Futures Market Failure?

Philip Garcia, Scott H. Irwin, and Aaron Smith[*]

October 2011

Abstract

In a well-functioning futures market, the futures price on the expiration date equals the price of the underlying asset on that date. This basic condition failed to hold in corn, soybeans, and wheat for most futures contract expirations between 2005 and 2010. During this episode, futures contracts expired at prices up to 35 percent greater than the prevailing cash grain price. Using a rational expectations commodity storage model, we demonstrate how such non-convergence can be produced in equilibrium by the institutional structure of the delivery market. Specifically, we show how a wedge between the price of storing the physical commodity and the cost of carrying the delivery instrument causes non-convergence. This wedge can be conceptualized as the coupon in a present value equation, so our results provide an example of a significant and complex market in which a present value model explains prices.

Keywords: finance; agriculture; commodity futures; delivery; grain; storage; present value; asset pricing;

JEL Categories: G13, G14, Q11, Q13

---

[*] Philip Garcia is the T.A. Hieronymus Distinguished Chair in Futures Markets and Director of the Office for Futures and Options Research in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. Scott H. Irwin is the Laurence J. Norton Professor of Agricultural Marking in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. Aaron Smith is an Associate Professor in the Department of Agricultural and Resource Economics at the University of California at Davis. We thank Peter Koziara and Gene Kunda for assistance with data collection. Pat Fishe and Dwight Sanders provided useful feedback on the manuscript. We also thank conference participants at the 2011 CFTC Conference on Commodity Markets and the 2011 NCCC-134 Conference, and seminar participants at the UC-Davis Department of Agricultural and Resource Economics, Iowa State University Department of Economics, the Economic Research Service of the USDA, and the Chicago Mercantile Exchange for helpful comments. This material is based upon work supported by Cooperative Agreement with the Economic Research Service, U.S. Department of Agriculture, under Project Numbers 58-3000-9-0078 and 58-3000-9-0076. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture. Corresponding author: Aaron Smith, 3103 Social Sciences and Humanities, University of California, Davis, One Shields Avenue, Davis, CA 95616, email: adsmith@ucdavis.edu.

I. INTRODUCTION

In theory, the price of an asset equals the present discounted value of expected future payoffs. In reality, the link between prices and future payoffs often seems tenuous, which leads to assertions of bubble patterns in prices and debates about the efficacy of the theory. Such assertions and debates have a long history, dating back at least to the infamous South Sea Bubble of 1720. Harrison (2001) shows that analysts in the early eighteenth century used a dividend discount model to calculate the intrinsic value of the South Sea Company, but struggled to reconcile observed price movements with news about future dividend payments. More recently, boom and bust cycles in Internet stocks in 1999–2001, real estate in 2003–2006, and commodities in 2007–2008 have incited debates about the extent to which rational expectations can justify observed asset prices.

Futures prices for corn, wheat, and soybeans on the major U.S. exchanges have recently exhibited patterns that provide a unique setting in which to assess the rational expectations present value model. Specifically, the price for these futures contracts has repeatedly failed to converge to the price of the underlying commodity on the expiration date. Between 2005 and 2010 many Chicago Board of Trade (CBOT) corn, wheat, and soybean contract expirations exhibited convergence failure with futures contracts expiring at prices up to 35 percent greater than the prevailing cash grain price (see Figure 1). In spite of the convergence failure, these contracts continued to be actively traded. Average daily trading volume in the CBOT corn, soybean, and wheat contracts doubled between September 2005, when non-convergence first appeared, and September 2008 when non-convergence was at its worst.[1]

This wave of convergence failures in the face of increasing trading volume generated a heated public debate about possible causes. Many blamed irrational traders. For example, a report by the United States Senate Permanent Subcommittee on Investigations (USS/PSI 2009) claims commodity index trading caused the non-convergence in wheat markets. The USS/PSI report maintains that index fund capital overpowered arbitrageurs, who may have been limited by credit constraints and uncertainty over

---

[1] For comparison, average daily trading volume doubled in NYMEX crude oil, tripled in COMEX gold and remained the same in COMEX copper during this same period. None of these contracts had convergence failures.

the time it would take to realize arbitrage profits. The magnitude and persistence of this non-convergence distinguishes it from idiosyncratic pricing anomalies that have arisen in the past due to market manipulation in the form of "corners" and "squeezes" (Kyle 1984, Peck and Williams 1991, Pirrong 1993, 2004, Allen, Litov, and Mei 2006)

We show in this article that persistent non-convergence can be explained using a rational expectations commodity storage model, without the need to appeal to bubbles or irrational traders. The root of the problem is that, unlike many commodities, grain futures contracts are not settled by physical delivery of the commodity.[2] With physical delivery, arbitrageurs would be able to force convergence by acquiring inexpensive grain in the cash market and delivering it at futures contract prices until the cash and futures prices equalized. When delivery occurs on a grain futures contract the firm on the short side of the market provides a delivery instrument (a warehouse receipt or shipping certificate) to the firm on the long side of the market. The delivery instrument is a security that can be exchanged for grain at any time. The long taking delivery may hold the delivery instrument indefinitely, although the firm must pay a daily storage fee while it holds the instrument.

Our theoretical model shows that non-convergence arises in equilibrium when the market price of physical grain storage exceeds the cost of holding delivery instruments. The storage fee on delivery instruments is capped by the futures exchange and has not varied much over time. However, the price of physical grain storage has varied substantially over time with changing inventory levels. Plentiful inventories generate a high price of physical storage and small inventories cause the price of storage to become negative as firms receive a convenience yield for holding the commodity (Working 1948, 1949; Brennan 1958). We call the difference between the price of carrying physical grain and the cost of carrying delivery instruments the *wedge*. The wedge represents the difference in payoffs in our present

---

[2] For example, energy futures traded on NYMEX, such as WTI crude oil, heating oil, RBOB gasoline, and Henry Hub natural gas, are settled by F.O.B. delivery into a pipeline or storage facility. Delivery on COMEX metals futures such as gold, silver, and copper are settled by a transfer of title (or warrant) to units of the commodity held in a licensed facility. For more information on energy and metals delivery specifications, see http://www.cmegroup.com/rulebook/NYMEX/index.html.

value model. We show that the magnitude of the non-convergence equals the present discounted value of future wedges over the expected period of non-convergence.

In our model, the delivery instrument represents a claim on the commodity at some unknown future date. This claim has a stream of negative payoffs associated with it (the storage fee) that differ from those associated with the commodity itself (the price of physical storage). Thus, although cash grain and the delivery instrument represent claims on the same asset (future grain), they differ in price because of their different payoff streams. This setting provides a unique test of the present value model because we are comparing two securities at each point in time and varying only the expected payoff flows. The two securities represent claims on the same asset and the arbitrage relationship between them implies that a risk premium cannot explain their difference. Moreover, this setting provides a significant challenge for the present value model. The wedge changes frequently and market participants do not know when the delivery instrument and the commodity prices will equalize. In addition, there has been considerable popular disagreement about the underlying causes of convergence failure.[3]

There is a large literature on the law of one price. One notable example that connects well with our topic is the closed-end fund puzzle, in which a bundle of stocks (the closed-end fund) sells for a different price than the basket of underlying stocks. Current theories to explain this puzzle include agency problems associated with the fund managers (e.g., Berk and Stanton 2007), liquidity differences between the fund and the underlying stocks (Cherkes, Sagi, and Stanton 2009), or limits to arbitrage (e.g., Lee, Shleifer, and Thaler 1991). We see no reason to suspect that agency or liquidity differences are relevant to non-convergence. Thus, we can directly compare the present value model to a limited-arbitrage, or bubble, model.

Divergence of prices from the present value of future payoffs defines a bubble. Rational bubble models provide an alternative to the present value model by specifying a price process that includes a bubble component, which is expected to grow on average at the discount rate (e.g., Diba and Grossman

---

[3] See Irwin et al. (2011) and USS/PSI (2009) for thorough reviews of the recent convergence problems in grain futures markets and the debate about causes.

1988, Froot and Obstfeld 1991). In such models, there is no arbitrage opportunity for rational traders to exploit. To incorporate this notion, we include in our model the possibility of a bubble solution in which the magnitude of the non-convergence (i.e., the basis) is driven by a non-fundamental noise term. We test empirically for the presence of such a bubble and find no evidence to support it.

A separate branch of literature dispenses with the assumption of rational expectations for at least some traders and produces bubbles by introducing frictions such as noise trader risk (e.g., DeLong et al. 1990, Shleifer and Vishny 1997), synchronization risk (Abreu and Brunnermeier 2003), or short sales constraints (Scheinkman and Xiong 2003). These frictions prevent arbitrageurs from eliminating bubbles. Market commentators have suggested that frictions may underlie convergence failure. For example, on July 21, 2009, Thomas Coyle, the Chairman of the National Grain and Feed Association, testified to the USS/PSI that "…disproportionate participation of investment capital has been the significant contributing factor to a disconnect between cash wheat values and wheat futures prices." Coyle's comments were echoed by the aforementioned USS/PSI report. By fitting our rational expectations model to the data, we rule out frictions-based models as explanations of convergence failure.

We estimate an econometric model to test the predictions of our theory for CBOT corn, soybeans, and wheat, and the Kansas City Board of Trade (KCBOT) wheat markets. The dependent variable in the econometric model is a measure of the wedge between the market price of physical grain storage and the cost of holding delivery instruments. The explanatory variables include: storage fee on delivery instruments, grain inventory at deliverable locations and inventories of materials and supplies divided by total sales for food products manufacturing firms (to capture convenience yield), a credit spread measure (e.g., 3-month commercial paper minus 3-month treasury bills), and the market position of commodity index traders (to represent a bubble component). The empirical evidence strongly supports our rational expectations model. Specifically, we find that high inventory in deliverable locations raises the wedge and the wedge is greatest early in the crop year when inventory is at its largest. We find no evidence of a futures bubble caused by commodity index traders. Graphical analysis shows how the wedge, driven by inventory levels, explains the occurrence and magnitude non-convergence.

Three recent studies examine the causes of the convergence failures. Heath's (2009) work is closest to ours. He develops a two-period theoretical model and shows how a cap on futures contract storage rates can induce non-convergence when the "capped" rate is below the market price of physical storage. However, his theoretical model does not show how relatively small differences in storage costs could lead to the large magnitudes of non-convergence that occurred recently and he does not conduct formal tests of model predictions, and it does not explain why the market price of storage may vary. Irwin, Garcia, Good, and Kunda (2011) identify a correlation between non-convergence and the occurrence of large carrying charges in corn, soybeans, and wheat. When carrying charges (spreads between futures prices with different delivery dates) are high enough, Irwin et al. argue that an incentive is created for takers of delivery (longs) to hold the delivery instruments and "earn the carry" rather than cancelling the instrument by converting it into grain (loading out), which would bring about convergence. The authors did not develop a formal theoretical model to explain how non-convergence could arise in equilibrium. Aulerich, Fishe, and Harris (2011) argue that the long's carry-induced incentive to hold delivery instruments can be modeled as an embedded real option to exchange the delivery instrument for another contract. The option becomes more valuable as the volatility of cash and futures prices increases; if the option value becomes large enough the cash and futures market can disconnect. A limitation of this framework is that it does not explain what drives the initial non-convergence.

While each of the previous studies is informative, they do not provide a complete picture of the underlying economic forces that generate such large episodes of convergence failure. We expand their contributions by showing how the wedge drives the occurrence and magnitude of non-convergence. Our model also reveals the important role of heterogeneous capital costs in the function of futures markets. In general, a necessary condition for the delivery instrument market to clear is that the takers of delivery (longs) have lower capital costs than the makers of delivery (shorts). We show in addition that lowering the cost of delivery does not provide a solution to non-convergence. The CME tried expanding delivery points in the wheat market in 2009 to no avail, which is consistent with our theory.[4] Moreover, we show that market power in the delivery instrument market is irrelevant for non-convergence.

Our results imply that convergence can be restored by raising the storage fee on delivery instruments. The CME recognized this point (Seamon 2009) and introduced the Variable Storage Rate (VSR) system in 2010, which raises and lowers the storage fee for CBOT wheat depending on market conditions. The VSR system has produced a larger storage fee, which has reduced the wedge and restored convergence in CBOT wheat. However, the VSR has been controversial. Market commentators have blamed the VSR for causing the price of grain storage to increase. Our theoretical and empirical results show that the causation operates in the opposite direction.[5]

The paper proceeds as follows. In Section II, we develop a theoretical model of the delivery market to explain non-convergence. We test the model empirically in Section III, and Section IV concludes the paper.

II. THEORETICAL MODEL OF THE DELIVERY MARKET

A. Institutional Background

As noted by Pirrong, Kormendi, and Haddock (1993, p. 9), "The delivery terms of futures contracts specify the types and grades of deliverable goods, and denote the places and times of delivery that must be met to avoid default on an outstanding contract." These terms evolve over time to reflect changes in the commercial standards for market transactions. If futures contract terms become misaligned with prevailing standards then the contract may no longer serve as a useful hedging instrument and its continued existence is threatened. Grain futures contracts traded at the CBOT and KCBOT specify a par

---

[4] The CBOT merged with the Chicago Mercantile Exchange (CME) in 2007 and became known as the CME Group. We refer to the corn, soybean and wheat contracts as CBOT contracts because they originated on the Chicago Board of Trade, but we use the term CME when describing actions taken by the CME group.

