WORKING PAPER SERIES

Contango in Cushing? Evidence on Financial-Physical Interactions in the U.S. Crude Oil Market

Louis H. Ederington, University of Oklahoma
Chitru S. Fernano, University of Oklahoma
Kateryna Holland, University of Oklahoma
Thomas K. Lee, U.S. Energy Information Administration

March, 2012

Contango in Cushing?
Evidence on Financial-Physical Interactions in the U.S. Crude Oil Market [*]

Louis H. Ederington, Chitru S. Fernando and Kateryna Holland
Price College of Business
University of Oklahoma
307 West Brooks Street, Norman, OK 73019

Thomas K. Lee
Energy Information Administration
U.S. Department of Energy
1000 Independence Ave., SW, Washington, DC 20585

March 2012

JEL Classifications: oil prices, spot oil markets, oil futures markets, oil storage, cash & carry arbitrage, speculation, index investors.

Keywords: G13, G18, Q41.

---

[*] Author contact information: Ederington: lederington@ou.edu, (405)325-5697; Fernando: cfernando@ou.edu, (405)325-2906; Holland: kateryna.holland@ou.edu; (405)325-5591; Lee: Thomas.Lee@eia.gov; (202)586-0829.

We thank Bruce Bawks, John Conti, Scott Linn, Anthony May, Glen Sweetnam, John Zyren, and seminar participants at the U.S. Energy Information Administration and the Second IEA IEF OPEC Workshop on the interactions between physical and financial energy markets for valuable discussions and comments. We gratefully acknowledge financial support from the U.S. Department of Energy -- Energy Information Administration and the University of Oklahoma Office of the Vice President for Research. The views expressed in this paper reflect the opinions of the authors only, and do not necessarily reflect the views of the Energy Information Administration or the U.S. Department of Energy. The authors are solely responsible for all errors and omissions.

Contango in Cushing?
Evidence on Financial-Physical Interactions in the U.S. Crude Oil Market

Abstract

We revisit an unresolved question that lies at the heart of the debate on the influence of financial speculators on spot oil prices and in which interest has intensified following the 2007-2008 oil price spike – whether there is a viable and active cash and carry market in crude oil. We study the relationship between U.S. crude oil inventories and the spread between WTI crude oil futures, while carefully accounting for the links between the U.S. and global oil markets, controlling for supply and demand shocks that affect both prices and inventories, and allowing for inventory response time lags. We find that over the 2004-2011 period, crude oil inventories at Cushing were a significant positive function of the spread between the two- and one-month NYMEX WTI crude oil futures with a lag. Over the 1992-2004 period (before Cushing inventories were reported separately), total U.S. non-SPR inventories and inventories in the PADD2 district (which includes Cushing) were positive functions of lagged spreads. However, over the 2004-2011 period, neither total U.S. non-SPR inventories nor PADD2 inventories were significant functions of the spread once Cushing inventories were excluded. These findings are consistent with Cushing being the WTI pricing and physical settlement hub. None of the other four PADD inventories are significantly related to the spread over either period, which may reflect significant limits to financial arbitrage at locations away from Cushing. Current crude oil inventories appear to be influenced by spreads over the last eight weeks, which suggests that current spreads likely lead to contracts for forward delivery that do not result in a change in actual inventory levels until delivery occurs sometime in the future. We further find evidence that total U.S and most individual PADD inventories (but not at Cushing) are a negative (positive) function of the change in current (next week) refinery inputs and a positive (negative) function of the current (next week) imports, indicating that storage operators are able to partially anticipate crude oil shortages and surpluses, and adjust their inventories accordingly. These findings establish, to our knowledge, the first tangible evidence documented in the literature of a causal link between oil futures and spot markets via inventory changes resulting from cash and carry arbitrage, and provide an important foundation for future research on the impact of financial traders on the spot markets, especially the twin questions of (a) whether financial traders exacerbate or attenuate spot price volatility, and (b) whether they systematically affect the spot oil price level.

JEL Classifications: oil prices, spot oil markets, oil futures markets, oil storage, cash & carry arbitrage, speculation, index investors.

Keywords: G13, G18, Q41.

