_{12}^{12} X_{2k}^2 + \sigma_1^{12} X_{2k}^2 + \sigma_1^{13} X_{2k}^3 \qquad (A.1)$$

$$\sigma_1^{12} X_{1,k}^1 + (\sigma_1^2)^2 X_{1,k}^2 + \sigma_1^{23} X_{1,k}^3 = \frac{E_0[P_1^2 - P_0^2]}{\alpha} - \sigma_{13}^{23}(X_{3,k}^3 + C_k) - \sigma_{12}^{23}(X_{2k}^3 - X_{3k}^3) - \sigma_{12}^{2} X_{2k}^2 + \sigma_1^{23} X_{2k}^3 + (\sigma_1^2)^2 X_{2k}^2 \qquad (A.2)$$

$$\sigma_1^{13} X_{1,k}^1 + \sigma_1^{23} X_{1,k}^2 + (\sigma_1^3)^2 X_{1,k}^3 = \frac{E_0[P_1^3 - P_0^3]}{\alpha} - \sigma_{13}^{3}(X_{3,k}^3 + C_k) - \sigma_{12}^{3}(X_{2k}^3 - X_{3k}^3) - \sigma_{12}^{23} X_{2k}^2 + \sigma_1^{23} X_{2k}^2 + (\sigma_1^3)^2 X_{2k}^3 \qquad (A.3)$$

From the analysis in the text, we know that $X_{i,S}^3 = X_{i,H}^3 + C$ for $i = 2{,}3$, and $X_{2,S}^2 = X_{2,H}^2$ which means that the terms involving $X_{i,k}^j$ on the right-hand sides of equations (A.1) –

(A.3), are each the same for hedgers and market-makers, with the exception of $X_{i,k}^3$ (which differs by C between S and H). Solving this system of equations shows that $X_{1,S}^3 = X_{1,H}^3 + C$, and $X_{i,S}^j = X_{i,H}^j$ for j = 1,2. We can then use the market clearing conditions to find that $X_{1,H}^3 = \frac{-(I_1^3+N_S C)}{N_S + N_H}$ and $X_{1,H}^j = \frac{-I_1^j}{N_S + N_H}$, j = 1,2 (while noting that $I_1^3 = 0$ by assumption). Hence, we find that unlike the case when there are only two periods, some of the X_1^j are functions of C, while others are functions of I_1^j.

Two implications of this analysis are that in that we would expect the spread to be decreasing in C, and that this conclusion will be true under more general conditions in period 1 than period 0. In that sense, the prediction is that $\frac{\partial S_0}{\partial C}$ will be smaller (i.e., more negative) in post-harvest periods

$$\frac{\partial S_0}{\partial C} =$$

$$\frac{1}{N_H + N_S}[(\sigma_1^{13} - \sigma_1^{23})N_S + (N_H)(\sigma_{13}^{13} - \sigma_{13}^{23})R_2 + N_S R_1(\sigma_1^{23} - \sigma_1^{13}) - N_S(R_1 - R_2)(\sigma_{12}^{13} - \sigma_{12}^{23})]$$

$$\frac{\partial S_1}{\partial C} = \frac{1}{N_H + N_S}[((\sigma_2^3)^2 - \sigma_2^{23})N_S + (N_H)(\sigma_{23}^{23} - \sigma_{23}^3)R_2 + N_S R_2(\sigma_2^{23} - (\sigma_2^3)^{2)}]$$

$$= \frac{1}{N_H+N_S}[N_S(R_2 - 1)(\sigma_2^{23} - (\sigma_2^3)^2 + N_H(\sigma_{23}^{23} - \sigma_{23}^3).$$

$\frac{\partial S_1}{\partial C}$ will be negative if $R_2 \geq 1$ and $\sigma_{23}^3 > \sigma_{23}^{23}$. The latter condition seems reasonable, since we would expect the correlation between movements in the period 2 and period 3 futures prices would likely be tighter between a single maturity than adjacent maturities. $\frac{\partial S_1}{\partial C}$ will be negative under similar conditions, but requires additional assumptions. In that sense, the prediction that $\frac{\partial S_1}{\partial C} < 0$ is stronger than the prediction that $\frac{\partial S_0}{\partial C} < 0$. In our empirical work, we allow the effect of hedger's cash positions on the spread to vary over the harvest cycle.