[5] Of course, convergence could be forced by changing the delivery mechanism to one that forced load out (like NYMEX energy futures contracts, see footnote 2) or cash-settled to a spot price index. These mechanisms would have their own challenges such as managing potential bottlenecks in the flow of grain and finding a spot price that is representative and free from manipulation. We do not consider these mechanisms further in this paper.

delivery location and grade for each contract. Delivery at non-par locations and grades is permissible at fixed premiums or discounts.[6]

In a competitive market with costless physical delivery at one particular location and date, arbitrage will force the futures price at expiration to equal the cash price. If the futures price exceeds the cash price, the cash commodity would be bought, futures sold, and delivery made. If the cash price exceeds futures, then futures would be bought, the buyer would stand for delivery, and then would sell the commodity in the cash market. This type of riskless arbitrage prevents the law of one price from being violated. In such a well-functioning delivery system, only a minimal number of futures deliveries are needed because long and short futures position holders are indifferent to offsetting their positions rather than making and taking delivery. As Hieronymus (1977, p. 340) notes, "A futures contract is a temporary substitute for an eventual cash transaction. In markets that work, delivery is rarely made and taken; futures contracts are entered into for reasons other than exchange of title."

In reality, delivery arbitrage is more complex than this simple description. When a futures contract allows multiple delivery days, locations and grades, as is the case for CBOT and KCBOT grain contracts, delivery will occur at the "cheapest-to-deliver" date, location, and grade because the shorts, or makers of delivery, seek the lowest cost alternative for sourcing the commodity to satisfy delivery obligations (Stulz 1982, Johnson 1987). The value of these delivery options to the short (timing, location, and grade) in grain markets may vary over time (Hranaiova, and Tomek 2002, Hranaiova, Jarrow, and Tomek 2005). Furthermore, both longs and shorts involved in the delivery process incur costs, which in turn determine arbitrage bounds for the convergence of cash and futures prices at delivery locations.

Delivery on CBOT and KCBOT grain futures contracts is not satisfied directly by physical grain, but instead by issuing a warehouse receipt in the case of KCBOT wheat or a shipping certificate in the

---

[6] Additional details on the CBOT and KCBOT delivery systems can be found in Irwin et al. (2011), Aulerich, Fishe, and Harris (2011), and the Exchange Rule Books located at http://www.cmegroup.com/rulebook/CBOT/I/7/7.pdf and http://www.kcbt.com/rule_book_kcbt.html.

case of CBOT corn, soybeans, and wheat.[7] A warehouse receipt is a legal document that provides proof of ownership (title) of a certain grade and quantity of a commodity at a given storage facility; e.g., 5,000 bushels of number one hard red winter wheat in firm *x*'s warehouse in Kansas City, Kansas. Crucially, warehouse receipts used in the futures delivery process are negotiable, and thus transferable between parties. A shipping certificate is also a legal document, but rather than representing actual grain in store, it gives the holder the right but not the obligation to demand load-out of the designated commodity from a particular shipping station; e.g., 5,000 bushels of number two yellow corn loaded on a barge at firm *y*'s shipping station on the Illinois River at LaSalle, Illinois.[8] The advantage of a shipping certificate is the flexibility it offers to makers of delivery since the grain can be sourced over time and space. Like warehouse receipts, shipping certificates are transferable. Neither warehouse receipts nor shipping certificates have expiration dates, and hence, are in theory infinitely lived instruments.

"Regular firms" play a key role in the CBOT and KCBOT delivery systems. Only firms approved by the exchange as regular for delivery are allowed to issue warehouse receipts or shipping certificates. Firms must meet certain exchange requirements to be eligible for regularity, such as a minimum net worth of $5 million, and have storage warehouses or shipping stations within the delivery territory of the futures contract. Regular firms are the source of all delivery instruments for their designated warehouses or shipping stations. If a maker of delivery is not a regular firm, he/she must buy a receipt or certificate from a regular firm, another holder of a receipt or certificate, or have taken delivery on a previous long futures position. A regular firm that is short is the only party that has the ability to make an "original" delivery with a newly issued delivery instrument. Regular firms are typically large commercial grain firms, such as Cargill, Bunge, and Archer Daniels Midland.

The delivery process for CBOT and KCBOT grain futures contracts consists of a three-day sequence: 1) an intention day where the short declares their intention for delivery to the exchange

---

[7] CBOT corn and soybean delivery was based on warehouse receipts prior to the March 2000 contract. CBOT wheat delivery was via a warehouse receipt prior to the July 2008 contract.

[8] In the case of a shipping certificate, title to the grain does not change hands until load out of grain occurs at the shipping station.

clearinghouse, 2) a notice day where the clearinghouse notifies the oldest outstanding long position holder with an invoice for delivery, and 3) a delivery day where the seller and the buyer exchange delivery instruments and payment. The first three-day sequence can be initiated two business days before the first business day of the expiration month and the last three-day sequence can be initiated on the business day prior to the 15th calendar day of the expiration month. This results in a total delivery period of about 10 business days for each contract.

Firms issuing delivery instruments must either have an equal quantity of grain in store or in a nearby location to back those instruments. This grain is costly to store, so a long taking delivery on a CBOT or KCBOT grain futures contract incurs a storage fee for as long as it holds the delivery instrument. This fee is assessed daily and is set by the futures exchange rather than the market. This daily storage fee is actually the maximum that can be charged by regular firms, but there is little evidence that regular firms have charged less than the maximum fee to takers of delivery. Historically, the storage fee has been fixed for long periods of time. As an example, the storage rate for CBOT wheat futures contracts was fixed at 15/100 of a cent per bushel per day (about 4.5 cents per month) from the early 1980s until June 2008.[9] In an attempt to address recent convergence problems, the CME broke with its long practice of fixed storage fees and implemented a variable storage rate (VSR) rule for wheat beginning with the September 2010 contract (Seamon 2009). The VSR is keyed to the level of term spreads in the period immediately preceding the expiration of the nearby contract. The impact of the new rule on market performance is a source of considerable controversy, particularly among grain traders.

Regular firms often argue that giving up warehouse space or tying up shipping facilities in the delivery process is costly. In particular, by issuing delivery instruments the firm gives up title to, and therefore control of, the grain that backs them. As pointed out by Hieronymous (1977, p. 159), grain in store that is under control of the firm potentially has a high option value due to the possibility of blending with grain of different quality to increase total revenue and other merchandizing opportunities that may

---

[9] We were unable to determine the exact date when this rate was first implemented for the CBOT wheat futures contracts as the records are currently sealed in the exchange archive. Figure 3.4 in Peck and Williams (1991) suggests the rate was implemented in the early 1980s.

arise (also see Paul 1970). In addition, when load-out of grain occurs, the long also incurs a load-out fee, as well as costs of weighing, grading, elevation, trimming, and blending the delivered grain (Pirrong, Haddock, and Kormendi 1993).[10] Thus, the regular firm will only deliver on futures if it receives a premium to compensate for the "inconvenience costs" associated with the delivery process.

B. Two-Period Model

We model three markets connected to a storable grain commodity: the cash market, the futures market, and the delivery instrument market. These markets are populated by two representative traders: the regular firm and the financial firm. The regular firm may issue delivery instruments (warehouse receipts or shipping certificates depending on the particular market) and has the capacity to store grain, whereas the financial firm may not issue delivery instruments and cannot store grain. The financial firm has capital cost r^f, which may be less than the capital cost r faced by the regular firm, reflecting possible advantages in capital markets.[11] The two firms operate with identical information sets, have homogeneous rational expectations, and behave competitively in all markets. In this section, we develop a two-period version of the model to elucidate the main results before moving to a dynamic infinite horizon model in Section II.C.

The regular firm enters period 1 with an endowment of I_0 units of the commodity. In period 1, it chooses how much of the commodity to store for sale in the second period (I_1) and how much to sell in the current period (I_0–I_1). It faces market net inverse demand curves $P_1 = f(I_0 - I_1)$ and $P_2 = f(I_1, \varepsilon_2)$ in periods 1 and 2, respectively. The net demand shock ε_2 is the only source of uncertainty in the model. We allow the regular firm to be risk-averse with a constant risk premium π, and we also impose a

---

[10] The non-convergence of futures prices since 2005 shown in Figure 1 far exceeds any reasonable estimates of the cost of delivery arbitrage. See Irwin et al. (2011) for further discussion on this point.

[11] In reality, the financial firm represents the firm with access to the cheapest capital, which creates the incentive for it to hold delivery instruments. The financial firm does not have to be a bank or a hedge fund, although some banks and hedge funds hold delivery instruments in these markets.

constant cost of physical storage, δ. In our formulation, the cost of physical storage includes the rental fee for warehouse space, handling and in- and out-charges, and insurance;.

The market price of physical storage may differ substantially from δ due to convenience yield, a concept introduced by Kaldor (1939) and developed by Brennan (1958). Convenience yield is typically motivated as an option value generated by transactions costs associated with sourcing the commodity (Telser 1958) or by the possibility that inventories could be driven to their lower bound (Routledge, Seppi and Spatt 2000). Two options are particularly relevant in grain markets. First, having grain in store allows the regular firm to take advantage of merchandizing opportunities that require immediate access to grain. Second, filling binspace with grain imposes an opportunity cost because that space cannot be used for anything else (Paul 1970). For example, holding wheat in store reduces opportunities to take advantage of merchandizing opportunities in corn. The first option increases willingness to hold inventory and is likely to have high value when inventory levels are low, whereas the second reduces willingness to hold inventory and is likely to have high value when inventory levels are high.[12]

We abstract from the underlying forces that generate convenience yield and specify it as a monotonically non-increasing function of inventory, denoted by $y(I_1)$. This formulation is consistent with the empirical regularity first established by Working (1948, 1949), who used Chicago wheat storage data to show that the price of storage increases with the level of inventories. Combining the cost of storage and convenience yield, the market price of storage in our model is $\delta - y(I_1)$. Under rational expectations, the equilibrium period 1 cash commodity price is[13]

$$P_1 = \frac{E[P_2] - \pi}{1 + r} - \delta + y(I_1) \,. \tag{1}$$

---

[12] The second option implies a negative convenience yield, which is not common in the literature but is implied by a model of the opportunity cost of binspace (Paul 1970). Botterud, Kristiansen, and Ilic (2009) find evidence of a negative convenience yield in the hyrdo-dominated Nord Pool electricity market. When hydro reservoirs are close to full, the possibility of overflow implies an opportunity cost to holding water in the reservoir. This is the mirror image of the option value generated in the rational storage model of Routledge, Seppi, and Spatt (2000) by the possibility that inventories may hit their lower bound.

[13] For simplicity, we specify the net demand and convenience yield functions to be such that inventories are never driven to zero, i.e., $f(I_0) < \left(E[f(0,\varepsilon_2)] - \pi\right)/(1+r) - \delta + y(0)$.

This representation is essentially the same as in the seminal paper by Brennan (1958), except that we specify δ and π as constants. In the next section, we allow δ, π, and r to vary over time.

We denote the futures price in period i for delivery in period j as $F_{i,j}$. Futures contracts are settled via delivery instruments, which can only be issued by the regular firm. Delivery instruments issued in period 1 may be held by the financial firm until period 2 through payment of a storage fee γ to the regular firm, or the financial firm may convert the instruments immediately into grain at zero cost. Converting the delivery instrument into grain (i.e., loading out) would mean that the regular firm had effectively sold grain on the cash market at the expiring futures price, $F_{1,1}$, which it would never do if $F_{1,1}$ were less than the cash price. Thus, the absence of arbitrage requires $F_{1,1} \geq P_1$.[14] The point of our model is to identify the conditions under which the equilibrium outcome is to hold delivery instruments also imply non-convergence in period 1, i.e., $F_{1,1} > P_1$. In period 2, outstanding delivery instruments are automatically converted into grain after markets clear. Thus, the absence of arbitrage requires $F_{2,2} = P_2$.

There are three potential ways for the regular firm to realize profits from storing the commodity: (i) wait and sell at the prevailing cash price in period 2 (store in the physical market), (ii) take a short futures position in period 1 for delivery in period 2 (store in the futures market) or (iii) issue a delivery instrument in period 1 that will be converted into grain in period 2 (store in the delivery instrument market). Combining (i) and (ii), the absence of arbitrage between the physical cash and futures markets ensures that the period 1 futures price for delivery in period 2 is

$$F_{1,2} = (1+r)\left(P_1 + \delta - y(I_1)\right). \tag{2}$$

Below, we combine (ii) and (iii) to explain how non-convergence can occur in period 1.

The delivery instrument market is the market for an expiring futures contract. The financial firm can take delivery of a delivery instrument in period 1 at price $F_{1,1}$. After taking delivery, the firm can enter a futures contract to deliver the certificate back to the regular firm in period 2 at price $F_{1,2}$. Thus,

---

[14] We specify only one location, grade and time period for delivery. As such, we abstract away from delivery options, which have been shown to be small in magnitude. For more on delivery options, see Hranaiova, and Tomek (2002) and Hranaiova, Jarrow, and Tomek (2005).

the discounted payoff to taking delivery equals $F_{1,2}/(1+r^f)-F_{1,1}$. The firm must pay the storage rate γ to hold the delivery instrument, so it will be willing to engage in this transaction if $F_{1,2}/(1+r^f)-F_{1,1} \geq \gamma$. Because the financial firm behaves competitively, the demand curve for delivery instruments is

$$F_{1,1} = \frac{F_{1,2}}{1+r^f} - \gamma \,. \tag{3}$$

Thus, the demand for delivery instruments is perfectly elastic; the financial firm will absorb as many instruments as the regular firm is willing to issue at this price.[15]

The regular firm incurs a cost from issuing delivery instruments in period 1 that are held until period 2, which we term the inconvenience cost. This cost is incremental to the price of storing grain in the physical market, $\delta - y(I_1)$,and may be related to several aspects of the firm's operations. The inconvenience cost may reflect the loss of flexibility from, in essence, giving up control of the grain in storage that backs outstanding certificates (Hieronymous 1977, p. 159). Alternatively, it could reflect futures exchange rules that require the firm to store grain that backs delivery instruments in a more costly location than it would use for unconstrained storage or because the possibility of load out reduces the firm's available throughput capacity and thereby reduces its ability to handle grain for customers (Paul 1970). Finally, the inconvenience cost can also represent transactions costs associated with issuing delivery instruments. We specify the inconvenience cost as a monotonically non-decreasing function of the number of certificates issued and held, C_1. We denote this function by $x(C_1)$.

The payoff to issuing a delivery instrument that is held equals $F_{1,1}+\gamma - x(C_1)$, and the cost is the discounted price of buying the certificate back next period which equals the expected cost of sourcing grain, i.e., $F_{1,2}/(1+r)$. The risk premium does not affect willingness to issue delivery instruments because the payoff and cost are both risk-free. Thus, the regular firm would be willing to issue C_1 delivery instruments if

---

[15] Until recently there was no limit on the number of warehouse receipts or shipping certificates that financial (non-commercial) firms could hold in grain futures markets. Under a CME rule instituted in February 2009, a financial firm is limited to holding no more than 600 receipts or certificates in corn, soybeans, and wheat. See http://www.cmegroup.com/rulebook/files/CBOT_RA0903-1.pdf.

$$\frac{F_{1,2}}{1+r} - F_{1,1} \leq \gamma - x(C_1)\,. \tag{4}$$

Given the no-arbitrage condition $F_{1,1} \geq P_1$, (4) implies that the expiring futures price must satisfy

$$F_{1,1} \geq \max\left(P_1, \frac{F_{1,2}}{1+r} - \gamma + x(C_1)\right). \tag{5}$$

Inserting (2) into (5) and using the fact that both firms act competitively, we see that the supply curve for delivery instruments is given by

$$F_{1,1} = P_1 + \max\left(0, \delta - y(I_1) - \gamma + x(C_1)\right). \tag{6}$$

This curve is weakly upward sloping because $x(C_1)$ is monotonically nondecreasing in C_1. The greater is the inconvenience cost, the greater the required price $F_{1,1}$ to issue delivery instruments.

Next, we show that a positive inconvenience cost implies that the delivery instrument market clears only if the regular firm faces a greater capital cost than the financial firm. The delivery instrument market clears if there exists a value C_1 such that supply (6) equals demand (3), i.e., if there exists C_1 such that

$$P_1 + \max\left(0, \delta - y(I_1) - \gamma + x(C_1)\right) = \frac{F_{1,2}}{1+r^f} - \gamma$$

which, using (2), we can re-write as

$$\begin{aligned} \max\left(y(I_1) - \delta + \gamma, x(C_1)\right) &= \frac{F_{1,2}}{1+r^f} - P_1 + y(I_1) - \delta \\ &= \frac{(1+r)(P_1 + y(I_1) - \delta)}{1+r^f} - P_1 + y(I_1) - \delta \\ &= (r - r^f)\frac{P_1 + \delta - y(I_1)}{1+r^f}. \end{aligned} \tag{7}$$

This market clearing condition describes a competitive equilibrium in which expected profit equals zero for both firms. It implies that delivery instruments will not be issued if the inconvenience cost to the regular firm exceeds the scaled difference in capital cost between the two firms. This result arises because the term spread in the futures market is determined by the financial firm's cost of capital, i.e., from (3), we have $F_{1,2} = (1+r^f)(F_{1,1} + \gamma)$. By issuing a delivery instrument, the regular firm incurs the inconvenience cost, but by selling today at price $F_{1,1} + \gamma$ and buying back next period at the price $F_{1,2}$, it essentially gains access to credit at rate r^f. This lower cost of capital compensates the regular firm for the inconvenience cost. Thus, assuming a nonzero inconvenience cost, the delivery instrument market clears at positive C_1 only if the financial firm has a lower cost of capital than the regular firm.

Equation (7) provides a condition for existence of a competitive equilibrium in the delivery instrument market. To obtain the price of the delivery instrument in this equilibrium, note that (6) and (7), along with the condition that $F_{1,1} \geq P_1$, imply that the basis is

$$\begin{aligned} F_{1,1} - P_1 &= \max\left(0, \delta - y(I_1) - \gamma + x(C_1)\right) \\ &= \max\left(0, \delta - y(I_1) - \gamma + (r - r^f)\frac{P_1 + \delta - y(I_1)}{1 + r^f}\right). \end{aligned} \tag{8}$$

Thus, the following two market relationships determine the extent of any convergence failure:

1. *The price of storing the physical commodity relative to the cost of holding delivery instruments.* If γ is set too low relative to the market price of physical grain storage, $\delta - y(I_1)$, then non-convergence may arise.
2. *Difference in capital costs between the holders of delivery instruments and the holders of grain inventory.* A large credit spread provides a means for the regular firm to access cheaper credit. This difference can be accentuated by low interest rates and high cash prices.

Equation (8) shows that non-convergence arises from a wedge between the price of carrying the commodity and the cost of holding delivery instruments. The wedge is

$$W_1 \equiv \delta - y(I_1) - \gamma + (r - r^f)(P_1 + \delta - y(I_1)) / (1 + r^f).$$

To explore the wedge further, we define the "excess" term spread as the extent to which futures prices are below full carry. Full carry occurs when the futures term spread is pushed to its maximum value given the fixed storage rate for delivery instruments (γ). The spread cannot exceed the maximum value; otherwise a riskless arbitrage opportunity is created. Specifically,

$$
\begin{aligned}
S_1 &\equiv \frac{F_{1,2}}{1+r^f} - F_{1,1} - \gamma \\
&= \frac{(1+r)\left(P_1+\delta-y(I_1)\right)}{1+r^f} - P_1 - \max\left(0, \delta - y(I_1) - \gamma + (r-r^f)\frac{P_1+\delta-y(I_1)}{1+r^f}\right) - \gamma \\
&= -\max\left(-\delta + y(I_1) + \gamma - (r-r^f)\frac{\left(P_1+\delta-y(I_1)\right)}{1+r^f}, 0\right) \\
&= \min\left(0, W_1\right).
\end{aligned}
$$

We also define the basis as $B_1 \equiv F_{1,1} - P_1$. With these definitions, we can write the basis and the excess spread as

$$B_1 = \max\left(W_1, 0\right) \qquad \text{and} \qquad S_1 = \min\left(W_1, 0\right) \tag{9}$$

which implies $W_1 = B_1 + S_1$. Thus, the wedge drives both the basis and the excess spread. When the wedge is positive the basis is positive and the futures market is at full carry (S_1=0); when the wedge is negative the expiring futures contract converges (B_1=0) and the deferred futures contract is priced below full carry (S_1<0).

To elucidate this result further, we place it in the classic supply of storage framework developed by Working (1948, 1949) and Brennan (1958). By rewriting (1), we obtain the supply of physical commodity storage as an increasing function of inventories because the convenience yield is decreasing in inventories

$$E[P_2] - P_1 = rP_1 + \pi + (1+r)(\delta - y(I_1)). \tag{10}$$

The demand for physical commodity storage is determined by the relative market net demand curves in the two periods because of the price tradeoff between selling the commodity in period 1 versus period 2. The demand for storage is

$$E[P_2] - P_1 = E\left[f(I_1, \varepsilon_2)\right] - f(I_0 - I_1). \tag{11}$$

From (6), we write delivery instrument supply as a function of the term spread in the futures market, i.e.,

$$F_{1,2} - F_{1,1} = F_{1,2} - P_1 - \max\left(0, \delta - y(I_1) - \gamma + x(C_1)\right). \tag{12}$$

This equation describes the maximum term spread at which the regular firm would be willing to issue C_1

delivery instruments. From (3), the demand for delivery instruments is perfectly elastic; the minimum term spread at which the financial firm would be willing to hold delivery instruments is $F_{1,2} - F_{1,1} = r^f F_{1,1} + (1 + r^f)\gamma$;

Based on (3) and (10)-(12), Figure 2 plots the supply and demand for commodity storage and delivery instruments for five different cases.[16] Panel A displays a case with high inventory (I_1) and therefore low convenience yield and a high price of storing the physical commodity. The demand for delivery instruments sets the futures term spread below the price of physical grain storage and implies a positive wedge. From (9), a positive wedge implies a futures market at full carry (S_1=0) and a basis equal to the wedge (B_1=W_1>0). To understand this result, recall that a delivery instrument issued in period 1 and held until period 2 becomes equivalent to owning the physical commodity in period 2. Thus, for the regular firm, issuing such a certificate is identical to taking a short futures position in period 1 to deliver the commodity in period 2. It follows that the returns to storage must be identical under the two approaches. The expected returns to storage for the regular firm are

$$\underbrace{E[P_2] - P_1}_{\text{expected price of physical storage}} = \underbrace{F_{1,2} - F_{1,1}}_{\text{futures term spread}} + \underbrace{F_{1,1} - P_1}_{\text{basis}} + \underbrace{\pi}_{\text{risk premium}} \qquad (13)$$

where we use the equilibrium condition $F_{1,2} = E[P_2] - \pi$. The regular firm is compensated for a low futures term spread by a positive basis.

Equation (13) reveals a breakdown in Working's famous argument about the role of a futures market in the temporal allocation of stocks

> *Thus existence of a futures market, coupled with the practice of hedging, gives potential holders of wheat a precise or at least a good approximate index of the return to be expected from storing wheat. This is an important fact which has been too much neglected in discussion of the economics of futures trading. It is through supplying a direct measure of the return to be expected from storage, and a means, through hedging, of assuring receipt of that return, or of approximately that return, that a futures market makes its most direct and powerful contribution to the economical distribution of supplies of a commodity over time.* (Working 1949, pp. 1257-58)

---

[16] This supply curve is downward sloping because we write the supply of delivery instruments in (12) in terms of the spread $F_{1,2} - F_{1,1}$; a larger spread implies a smaller price of the delivery instrument $F_{1,1}$.

Equation (13) shows that the term spread in a futures market provides 'a direct measure of the return to be expected from storage' only if the price of storing the physical commodity is less than the cost of holding the delivery instrument. Otherwise, the term spread provides a downward-biased estimate of the expected returns from storage. If market participants are less well informed than the two firms in our model and use the term spread to guide physical storage decisions, then they will underestimate the returns to storage and therefore store less than is optimal. This result shows how the futures term spread fails to reveal the price of storage in the presence of non-convergence.

Returning to Figure 2, we change a single parameter in each of Panels, B, C, D and E relative to Panel A. We reduce inventories (I_1) in Panel B, set the risk premium (π) to zero in Panel C, raise the cost of capital faced by the financial firm (r^f) in Panel D, and set the inconvenience cost ($x(C_1)$) to zero in Panel E. Panel B shows a case with low inventory carryover and therefore a large convenience yield that drives the price of physical storage below the level required for the financial firm to hold delivery instruments. There is no positive wedge between the futures term spread and the price of physical storage, and the futures market converges. Another way to see that no delivery instruments are issued is to note that the high convenience yield drives down the supply curve for delivery instruments so that it does not cross the demand curve for positive C_1. As drawn in Panel B, the supply of delivery instruments in (12) becomes flat because the convenience yield is too large at all values of C_1 to produce $\max\left(0, \delta - y(I_1) - \gamma + x(C_1)\right) > 0$.

The demand for delivery instrument curve defines "full carry" in the futures market. As noted earlier, full carry occurs when the futures term spread is pushed to its maximum value given the fixed storage rate for delivery instruments (γ), which corresponds to $S_t = 0$ in (9). We thus generate the regularity Irwin et al. (2011) identified; namely that the futures market is at full carry when the basis is inflated. Our model shows that full carry occurs when plentiful inventories drive the price of physical storage above the cost of holding delivery instruments.

The contrast between Panels A and B in Figure 2 illustrates the important role of inventory levels in producing non-convergence. Other components of the model are important as well. An increase in the storage rate for delivery instruments (γ) or the cost of capital faced by the financial firm (r^f) shifts the demand for certificates up and therefore reduces the basis. Thus, convergence could be achieved by raising the storage rate high enough that holding delivery instruments is unattractive. This observation underlies the VSR system of dynamic storage rates, which the CME implemented in CBOT wheat in 2010. Similarly, a smaller value for the fixed warehouse cost (δ) or the cost of capital faced by the regular firm (r) moves the supply of storage curve down and reduces the basis.

Panel C of Figure 2 shows that eliminating the risk premium (π) causes a parallel shift down in the supply of storage curve because a risk-neutral regular firm is more willing to hold inventory than a risk-averse firm. This shift increases the equilibrium amount of inventory held as the market slides down the demand for storage curve. To the extent that risk aversion reduces inventory holdings, eliminating risk aversion increases inventory, reduces convenience yield and expands the basis. This result is akin to the standard result that the producer bears some of the incidence of a sales tax if demand is less than perfectly inelastic; in our case, the basis bears some of the incidence of the risk premium. Thus, risk aversion affects the basis only through its effect on inventory from shifting the supply of storage curve; it does not affect the basis directly because the basis is determined in (8) by a no-arbitrage condition.