The other extension we wish to consider involves multiple endowments. The basic model in Section IV features an endowment which is realized (i.e., priced) as of the final period. As such, it does not allow one to consider pricing behavior over the course of a period following the valuation of the endowment. To examine that, we modify the model to allow hedgers to have endowments in two periods (t =1 and t = 3). This corresponds to an environment in which the market has a periodic change in supply or demand (e.g., an annual harvest).[43] Specifically, suppose the hedger has a crop that will mature in period 1 (C^{curr}) as well as period 3 (C^{next}), and also consumes some of his wealth in period 1 (φ). We can think of periods 1 and 3 as the post-harvest periods in consecutive years. We now can write his utility function as

$$U\big[W_0 + \; X_2^3(P_2^3 - P_1^3) + X_3^3(P_3^3 - P_2^3) + X_2^2(P_2^2 - P_1^2) + (P_3^3)C_k^{next} + X_1^2(P_1^2 - P_0^2) + X_1^3(P_1^3 - P_0^3) + X_1^1(P_1^1 - P_0^1) + \; P_1^1 C_k^{Curr} - \varphi_k\big] + U[\varphi_k]$$

The optimization at t = 2 is identical to that characterized in equation (1), except that hedger's wealth at t = 3 increases by $P_1^1 C_k^{curr}$ and decreases by φ. Since those change do not affect the first-order conditions at t = 1 and t = 2, they do not alter the optimal choice of X_i^j in those periods, and hence do not affect P_j^i in those periods.

The choice of φ at $t = 1$ involves a first-order condition of the general form

$$\int \exp(-(\alpha(W_0 - \varphi) + P_3^3 C))f(P_3^3)dP_3^3 = \exp -(\alpha\varphi)$$

(i.e., equating the marginal utility of income in the two periods).

Finally, consider the choice of positions at t = 0. The first-order conditions with respect to positions are similar to those in equations (A.1) – (A.3);

$$(\sigma_1^1)^2 X_{1,k}^1 + \; \sigma_1^{12} X_{1,k}^2 + \; \sigma_1^{13} X_{1,k}^3 \; = \; \frac{E_0[P_1^1 - P_0^1]}{\alpha} - \sigma_{13}^{13}(X_{3,k}^3 + C_k^{next}) - \sigma_{12}^{13}(X_{2k}^3 - X_{3k}^3) - \sigma_{12}^{12} X_{2k}^2 + \sigma_1^{12} X_{2k}^2 + \sigma_1^{13} X_{2k}^3 - (\sigma_1^1)^2 C_k^{curr} \qquad \text{(A.1')}$$

$$\sigma_1^{12} X_{1,k}^1 + \; (\sigma_1^2)^2 X_{1,k}^2 \; + \; \sigma_1^{23} X_{1,k}^3 = \; \frac{E_0[P_1^2 - P_0^2]}{\alpha} - \sigma_{13}^{23}(X_{3,k}^3 + C_k^{next}) - \sigma_{12}^{23}(X_{2k}^3 - X_{3k}^3) - \sigma_{12}^{2} X_{2k}^2 + \sigma_1^{23} X_{2k}^3 + (\sigma_1^2)^2 X_{2k}^2 - \sigma_1^{12} C_k^{curr} \qquad \text{(A.2')}$$

$$\sigma_1^{13} X_{1,k}^1 + \; \sigma_1^{23} X_{1,k}^2 \; + \; (\sigma_1^3)^2 X_{1,k}^3 = \; \frac{E_0[P_1^3 - P_0^3]}{\alpha} - \sigma_{13}^{3}(X_{3,k}^3 + C_k^{next}) - \sigma_{12}^{3}(X_{2k}^3 - X_{3k}^3) - \sigma_{12}^{23} X_{2k}^2 + \sigma_1^{23} X_{2k}^2 + (\sigma_1^3)^2 X_{2k}^3 - \sigma_1^{13} C_k^{curr} \qquad \text{(A.3')}$$

As above, we can use the relationship $X_{i,H}^3 = X_{i,S}^3 - C^{next}$ for i = 2,3, along with (A.1') – (A.3') to derive the relationship between $X_{1,H}^j$ and $X_{1,S}^j$: $X_{1,H}^1 = X_{1,S}^1 - C^{curr}$, $X_{1,H}^2 = X_{1,S}^2$ and $X_{1,H}^3 = X_{1,S}^3 - C^{next}$,

[43] While this interpretation would be more complete with a 4-period model (so that endowments are receiving every second period), no additional insight would be obtained from that model, while it would add considerable notational clutter.

It follows that using the market-clearing condition and (A.1') - (A.3'), we can express prices in terms of the exogenous variables, I, C^{next} and C^{curr}.

This yields one additional implication: the spread in period 0 will be is increasing in C^{curr}:

$$\frac{\partial S_0}{\partial C^{curr}} = [(\sigma_1^1)^2 - \sigma_1^{12}\left[1 - \frac{N_H}{N_H + N_S}\right] > 0.$$

That is, the spread in a post-harvest period will be increasing in hedger's current-year physical positions.