Panels D and E of Figure 2 reveal the role of heterogeneous capital costs and the inconvenience cost in clearing the delivery instrument market. In Panel D, we raise the financial firm's capital cost to equal that of the regular firm, i.e., $r^f = r$. This shift increases the cost to the financial firm of holding certificates, but high inventory means that the physical price of storage is high enough that the financial firm still demands delivery instruments (i.e., $E[P_2] - P_1 > r^f F_{1,1} + (1 + r^f)\gamma + \pi$). However, because capital costs are equal, the regular firm receives no compensation for the inconvenience cost and so is unwilling to issue any delivery instruments. In terms of the market clearing condition (7), high inventory implies $\max\left(y(I_1) - \delta + \gamma, x(C_1)\right) = x(C_1) > 0$. Thus, the delivery market fails to clear; no delivery instruments are issued. The fact that delivery instruments have been issued throughout the periods of non-convergence suggests that this equilibrium is not important empirically (see Figure 3 in Irwin et al. 2011).

In Panel E, we set the inconvenience cost to zero. The regular firm now makes arbitrage profits from issuing delivery instruments and will issue as many as it can.[17] Because the regular firm is selling today at price $F_{1,1}$ and buying back next period at the price $F_{1,2}$, it gains access to credit at rate r^f. However, because its cost of capital is $r > r^f$, the regular firm is essentially arbitraging the difference in capital costs. Panel E also shows that making it less costly for the regular firm to issue delivery instruments may increase the number of outstanding certificates and the profits of the regular firm, but it would not affect the basis. The basis is determined by the wedge in carrying costs, which the inconvenience cost does not enter. This case is quite relevant to the recent experience in CBOT wheat. In July 2009, the CBOT more than doubled the contract's delivery capacity by adding new delivery locations in Northwest Ohio, on the Ohio River, and on the Mississippi River. This change gave greater flexibility to regular firms and thus lowered the inconvenience cost of making delivery. However, it had no effect on the basis because it was inconsequential for the price of storing wheat in the physical market and therefore had no effect on the wedge.[18]

We model both firms as behaving competitively. A relevant empirical alternative could be that the regular firm has monopoly power. In the model as presented above, the regular firm faces a perfectly elastic demand for delivery instruments, so it could not extract monopoly rents by reducing supply in this market. However, if the demand for certificates were less than perfectly elastic, which could arise if the financial firm faced an increasing marginal cost of capital, then the firm would have an incentive to issue fewer delivery instruments. The effect on the model would be the same as the inconvenience cost because issuing an additional certificate would lower the price received on all certificates. If the regular

---

[17] Panel E arbitrarily caps the maximum allowable stock of delivery instruments at C_1. A cap is not unrealistic since the total number of warehouse receipts is limited to registered warehouse capacity in a warehouse delivery system and a maximum number of outstanding shipping certificates is imposed on each regular firm in a shipping certificate system (e.g., CME rules limit the maximum outstanding shipping certificates to 20 times a firm's daily load out rate).

[18] This is consistent with Aulerich, Fishe, and Harris's (2011) finding that cash market liquidity in CBOT corn, soybean, and wheat delivery territories was similar during delivery and non-delivery periods.

firm were to have market power in the grain storage market, then the supply of storage curve would shift up, thereby raising the expected price of storage. Such market power would expand the wedge. However, just as for the risk premium, it does so indirectly through the market for physical grain storage.

C. Dynamic Infinite Horizon Model

In this section we expand the model to an infinite horizon to accommodate the empirical reality that convergence is not forced at any particular future date. Other than the provision that the two firms maximize over an infinite horizon, the structure of the model remains the same as in the previous section. The results and intuition gained there continue to apply. We add time subscripts to π, δ, γ, r, and r^f to allow the possibility that they vary over time, but for simplicity we treat these quantities as exogenous and deterministic.

Uncertainty enters the model through the cash market net demand $P_t = f(I_{t-1} - I_t, \varepsilon_t)$, where we specify the shock sequence $\{\varepsilon_t\}$ to be stationary and ergodic. This shock is the only source of uncertainty in the model. Following Williams and Wright (1991) and Routledge, Seppi, and Spatt (2000) among others, a stationary rational expectation equilibrium exists and implies[19]

$$P_t = \frac{E_t[P_{t+1}] - \pi_t}{1 + r_t} - \delta_t + y(I_t) . \qquad (14)$$

Futures market equilibrium implies $F_{t,t+1} = E_t[F_{t+1,t+1}] - \pi_t$.

The financial firm's period t demand for delivery instruments is $F_{t,t+1} = (1 + r_t^f)(F_{t,t} + \gamma_t)$, as in the previous section. Because the option always exists to immediately convert a delivery instrument into grain, the absence of arbitrage implies $F_{t,t} \geq P_t$. Thus,

---

[19] In Williams and Wright (1991) and Routledge, Seppi, and Spatt (2000), the convenience yield is implicit and is only nonzero if I_t=0, i.e., in our notation, they would write the equilibrium as $P_t = \left(E_t[P_{t+1}] - \pi_t\right)/(1 + r_t) - \delta_t$ if I_t>0 and $P_t > \left(E_t[P_{t+1}] - \pi_t\right)/(1 + r_t) - \delta_t$ if I_t=0. Equation (14) includes this setting as a special case, but also permits more general dependence of the price of storage on inventory through the convenience yield function $y(I_t)$.

$$F_{t,t} = \max\left(\frac{F_{t,t+1}}{1+r_t^f} - \gamma_t, P_t\right).$$

It follows that the basis is

$$\begin{aligned}
F_{t,t} - P_t &= \max\left(0, \frac{E_t[F_{t+1,t+1}] - \pi_t}{1+r_t^f} - P_t - \gamma_t\right) \\
&= \max\left(0, \frac{E_t[F_{t+1,t+1}] - \pi_t}{1+r_t^f} - \frac{E_t[P_{t+1}] - \pi_t}{1+r_t} + \delta_t - y(I_t) - \gamma_t\right) \\
&= \max\left(0, \frac{E_t[F_{t+1,t+1} - P_{t+1}]}{1+r_t^f} + \delta_t - y(I_t) - \gamma_t + (r_t - r_t^f)\frac{E_t[P_{t+1}] - \pi_t}{(1+r_t)(1+r_t^f)}\right) \\
&= \max\left(0, \frac{E_t[F_{t+1,t+1} - P_{t+1}]}{1+r_t^f} + \delta_t - y(I_t) - \gamma_t + (r_t - r_t^f)\frac{P_t + \delta_t - y(I_t)}{1+r_t^f}\right)
\end{aligned} \tag{15}$$

where we use $F_{t,t+1} = E_t[F_{t+1,t+1}] - \pi_t$, and $P_t = (E_t[P_{t+1}] - \pi_t)/(1+r_t) - \delta_t + y(I_t)$. Apart from the term $E_t[F_{t+1,t+1} - P_{t+1}]/(1+r_t^f)$, this is the same expression as the second line in equation (8).

Equation (15) presents the expiring futures market basis assuming that the delivery instrument market clears. As was the case in equations (4)-(7), the regular firm would be willing to exit period *t* with C_t outstanding delivery instruments if

$$\frac{F_{t,t+1}}{1+r_t} - F_{t,t} \le \gamma_t - x(C_t) \tag{16}$$

which implies a delivery instrument supply curve of

$$F_{t,t} = \max\left(P_t, \frac{F_{t,t+1}}{1+r_t} - \gamma_t + x(C_t)\right) \tag{17}$$

and the same market clearing condition as in (7).

As in the two-period model, we define the excess futures spread as the extent to which prices depart from full carry, i.e.,

$$\begin{aligned} S_t &\equiv \frac{F_{t,t+1}}{1+r_t^f} - F_{t,t} - \gamma_t \\ &= \frac{E_t[F_{t+1,t+1}] - \pi_t}{1+r_t^f} - \gamma_t - \max\left(\frac{E_t[F_{t+1,t+1}] - \pi_t}{1+r_t^f} - \gamma_t, \frac{E_t[P_{t+1}] - \pi_t}{1+r_t} - \delta_t + y(I_t) \right) \\ &= -\max\left(0, -\frac{E_t[F_{t+1,t+1}] - \pi_t}{1+r_t^f} - \delta_t + y(I_t) + \gamma_t + \frac{E_t[P_{t+1}] - \pi_t}{1+r_t} \right) \\ &= \min\left(0, \frac{E_t[F_{t+1,t+1} - P_{t+1}]}{1+r_t^f} + \delta_t - \gamma_t - y(I_t) + (r_t - r_t^f) \frac{P_t + \delta_t - y(I_t)}{1+r_t^f} \right), \end{aligned} \tag{18}$$

and the wedge between the price of carrying the commodity and the cost of holding certificates as

$$W_t \equiv \delta_t - y(I_t) - \gamma_t + (r_t - r_t^f)(P_t + \delta_t - y(I_t)) / (1 + r_t^f) . \tag{19}$$

With these definitions, (15) and (18) imply that the basis and excess spread are

$$\begin{aligned} B_t &= \max\left(\frac{E_t[B_{t+1}]}{1+r_t^f} + W_t, 0 \right) \\ S_t &= \min\left(\frac{E_t[B_{t+1}]}{1+r_t^f} + W_t, 0 \right) \end{aligned} . \tag{20}$$

which implies the linear expression

$$B_t + S_t = \frac{E_t[B_{t+1}]}{1+r_t^f} + W_t. \tag{21}$$

Equation (20) implies if $E_t[B_{t+1}] / (1 + r_t^f) + W_t > 0$, then we have a futures market at full carry (S_t=0) and non-convergence. Conversely if $E_t[B_{t+1}] / (1 + r_t^f) + W_t < 0$ then we have a futures market at less than full carry (S_t<0) and convergence.

Equation (20) also shows that the level of the basis in *t* depends on the expected basis in *t*+1, which depends on the expected basis in *t*+2, etc. Specifically,

$$\begin{aligned} B_t &= \max\left(\frac{E_t[B_{t+1}]}{1+r_t^f} + W_t, 0 \right) \\ &= \max\left(E_t\left[\max\left(\frac{E_{t+1}[B_{t+2}]}{(1+r_{t+1}^f)(1+r_t^f)} + \frac{W_{t+1}}{1+r_t^f}, 0 \right) \right] + W_t, 0 \right) \\ &= \text{ etc}\ldots . \end{aligned}$$

To obtain a more easily interpretable expression for the basis, we define $D_{t+i} \equiv 1\left(B_{t+i} > 0\right)$ as an indicator function for whether the basis is positive in period $t+i$. Then we can write the basis as

$$\begin{aligned} B_t &= \max\left(\frac{1}{1+r_t^f} E_t \left[\frac{D_{t+1}}{1+r_{t+1}^f} E_{t+1}\left[B_{t+2}\right] + D_{t+1} W_{t+1} \right] + W_t, 0 \right) \\ &= \max\left(W_t + \sum_{i=1}^{\infty} E_t \left[D_{t+i} W_{t+i} \left(\prod_{j=0}^{i-1} \frac{D_{t+j}}{1+r_{t+j}^f} \right) \right], 0 \right). \end{aligned} \tag{22}$$

Hence, the basis equals the expected present discounted value of future wedges for as long as the basis is positive. This equation is the main addition to the model from the two-period case. It implies that a relatively small wedge term in period t can have a large effect on the basis if it is expected to persist for an extended period.

Only positive expected future basis values enter equation (20), so the basis has an option like payoff structure. Aulerich, Fishe, and Harris (2011) also view the long's incentive to hold delivery instruments an embedded real option. They price the option as a function of the relative volatility of cash and futures prices, whereas in our model the option value depends on expected future wedges. Our model makes no general prediction about the relative volatility of cash and futures prices during non-convergence episodes, but to the extent that changes in expectations of future positive wedges drive futures prices but not cash prices, our model is consistent with the results of Aulerich, Fishe, and Harris (2011).

The stationary equilibrium in (14), along with a transversality condition, ensures that the sequence in (22) converges as $s \to \infty$, where s reflects future states. This result follows from the fact that with probability one the model will enter a state with a zero basis, i.e., convergence will occur someday. Also, because convergence is guaranteed on some future date, we can view the delivery instrument and grain as claims on the same future asset, namely future grain. In other words, the delivery instrument and cash purchases of grain provide two ways of obtaining future grain. The wedge term implies that the expected stream of payoffs differs across these two claims and therefore that their current prices differ.

This setting provides a unique test of the present value model because we are comparing two securities at a point in time and varying only the expected payoff flows. The two securities represent claims on the same asset and the arbitrage relationship between them implies that a risk premium cannot explain their difference (see equation (8) and Panel C in Figure 2). Moreover, there appears little reason to think that the basis would represent a liquidity premium (Cherkes, Sagi, and Stanton 2008) or reflect agency costs (e.g., Berk and Stanton 2007), both of which have been proposed to explain the closed-end fund puzzle. Thus, we have a direct comparison between the present value model and a bubble-based model.

We also consider the possibility that the transversality condition does not hold and the basis contains a bubble component. Specifically, suppose the basis is

$$F_{t,t} - P_t = R_t + N_t \tag{23}$$

where $R_t = \max\left(E_t[R_{t+1}]/(1+r_t^f) + W_t, 0\right)$ denotes the rational component of the basis as in (22) and N_t denotes a nonnegative noise component. Some algebra shows that the equilibrium condition $F_{t,t+1} = E_t[F_{t+1,t+1}] - \pi_t$ holds if $N_t = E_t[N_{t+1}]/(1+r_t^f)$.

At prices characterized by $N_t > 0$, the regular firm would be willing to issue delivery instruments, and the financial firm would be willing to hold them as long as it could hedge at the price $F_{t,t+1} = E_t[F_{t+1,t+1}] - \pi_t = E_t[P_{t+1} + R_{t+1}] - \pi_t + (1+r_t^f)N_t$. Thus, the noise term could perpetuate itself if both firms believed it would continue, and it is an example of a rational bubble (Diba and Grossman, 1988). However, because the firms in our model are infinitely lived, neither would be willing to take the other side of this hedge. Both firms know that the bubble will burst at some future date and at that time the firm on the other side of this hedge would be left holding delivery instruments or grain for which it had overpaid (Tirole 1982). To the extent that firms do not display such rationality, a bubble could arise and persist.

In summary, our model shows that convergence failure occurs because of a wedge between the price of physical commodity storage and the cost of holding delivery instruments. The observed basis equals the present value of a nonlinear function of expected future wedges. In the next section, we develop econometric methods to assess the model and understand the causes of the recent episode of non-convergence in the CBOT and KCBOT grain futures markets.

III. EMPIRICAL ANALYSIS

We begin this section by describing our data. Next, we develop and estimate regression models to quantify the driving forces behind the wedge in corn, wheat, and soybeans. Then for the CBOT wheat market, we provide graphical evidence in support of our model. Convergence failures in this market have been greater in magnitude than those in other grain markets. Finally, we provide evidence on the extent to which our model explains the magnitude of recent basis non-convergence in each commodity.

A. Data

Using data from 1986-2010 for the CBOT contracts and 1990-2010 for the KCBOT wheat contract, we estimate regression models to determine drivers of the wedge and to assess the prediction of our theory that aggregate inventory levels at deliverable locations have strong explanatory power. Variables other than inventory that may cause changes in the wedge include credit spreads between commercial paper and the T-Bill yield, the exchange established storage rate, and the inventory-sales ratio for food manufacturing firms, which could affect convenience yield on a more macro level. We also include the open interest held by commodity index traders to capture the possibility of a bubble induced by the limited ability of the market to arbitrage against the index investment.

To compute basis at delivery locations, we use the settlement price of the expiring futures contract on the first day of the delivery month. The source for the futures prices is barchart.com. Cash prices are from the Agricultural Marketing Service of the U.S. Department of Agriculture (USDA) for Chicago, Illinois River North of Peoria, Illinois River South of Peoria, Toledo, and Kansas City. The USDA reports the range of spot bids at the specified location after 1:30 pm CST (soon after the close of the futures markets.) The data is generally available by 3:00 pm CST. Basis is calculated as the settlement futures price minus the cash bid.[20] Delivery location and grade differentials from the CBOT and KCBOT Rulebooks are applied as necessary. Contract storage rates (i.e., the storage fee on delivery instruments) are also collected from the Rulebooks.

For the interest rate faced by financial firms, we use the 3-month London Interbank Offered Rate (LIBOR). It is the most widely used benchmark rate for short-term interest rates and is compiled by the British Bankers Association in conjunction with Reuters and released to the market shortly after 11am London time each day. To approximate differences in the cost of capital between regular and financial firms, we use the spread between yield on 3-month non-financial commercial paper and 3-month Treasury Bills. Prior to 1997, financial and non-financial commercial paper yields were not reported separately; for this period we use the reported overall commercial paper rate. We obtained these data from the Federal Reserve Bank of St Louis.

Inventories of grain at deliverable locations are collected from Registrar Reports available from the CBOT and KCBOT. The reported inventories include deliverable grades, non-deliverable grades/ungraded, and Commodity Credit Corporation (CCC) stocks. Deliverable grades of grain meet the exchange quality requirements for futures delivery, excluding CCC-owned grain but including all non-CCC deliverable grades regardless of whether receipted and/or registered. Non-deliverable grades/ungraded is graded grain not meeting exchange quality requirements for futures delivery and ungraded grain, excluding CCC-owned grain. CCC stocks are owned by the CCC of the USDA and not deliverable. We report only results for total stocks at deliverable locations, but obtained nearly identical results when we used deliverable stocks only.

In recent years, manufacturing firms have developed more efficient inventory management systems using information technology. This change has reduced the willingness of these firms to hold

---

[20] Since the price series collected by the USDA represent bid prices and not necessarily transaction prices the computed basis could be contaminated by measurement error (Pirrong, Haddock, and Kormendi 1993, pp. 16-17). Heath (2009) compares USDA prices for soft red winter wheat in Chicago, Toledo, and St. Louis with prices obtained from a private data vendor for individual elevators that are regular for delivery. He finds that the median difference between the USDA and private elevator price is less than two cents per bushel. This evidence suggests USDA price data are a good reflection of the market price for grain in delivery locations.

inventory. In food markets, this change may manifest in a reduced convenience yield for grains. To approximate this component of convenience yield we use the ratio of inventories of materials and supplies held by food products manufacturing firms to sales of those firms, which is collected from the Bureau of Economic Analysis of the U.S. Commerce Department.

Positions of commodity index traders are drawn from the *Supplemental Commitments of Traders* report, which is generated by the Commodity Futures Trading Commission (CFTC). The report is commonly referred to as the Commodity Index Trader (CIT) report. The CIT data are released each Friday in conjunction with the traditional *Commitment of Traders* report and show the combined futures and options positions as of the previous Tuesday's market close. Positions are also aggregated across all contract maturities for a given commodity. The publically-available CIT data starts in 2006. The CFTC collected additional data for CBOT corn, soybeans, and wheat and KCBOT wheat over 2004-2005 at the request of the U.S. Senate Permanent Subcommittee on Investigations (USS/PSI 2009) and these additional observations are used in the analysis.[21] CIT positions are measured as the net long position (long minus short contracts) on the report date closest to the first day of delivery for the relevant futures contract. We assume zero values for CIT net long positions before 2004. We do not have access to CIT positions before this date. This is unlikely to be a significant limitation because index positions before 2004 were small (Sanders, Irwin, and Merrin 2010).

B. Regressions to Explain Variation in the Wedge

Our theory predicts that the basis is driven by a wedge between the price of physical grain storage and the cost of holding delivery instruments. Most of the terms in the wedge (19) are unobservable, but (21) suggests a way to approximate it. Defining the prediction error in the period t+1 basis as $\varepsilon_{t+1} \equiv \left(E_t[B_{t+1}] - B_{t+1}\right) / (1+r_t^f)$, equation (21) implies

---

[21] The authors are indebted to the staff of the U.S. Senate Permanent Subcommittee on Investigations for providing the 2004-2005 index trader position data.

$$B_t + S_t - \frac{B_{t+1}}{1 + r_t^f} = W_t + \varepsilon_{t+1}. \tag{24}$$

We observe B_t and essentially observe S_t because we observe the storage rate (γ_t), and can approximate the capital cost of financial firms (r_t^f) with the LIBOR. Thus, we can calculate a noisy version of the wedge using (24).

We seek to explain the wedge using a set of explanatory variables Z_t, i.e., we would like to estimate the following regression equation:

$$W_t = \beta' Z_t + u_t \tag{25}$$

where $E[u_t \mid Z_t] = 0$. This equation cannot be estimated because we do not observe W_t. Combining (24) and (25), we obtain a regression equation

$$S_t + B_t - \frac{B_{t+1}}{1 + r_t^f} = \beta' Z_t + v_{t+1} \tag{26}$$

where $v_{t+1} = u_t + \varepsilon_{t+1}$ has the property that $E[v_{t+1} \mid Z_t] = 0$. The left-hand-side variable in (26) is essentially the excess spread minus the change in the basis to the next contract expiration. Due to differences across contract months in time to the next expiration, we scale the wedge measure by m_t, the number of months until the next expiration. In our data, m_t is either 1, 2, or 3. The adjusted wedge and the left-hand-side variable in our regressions is

$$\tilde{W}_t = \frac{1}{m_t}\left(S_t + B_t - \frac{B_{t+1}}{1 + r_t^f} \right). \tag{27}$$

This adjusted wedge is measured in cents per bushel per month.

In addition to the variables described in the previous section, we include fixed effects for each contract month, as well as trend variables to purge the analysis of any systematic time-variant factors. Our dependent variable exhibits occasional large negative values when the futures term structure becomes deeply inverted. The observations are large enough to dominate in a regression model, but are irrelevant to our modeling because we are interested in the determinants of positive values of the wedge. Thus, we add a dummy variable for each dependent variable observation that is more than four standard deviations below the mean.

The estimated results for the three CBOT contracts individually and together are provided along with the KCBOT wheat contract in Table 1. For corn and soybeans, we use Toledo as the basis location until the end of 1999, when Toledo ceased to be a delivery location. Beginning in 2000, we switch to the Illinois River location. We include a dummy variable in the corn and soybeans regressions to allow a level shift in the wedge between these two periods, although its coefficient is insignificant in all models. The use of Toledo and Illinois River prices for the CBOT corn and soybean contract reflects the limited cash trade that flows through Chicago and concerns about the representativeness of reported Chicago cash prices. However, our results are similar if we use Chicago cash prices to measure the basis. We use Toledo as the delivery location for CBOT wheat and Kansas City for KCBOT wheat.

Consistent with our theoretical model, the primary driver in the relationship is inventory in deliverable locations, which is strongly related to the wedge in all cases and yields similar coefficient values across commodities. The inventory variable enters in logs, so a coefficient of 3.82 (CBOT wheat) implies that a 10% increase in inventory (in log terms) leads to a 0.38 cent increase the wedge. During the 2004 to 2008 period, deliverable stocks approximately doubled, which corresponds to an increase in log inventory of 0.69. In response to such a doubling of inventories, the coefficient implies an increase in the wedge of 3.82*0.69=2.62 cents per month. For comparison, the CBOT storage fee on delivery instruments was 4.5 cents per month during most of this period, so the inventory effect is substantial.

To assess the possibility that a high wedge draws inventory into delivery locations rather than the other way around, we use instrumental variables estimation. We use national crop-year beginning stocks of the commodity as an instrument. This variable is determined by the size of the most recent harvest and prior-year aggregate storage decisions, which are unlikely to be caused by an anticipated future wedge in storage costs. In all cases, this proved to be a strong instrument (1st stage F>10). The instrumental variables estimates were close to the OLS estimated reported in table 1 and in no case did a Hausman test suggest the presence of endogeneity.

In almost all equations, the contract-month fixed effects show that the wedge is lowest late in the crop year when inventories are low and highest around the harvest when inventories are plentiful. The corn and soybean harvest occurs around October. Corn and soybeans exhibit the lowest month effects in July and August, and the highest in September and November, respectively. (November is the omitted category in the definition of the dummy variables for soybeans.) The winter wheat harvest occurs in June and July, which is consistent with the wedge being smallest in March and largest in July for KCBOT wheat and September for CBOT wheat.

The coefficients on the other variables in the model are not statistically significant although they often possess the expected signs. The storage rate for corn and soybeans is negatively correlated with the wedge, and not statistically different from −1, which is anticipated from the theoretical model. The storage rate coefficients for wheat are positive but they demonstrate large standard errors. This coefficient is not well identified because of limited changes in the storage rate during the period (e.g., only one change during the sample period for KCBOT wheat), and when it did change in CBOT wheat it occurred in response to large basis movements which leads to a positive sign. For the inventory-sales ratio, which enters the model in log first difference form, only the wheat coefficients approach any modest level of statistical significance. These estimates suggest that the convenience yield for wheat in food manufacturing may have declined over time.

The credit spread only emerges as modestly significant in the corn equation with a positive sign, indicating that a difference in interest rates would widen the wedge. However, with a coefficient of 0.03 and the small differentials that have existed between these interest rates, it is unlikely that its effect on the wedge would be anything but small. This result is consistent with the implication of our theory that, although credit spreads are an important determinant of the equilibrium number of delivery instruments issued and the possible profits earned by the participating firms, they are only a small determinant of the magnitude of the wedge, which is driven primarily by inventories.

Finally, the findings do not support the notion in the USS/PSI (2009) report that commodity index traders overpowered arbitrageurs and thereby expanded the basis. The coefficient on CIT positions has no consistent sign across models and is far from statistically significant in all cases. This result is consistent with Stoll and Whaley (2010) and Irwin et al. (2010), who found no evidence that CIT positions significantly expand term spreads in CBOT corn, soybeans, and wheat. Mou (2011) finds evidence that term spreads expand during the period when commodity index funds roll from one contract to another. However, such effects are concentrated in commodities such as energy and livestock, which exhibit much greater spread volatility than grains and oilseeds.

Tests for breaks in the estimated models provide mixed findings. The Elliott and Muller (2006) break tests consistently indicate stable structures in all cases. However, specific tests of a break in the inventory coefficient and the constant in 2006 suggest a structural change in both the CBOT and the KCBOT wheat markets at this time. We explore this result further in the following section, where we investigate the model predictions graphically for CBOT wheat, which exhibited particularly dramatic behavior in 2008-09. Overall, these results provide evidence of the importance of inventories at the deliverable locations as the key factor in explaining wedge behavior over time. As anticipated from the theory, inventories have a pervasive effect on the behavior of the wedge in all the markets.

C. Graphical Evidence in Support of the Model

The econometric results highlight the key role of inventories and the price of physical storage in explaining non-convergence. The results are consistent with our theory which predicts that, given a constant contract storage fee, the basis expands when grain inventory is plentiful and collapses when inventories are scarce. Specifically, non-convergence develops as cash prices drop after a temporary price shock. The temporary nature of the shock is important, because only temporary shocks such as an exceptionally good harvest will cause inventories to run up sharply and convenience yield to fall.

Figure 3 shows for CBOT wheat, the basis, the wedge and inventory in Toledo, which was typically the cheapest delivery location during this period. Similar to the data for the econometric model, each curve in Figure 3 is compiled from data measured on the first day of delivery for each contract expiration between 1990-2010. Panel A reveals three distinct episodes of non-convergence: July 1998-

July 2001, March 2006-July 2007, and May 2008-May 2010. The first and third of these episodes occurred after prices had descended from a peak and inventory had accumulated. The 2006-07 episode occurred during a period of rising prices, but Panel C shows that inventory was also relatively high in this period, consistent with the theory that the basis expands when inventory is high. The fact that prices were high in 2006-07 in spite of high inventory suggests that a persistent demand shock caused the price increase. This period was characterized by a strong increase in the demand for grain due to the massive expansion of corn-ethanol production (e.g., Abbot, Hurt, and Tyner 2011).

Panel B decomposes the wedge into two components. The first component is the storage rate on delivery instruments, γ_t. Aside from a small drop in 2000 and two separate increases in 2009-10, the storage rate remained constant at about 4.5 cents per bushel per month. The second line shows the remaining components of the wedge, i.e., it is

$$\begin{aligned} W_t + \gamma_t + \varepsilon_{t+1} &= B_t + S_t - \frac{B_{t+1}}{1+r_t^f} + \gamma_t + \varepsilon_{t+1} \\ &= \delta_t - y(I_t) + (r_t - r_t^f)\frac{P_t + \delta_t - y(I_t)}{1+r_t^f} + \varepsilon_{t+1}. \end{aligned} \tag{28}$$

We plot a three-period, centered moving average of this quantity to smooth out the shocks ε_{t+1}. We label this curve the price of physical grain storage, although it also includes a credit spread component. Comparing Panels B and C, we see that this curve is closely related to inventory, which suggests that the credit spread component plays a small role in determining the wedge. Rather, the dominant factor is inventories, as we found with the econometric model.

With the aid of Figure 3, the reasons for the recent and dramatic failures of convergence come into sharper focus. Demand and/or supply shocks to the underlying commodity created a surge in inventories, which in turn drove up the price of physical grain storage. This market price of storage substantially exceeded the storage rate being paid on CBOT and KCBOT grain futures contracts. In the case of CBOT wheat, Panel B shows that the gap between the two was very large for much of 2007-2010, and hence, the very large delivery location basis. The gap between the two series only began to narrow with the implementation of the VSR rule, which started with the September 2010 contract.

D. Explaining the Magnitude of the Basis

On September 2, 2008, the cash price of soft red winter wheat in Toledo was \$5.49, which was \$1.95 below the price of the expiring futures contract and by far the widest basis in our sample. Cash prices had dropped almost 50 percent in the previous six months as plentiful harvests had replenished global inventories and relieved supply pressures. In our theoretical model, the basis expands when rising inventories cause an increase in the price of carrying physical grain. Was the increase in the price of physical grain storage large enough to justify a basis of \$1.95? To answer this question, and similar questions for other expirations and commodities in our sample, we ask whether subsequent wedges were large enough to match the basis using the present value expression in (22).

Equation (22) predicts that the current basis equals the present value of expected future wedges over the period in which the basis is positive. Testing this prediction explicitly would require a model for market expectations. Instead, we assess whether the observed basis patterns are commensurate with the basis that would be implied by (22) under the assumption of perfect foresight. This approach is similar to that of Shiller (1981) and LeRoy and Porter (1981), who asked whether subsequent dividends vary enough to explain stock price variation. Here, we ask whether subsequent wedges can justify the observed difference in the prices of the delivery instrument and cash grain. Thus, we assess the present value model using two securities (delivery instrument and cash grain) that represent claims on the same asset (future grain), but which have different flows of services in the interim. Rather than comparing time series variation in asset prices as in Shiller (1981) and LeRoy and Porter (1981), we compare the difference between two prices (futures and cash) on a particular date, holding constant the future value of the asset and discount rates.

In addition to market expectations, we also do not observe the wedge. From (20), the basis implied by perfect foresight and a wedge proxy is

$$B_t^{PF} = \max\left(\frac{B_{t+1}^{PF}}{1 + r_t^f} + \hat{W}_t, 0 \right) \tag{29}$$

where $\hat{W}_t$ denotes a proxy for the wedge. To proxy for the wedge, we use two measures: the 3-period moving average of the proxy defined in (24) and the predicted values from the regression equation (26). In each case, we set the perfect-foresight basis for the last expiration in our sample equal to the actual basis. Then we solve backwards using (29) to generate a basis time series.

Figure 4 plots the observed basis for each of the four commodities along with the predicted basis from each of the two wedge proxies. The smoothed proxy tracks the CBOT wheat basis very closely. It predicts a \$1.70 basis in September 2008 and a steady decline the ensuing two years. To understand the source of this prediction, reconsider Figure 3, Panel B which shows that the price of physical storage averaged about 13 cents and the storage rate about 5 cents between September 2008 and July 2010 implying an average wedge of about 8 cents per month during this period. If traders believed in September 2008 that the high price of physical storage would persist until the July 2010 harvest (which it did), then they would expect 8 cent wedges for 22 months. Discounting at the average LIBOR rate during this period (0.2% per month), they would price the basis at \$1.72 and expect it to decline steadily at 8c per month to zero in July 2010. This predicted price is 23 cents lower than the actual basis, but is of the right order of magnitude. For example, a 9 cent monthly wedge would raise the predicted basis to \$1.94 and a 7 cent monthly wedge would lower the predicted basis to \$1.51. Between September 2008 and July 2010 the basis dropped steadily, reaching 12 cents in July 2010. This steady drop matches the predictions of our theory.

Predictions based on the regression model in (26) match the actual 2008-09 basis closely for CBOT wheat. The predictions for both CBOT and KCBOT wheat allow for the constant and the coefficient on inventory to change in 2006. This break was found to be statistically significant in Table 1, and it improves the basis predictions substantially. For instance, the model predicts a peak basis for CBOT wheat in September of 2008 of \$1.92, which is very close to the actual basis of \$1.95. Without allowing the break, the model predicts a basis of \$0.78 in September 2008. Figure 3, Panel C shows inventory levels in 2009 were similar to those in 2001, but the price of physical storage was much greater in the latter period. A break in 2006 enables the model to increase the predicted wedge in the latter period.

For the other commodities and time periods, the smoothed proxy also matches the 2008 basis spike in corn and the 2010 spike in KCBOT wheat. It matches less closely the 2008 KCBOT wheat spike and the high corn and KCBOT wheat basis in 1998-2001. Similarly, the regression predictions match the 1998-2001 and 2006 basis spikes well for CBOT wheat. For corn and KCBOT wheat, the model predicts well since 2006 but less well in 1998-2001. Neither proxy captures the high soybean basis in 2008. Overall, these results enable us to reject the bubble model because the perfect-foresight predictions generally explain the major basis spikes in our sample period.

IV. CONCLUSION

In this article, we develop a dynamic rational expectations model of commodity storage that explains recent convergence failures in grain futures markets. When delivery occurs on a grain futures contract, the firm on the short side of the market issues a delivery instrument (a warehouse receipt or shipping certificate) to the firm on the long side of the market. The delivery instrument is a security that can be exchanged for grain at any time. The long firm may hold the delivery instrument for as long as it wishes, although it must pay an exchange determined daily storage fee while it holds the instrument. We show that non-convergence arises in equilibrium when the market price of physical grain storage exceeds the cost of holding delivery instruments. We call the difference between the price of carrying physical grain and the cost of carrying delivery instruments the *wedge*. We show theoretically and empirically that the magnitude of the non-convergence equals the expected present discounted value of a function of future wedges. As such, our results provide an example of a significant and complex market in which a present value model explains prices.

Futures markets exist to facilitate effective risk-shifting and efficient price discovery (e.g., Telser and Higinbotham 1977, Pirrong, Haddock, and Kormendi 1993). Our model shows that if futures prices are driven by expectations of future wedges, then they fail to price cash grain. However, non-convergence does not induce any welfare losses in our model because market participants have full information and understand the equilibrium. Since we do not model the price discovery process, the futures markets does not perform this function in our framework. Moreover, firms that wanted to hedge cash price risk could design derivative contracts to, for example, price the option to load out from delivery instruments (Aulerich, Fishe, and Harris 2011).

The absence of welfare losses in our model is consistent with the empirical fact that trading in corn, soybean and wheat futures does not appear to have been affected by non-convergence. Figure 5 shows that daily average volume and open interest for CBOT corn and wheat increased significantly though the non-convergence period, and that these quantities increased at similar rates to NYMEX crude oil, which had no convergence problems. Moreover, volume increased at similar rates for corn and wheat, and open interest even increased slightly faster for wheat than corn, in spite of wheat exhibiting much greater non-convergence.

In contrast, numerous prominent futures contracts have undergone dramatic declines in trading when structural problems undermined their facility as hedging and price discovery tools. The most recent example is pork bellies, which ceased trading in 2011 because it was no longer useful as a hedging tool. In an older example, Working (1954) documents a dramatic decline in volume and open interest in Kansas City wheat futures in 1953 when a change in the value of delivery options caused the futures market to price soft rather than hard red winter wheat. The Kansas City contract lost about 70 percent of its trading volume in a two-month period and wheat trading in Kansas City only recovered when the exchange introduced a new futures contract that allowed only hard wheat to be delivered. A similar example arose in GNMA futures in the mid 1980s. This contract permitted flexibility of the bonds that would be delivered on the contract, but set the conversion factors in such a way that the cheapest-to-deliver bond was only weakly correlated with the instrument that market participants most wanted to hedge (Johnston and McConnell 1989). Trading volume declined rapidly between 1983 and 1985, and the contract ceased trading in 1987.

Although grain and oilseed markets have so far taken non-convergence in stride, it would be naive to dismiss convergence as irrelevant. The vigorous public debate surrounding the causes of non-convergence suggests that many market participants may not have fully understood the source of non-convergence and therefore may have misinterpreted the price signals received from futures prices. Our model shows there is a breakdown in Working's (1949) famous argument about the role of a futures market in the temporal allocation of stocks. Specifically, the term spread provides a downward-biased estimate of the expected returns from storage when the price of storing the physical commodity is more than the exchange-determined storage fee on the delivery instrument. If this bias is not recognized and understood, market participants may underestimate the returns to storage and store less than is optimal. Even fully informed traders, who could theoretically enter into derivative contracts to achieve a desired hedge, would incur additional transactions costs in the presence of non-convergence. Thus, if non-convergence were to persist, there would be potential for a competing set of futures contracts that enforces convergence to replace the current CBOT and KCBOT contracts.

CBOT wheat has been the poster child for non-convergence. The institution of a variable rate storage rule (VSR) appears to have resolved the immediate convergence problems in that market, but significant obstacles remain. The VSR is complicated and potentially prone to manipulation, and significant structural problems exist due to the fact that wheat production has moved away from Toledo, which remains the cheapest-to-deliver location (Irwin et al. 2011). More broadly, the physical price of grain storage has increased significantly in recent years, and in wheat markets it has increased by more than would be expected by the observed inventory levels. The source of this rise remains a question for future research.

References

Abreu, D. and M. K. Brunnermeier. 2003. “Bubbles and Crashes.” *Econometrica*. 71:173-204.

Abbot, P.C., C. Hurt, and W.E. Tyner. 2011. “What’s Driving Food Prices in 2011?” Issue Report, Farm Foundation. http://www.farmfoundation.org/news/articlefiles/1742-FoodPrices_web.pdf.

Allen, F., L. Litov, and J. Mei. 2006. “Large Investors, Price Manipulation, and Limits to Arbitrage: An Anatomy of Market Corners.” *Review of Finance* 10:645-693.

Aulerich, N.M., R.P.H. Fishe, and J.H. Harris. 2011. “Why Do Expiring Futures and Cash Prices Diverge for Grain Markets?” *Journal of Futures Markets* 31:503–533.

Berk, J. and R. Stanton. 2007. “Managerial Ability, Compensation, and the Closed-End Fund Discount.” *Journal of Finance* 62:529-556.

Botterud, A., T. Kristiansenb, and M D. Ilic. 2009. “The Relationship between Spot and Futures Prices in the Nord Pool Electricity Market.” *Energy Economics* 32:967-978.

Brennan, M.J. 1958. “The Supply of Storage.” *American Economic Review* 48:50-72.

Cherkes, M., J. Sagi, and R. Stanton. 2009. “A Liquidity-Based Theory of Closed-End Funds.” *Review of Financial Studies* 22:257-297.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann. 1990. “Noise Trader Risk in Financial Markets.” *Journal of Political Economy* 98:703-738.

Diba, B.T. and H.I. Grossman. 1988. “The Theory of Rational Bubbles in Stock Prices.” *The Economic Journal* 98:746-754.

Elliott, G. and U.K. Müller. 2006. “Optimally Testing General Breaking Processes in Linear Time Series Models.” *Review of Economic Studies* 73(4): 907-940.

Froot, K.A. and M. Obstfeld. 1991. “Intrinsic Bubbles: The Case of Stock Prices.” *American Economic Review* 81(5):1189-1214.

Harrison, P. 2001. “Rational Equity Valuation at the Time of the South Sea Bubble.” *History of Political Economy* 33(2): 269-281.

Heath, D. 2009. “Convergence Failure in CBOT Wheat Futures.” Working Paper. University of Chicago.

Hieronymus, T.A. 1977. *Economics of Futures Trading for Commercial and Personal Profit,* Second Edition. New York, NY: Commodity Research Bureau.

Hranaiova, J. R.A. Jarrow, and W.G. Tomek. 2005. “Estimating the Value of Delivery Options in Futures Contracts.” *Journal of Financial Research* 28:363-383.

Hranaiova, Jana, and William G. Tomek. 2002. “Role of Delivery Options in Basis Convergence.” *Journal of Futures Markets* 22:783-809.

Irwin, S.H., P. Garcia, D.L. Good, and E.L. Kunda. 2011. “Spreads and Non-Convergence in Chicago Board of Trade Corn, Soybean, and Wheat Futures: Are Index Funds to Blame? *Applied Economics Perspectives and Policy* 33:116-142.

Johnson, H. 1987. "Options on the Maximum or Minimum of Several Assets." *Journal of Financial and Quantitative Analysis* 22:277-283.

Johnston, E.T. and J.J. McConnell. 1989. "Requiem for a Market: An Analysis of the Rise and Fall of a Financial Futures Contract." *The Review of Financial Studies* 2:1-23.

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

Kyle, A. S. 1984. "A Theory of Futures Market Manipulation." *The Industrial Organization of Futures Markets*. Anderson, R.W. (ed.). Lexington, MA: Lexington Books.

Lee, C.M.C., A. Shleifer, and R.H. Thaler. 1991. "Investor Sentiment and the Closed-End Fund Puzzle." *The Journal of Finance* 46:75-109.

LeRoy, S.F.,R.D. Porter. 1981. "The Present Value Relation: Tests Based on Implied Variance Bounds." *Econometrica* 49:555-574.

Mou, Y. 2011. "Limits to Arbitrage and Commodity Index Investment: Front-Running the Goldman Roll." Working Paper, Columbia University.

Paul. A.B. 1970. "The Pricing of Binspace: A Contribution to the Theory of Storage." *American Journal of Agricultural Economics* 52:1-12.

Peck, A. E., and J.C. Williams. 1991. "An Evaluation of the Performance of the Chicago Board of Trade Wheat, Corn, and Soybean Futures Contracts during Delivery Periods from 1964-65 through 1988-89." *Food Research Institute Studies* 22:128-225.

Pirrong, S.C. 1993. "Manipulation of the Commodity Futures Market Delivery Process." *Journal of Business* 66:335-370.

Pirrong, S. 2004. "Detecting Manipulation in Futures Markets: The Ferruzzi Soybean Episode." *American Law and Economics Review* 6:28-71.

Pirrong, S.C., D. Haddock, and R. Kormendi. 1999. *Grain Futures Contracts: An Economic Appraisal*. Boston, MA: Kluwer Academic Publishers.

Routledge, B.R., D.J. Seppi, and C.S. Spatt. 2000. "Equilibrium Forward Curves for Commodities." *Journal of Finance* 55:1297-1328.

Sanders, D.R., S.H. Irwin, and R.P. Merrin. 2010. "The Adequacy of Speculation in Agricultural Futures Markets: Too Much of a Good Thing?" *Applied Economics Perspectives and Policy* 32:77-94.

Seamon, F. 2009. An Introduction to Variable Storage Rates in Wheat Futures. Research and Product Development, CME Group. http://www.cmegroup.com/trading/agricultural/files/VSR-White-Paper-FINAL.pdf.

Scheinkman J. and W. Xiong. 2003. "Overconfidence and Speculative Bubbles." *Journal of Political Economy* 111:1183-1219.

Shiller, R.J. 1981. "Do Stock Prices Move too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71:421-436.

Shleifer, A. and R.W. Vishny. 1997. "The Limits of Arbitrage." *Journal of Finance* 52: 35–55.

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

Stulz, R. 1982. "Options on the Minimum or the Maximum of Two Risky Assets: Analysis and Applications," *Journal of Financial Economics* 10:161-185.

Telser, L. 1958. "Futures Trading and the Storage of Cotton and Wheat." *Journal of Political Economy* 66: 233-255.

Telser, L., and H. Higinbotham. 1977. "Organized Futures Markets: Costs and Benefits." *Journal of Political Economy* 85: 969-1000.

Tirole, J. 1982. "On the Possibility of Speculation under Rational Expectations." *Econometrica* 50:1163-1182.

United States Senate, Permanent Subcommittee on Investigations (USS/PSI). 2009. *Excessive Speculation in the Wheat Market*. Washington, D.C.: U.S. Government Printing Office.

Williams, J.C., and B.D. Wright. 1991. *Storage and Commodity Markets*. Cambridge, UK: Cambridge University Press.

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

Working, H.1949. "The Theory of Price of Storage." *American Economic Review* 39(6): 1254-62.

Working, H.1954. "Whose Markets?: Evidence on Some Aspects of Futures Trading." *Journal of Marketing* 19(1): 1-11.

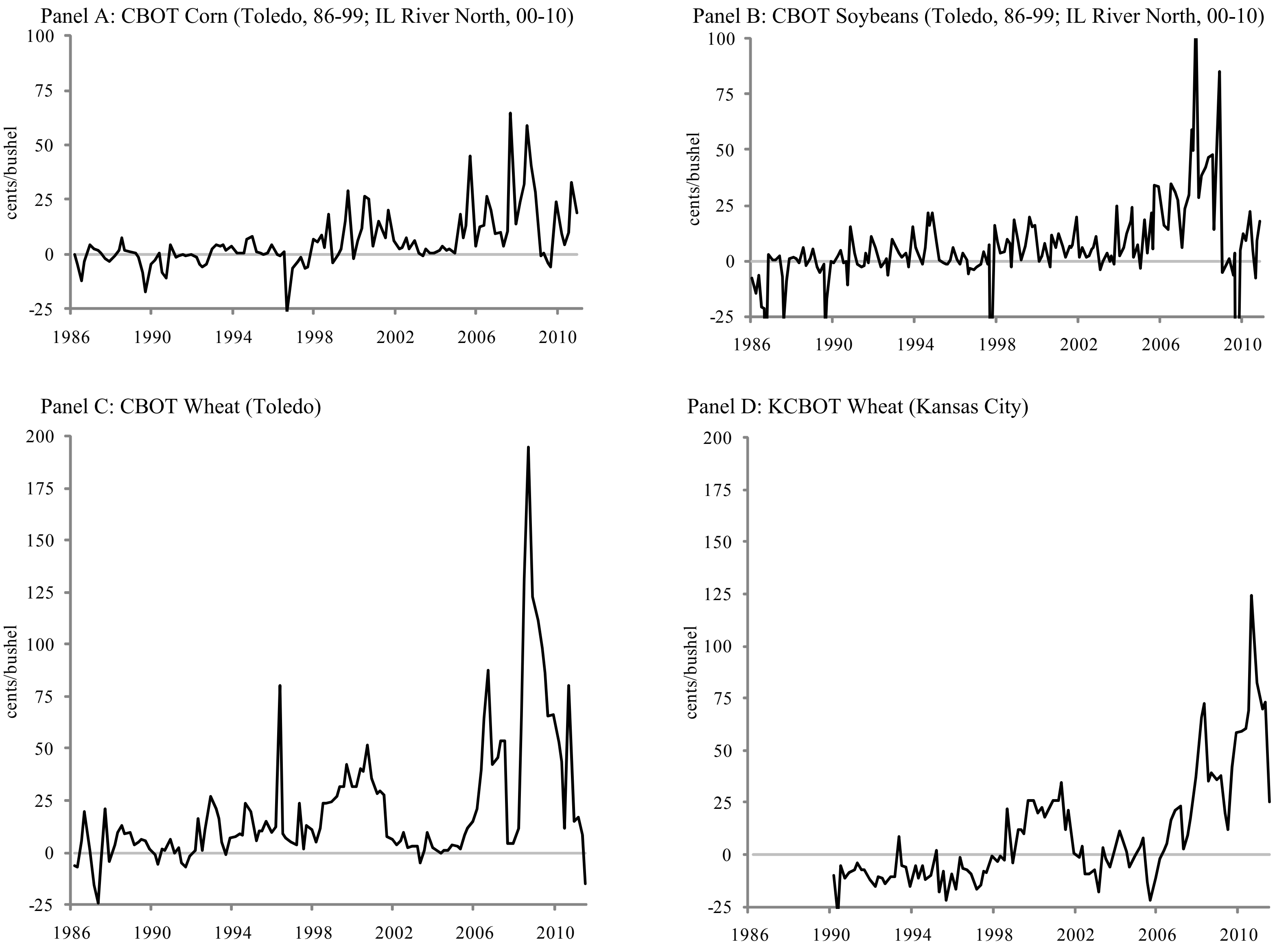
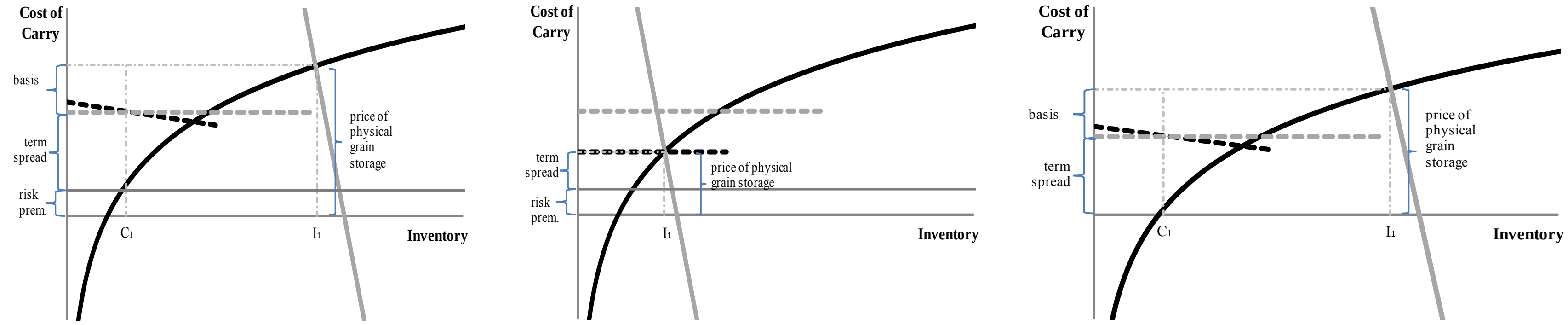


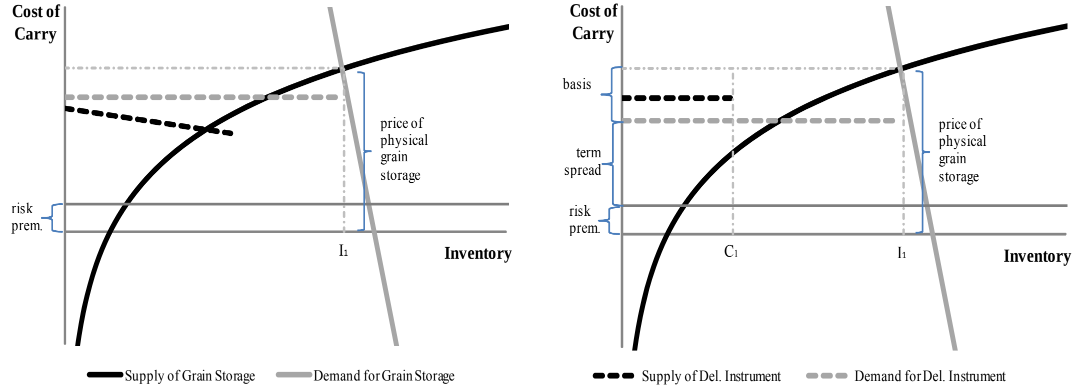
Figure 1. Futures Price at Expiration minus Cash Price, 1986-2010. City in parentheses is cash price location. Toledo ceased to be a delivery location for corn and soybeans in 1999, so we switch to the Illinois River North of Peoria at that time.



Panel A: Non-convergence (High Inventory)

Panel B: Convergence (Low Inventory)

Panel C: No Risk Aversion



Panel D: Market Doesn't Clear (Equal Capital Costs)

Panel E: Arbitrage Profits (No Inconv. Cost)

Figure 2. Equilibrium in the Delivery Instrument, Futures and Cash Markets

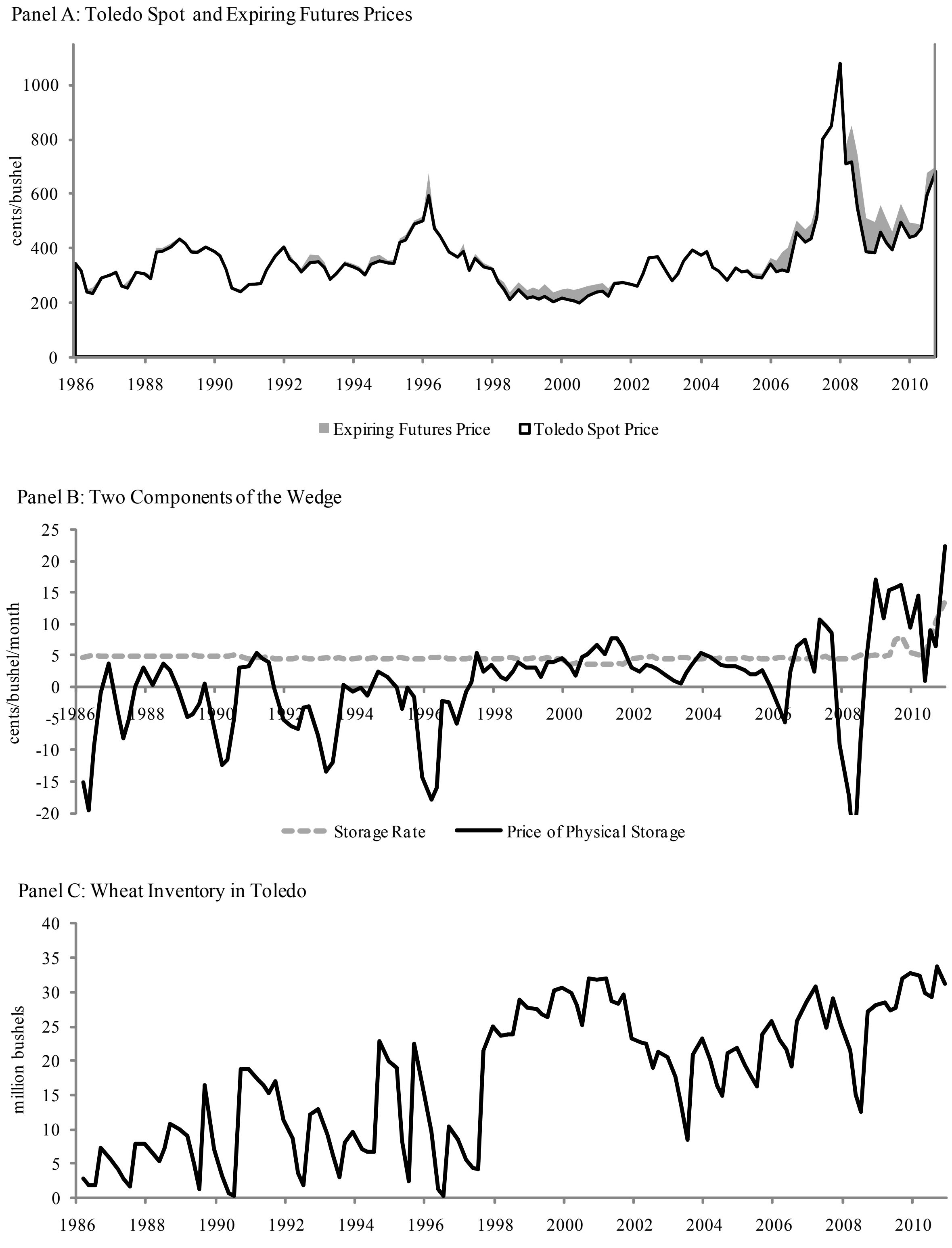
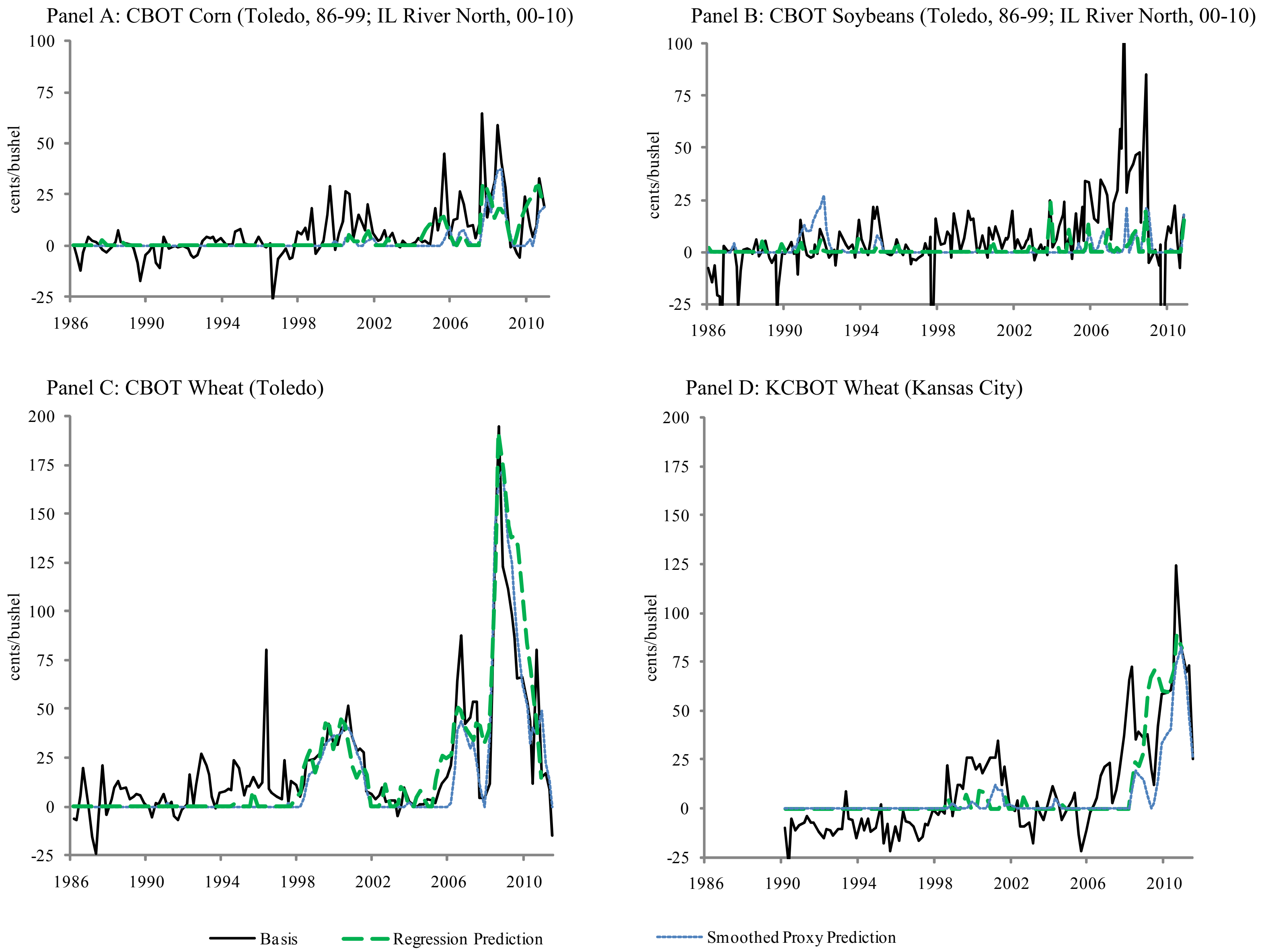


Figure 3. Elements of Non-convergence in CBOT Wheat, 1986-2010.


Panel A: CBOT Corn (Toledo, 86-99; IL River North, 00-10)

Panel B: CBOT Soybeans (Toledo, 86-99; IL River North, 00-10)

Panel C: CBOT Wheat (Toledo)

Panel D: KCBOT Wheat (Kansas City)

Figure 4. Perfect Foresight Basis: 1986-2010.

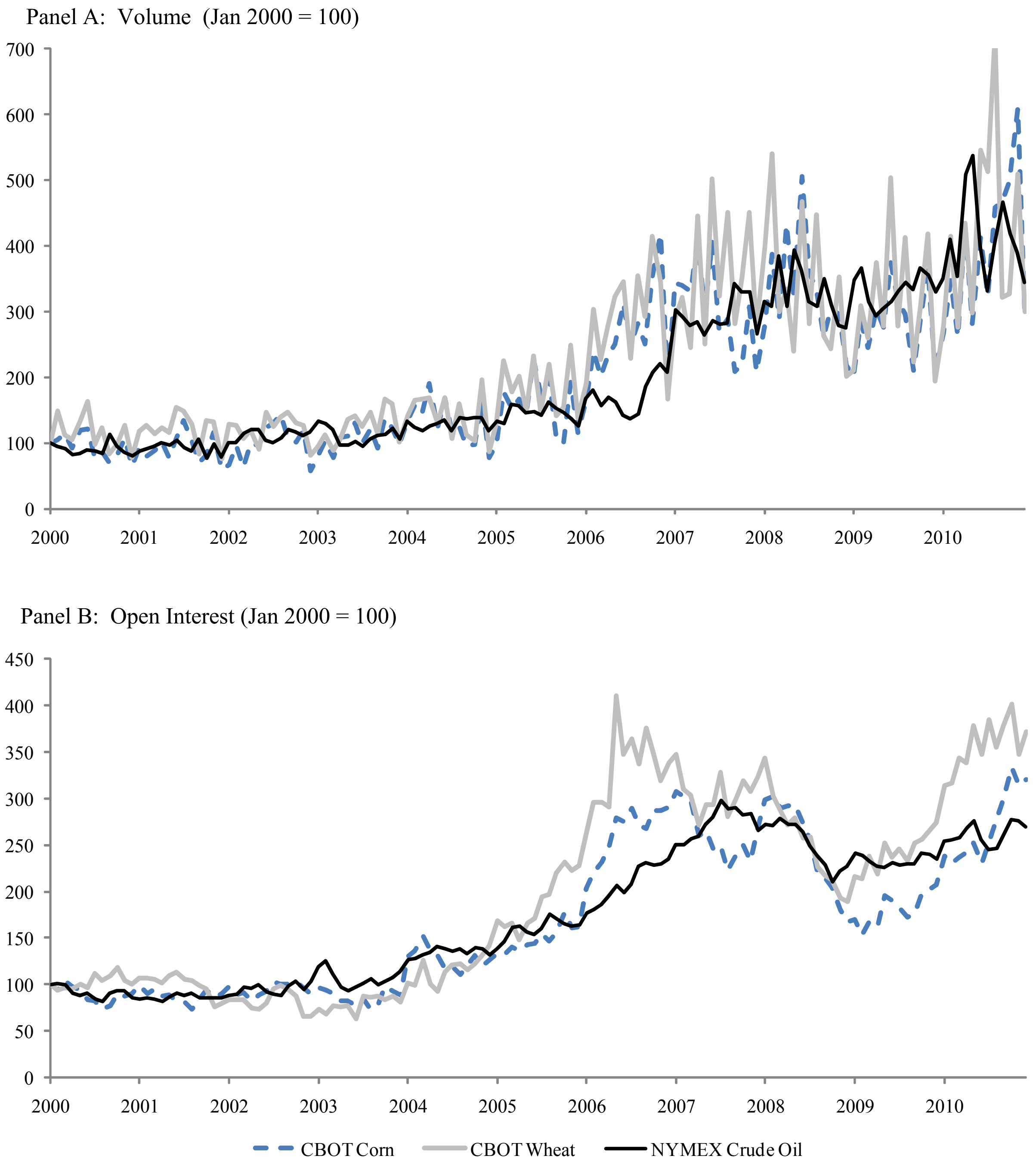


Figure 5. Average Daily Trading Volume and Open Interest, 2000-2010. Daily averages by Month. Indexed to 100 in January 2000. Aggregated across all open expirations.

Table 1. Wedge Regressions for CBOT Corn, Soybeans, and Wheat and KCBOT Wheat, 1986-2010

Commodity	**CBOT Corn**		**CBOT Soybeans**		**CBOT Wheat**		**All CBOT**		**KCBOT Wheat**	
Basis Location	**Toledo/IL River**		**Toledo/IL River**		**Toledo**		**Toledo/IL River**		**Kansas City**	
variable	**coeff**	**t-stat**	**coeff**	**t-stat**	**coeff**	**t-stat**	**coeff**	**t-stat**	**coeff**	**t-stat**
Inventory (Total)	3.17	3.25	4.56	3.13	3.82	2.75	3.13	5.55	3.03	2.85
Storage Rate	-1.91	-1.48	-1.78	-0.61	0.44	0.35	0.01	0.00	1.22	0.14
Inventory-Sales Ratio	0.29	0.45	0.87	1.08	-1.13	-1.46	-0.07	-0.15	-1.20	-2.00
Credit Spread	0.03	1.49	-0.02	-0.50	-0.03	-0.81	0.01	0.27	-0.01	-0.29
CIT	0.00	0.02	-0.46	-0.86	-0.02	-0.07	-0.06	-0.42	1.36	1.05
Constant	-858	-2.60	-475	-0.93	-158	-0.32	-514	-2.13	-667	-2.38
Trend	0.42	2.57	0.22	0.87	0.06	0.24	0.24	1.99	0.31	2.21
January			-4.66	-2.37			0.50	0.32		
March	0.37	0.40	-4.57	-2.38	-5.03	-2.67	-1.69	-1.76	-1.25	-0.72
May	-1.75	-1.42	-5.93	-2.66	-2.01	-0.82	-2.37	-1.91	2.08	1.13
July	-2.35	-1.27	-7.15	-2.02	0.68	0.25	-2.11	-1.64	3.13	1.54
August			-15.62	-3.41			-11.56	-2.45		
September	3.31	2.19	-10.38	-2.48	2.96	1.76	2.02	1.83	4.16	3.04
Corn Dummy							3.53	2.27		
Soybean Dummy							2.43	1.27		
Toledo Dummy (Corn)	-1.50	-0.91					-3.40	-2.12		
Toledo Dummy (Soy)			-3.21	-1.00			-2.41	-1.09		
Outlier Dummies										
19960301					-37.20	-22.34	-40.41	-36.12		
19960701	-44.95	-25.71					-45.91	-34.78		
20040701			-135.95	-39.17			-137.48	-77.12		
20090701			-100.21	-14.32			-106.47	-42.50		
20090901			-80.74	-20.34			-92.95	-46.93		
Diagnostics	**stat**	**crit. val.**	**stat**	**crit. val.**	**stat**	**crit. val.**	**stat**	**crit. val.**	**stat**	**crit. val.**
t_test: Storage Rate = -1	-0.70	±1.96	-0.27	±1.96	1.04	±1.96	0.82	±1.96	0.26	±1.96
1st Stage F-Stat	11.72	10.00	11.88	10.00	19.88	10.00	109.56	10.00	78.52	10.00
Hausman Test	0.37	3.84	0.17	3.84	0.66	3.84	0.04	3.84	0.51	3.84
Break in 2006	1.58	5.99	1.35	5.99	88.11	5.99	1.22	5.99	10.41	5.99
Elliott-Muller Break Test	-10.44	-14.32	-6.38	-14.32	-8.99	-14.32	-3.63	-14.32	-12.39	-14.32
R-square (all data)	0.53		0.63		0.37		0.49		0.23	
R-square (excluding outliers)	0.26		0.25		0.28		0.20		0.23	
Sample Size	124		174		124		422		104	

Notes: KCBOT sample begins in 1990. Std errors estimated using Newey-West method with one lag