Contango in Cushing?
Evidence on Financial-Physical Interactions in the U.S. Crude Oil Market

1. Introduction

While there has been considerable focus, especially in the aftermath of the 2007-08 oil price spike, on the role of financial speculators in influencing oil prices,[1] a question that lies at the heart of this debate -- how oil futures trading is related to spot oil prices – remains unresolved. A financial speculator who expects future oil prices to rise and wants to take a speculative position based on this expectation would typically go long in financial futures contracts. An index investor who wants to invest in oil will take a similar long position in futures contracts, which would be rolled over periodically.[2] If such speculative or investment activity increases the futures price sufficiently relative to the prevailing spot price,[3] a rational market response would be for arbitrageurs to step in to buy oil in the spot market and store it while simultaneously selling futures.[4] This "cash and carry" (C&C) arbitrage provides the mechanism that links oil futures and spot markets, since the withdrawal of oil from the market by arbitrageurs will cause spot prices to also increase.[5] Accordingly, a number of studies argue that if financial speculators or index investors drive up futures prices that, in turn, elevates spot oil prices above the level dictated by supply-

[1] See, for example, U.S. Senate Permanent Subcommittee on Investigations (2006), Masters (2008), Einloth (2009), Kaufmann and Ullman (2009), Sornette, Woodard, and Zhou (2009), Phillips and Yu (2010), Parsons (2010), and Singleton (2011).

[2] See, for example, Masters (2008).

[3] Singleton (2011) provides evidence of a significant effect of such investor flows on futures prices during the 2006-2010 period.

[4] This argument stems from standard financial market theory (reviewed in section 2) -- arbitrageurs have an incentive to simultaneously sell futures and buy oil in the spot market and put it in storage when the futures price exceeds the spot price by enough to cover net carrying costs (storage plus financing costs minus convenience yield), resulting in a riskless profit.

[5] Of course, in theory, financial speculators betting on a price run-up could also directly accumulate crude oil inventories, which would also increase spot oil prices. In practice, the higher financial leverage and lower transactions costs of trading futures relative to physical oil makes it much more likely that pure financial speculators will employ futures.

demand fundamentals, such an elevation in the oil price should be accompanied by a build-up in oil inventories.[6]

However, the available evidence of such an inventory build-up during the sharp 2007-08 oil price increase is mixed at best. Studies by the International Energy Agency (IEA) (2008), International Monetary Fund (IMF) (2008), and Organization for Economic Co-operation and Development (OECD) Working Party on Agricultural Policies and Markets (2010) find no evidence of a speculative increase in crude oil inventories in 2007-2008. The Interagency Task Force on Commodity Markets (ITFCM) (2008) argues that oil inventories were near historical levels in 2006-2008, while Hamilton (2009) concludes "in late 2007 and the first half of 2008, when the [oil] price increases were most dramatic, inventories were significantly below normal." Krugman (2008) makes the same point regarding the 2008 price run-up but does believe speculation contributed to higher prices in 2009 (Krugman, 2009). On the other hand, the U.S. Senate Permanent Subcommittee on Investigations (2006) argues that the behavior of inventories was consistent with speculation impacting cash prices and Einloth (2009) argues in support of a speculative build-up of inventory that accompanied the 2008 increase of oil prices from $100 to $140 a barrel but not during the preceding period.

The prerequisite for an inventory build-up as predicted above is a viable and active C&C market in crude oil. The existence of such a market cannot be simply assumed since there are many limits to arbitrage that would impede the functioning of such a market, such as the unavailability of non-operational storage (i.e., storage that is not reserved for operating purposes), pipeline and other transportation constraints, and financing barriers.[7] To our knowledge, there has been no in-depth research on the existence and functioning of a C&C market in oil, i.e., how oil inventories respond to changes in the futures-spot price spread, which should be the mechanism connecting financial market speculation and

---

[6] See, for example, ITFCM (2008), IMF (2008), IEA (2008), Krugman (2008), Hamilton (2009), Irwin, Sanders, and Merrin (2009), Smith (2009), and Kilian and Murphy (2010).

[7] See, for example, Shleifer and Vishny (1997), Etula (2010), and Acharya, Lochstoer, Ramadorai (2011).

spot oil prices.

This study focuses on the causal relationships between oil spot prices, futures prices and storage, specifically how storage is impacted by contango versus normal backwardation in oil futures prices. In other words, the study examines the relation between oil inventories and the spread between crude oil futures contracts.[8] Gaining an in-depth knowledge of this relationship is an important topic for academics, energy companies and traders, policymakers, regulators, as well as the general public, since it can deepen our understanding of the factors that move oil prices. No direct connection exists between the financial futures and physical spot prices of crude oil as contracts are rarely settled through delivery (Smith, 2009; IEA, 2008). The physical crude oil market is a highly competitive market in which prices are set by supply and demand. Thus, if crude oil futures trading impacts physical prices, it must do so by impacting either the physical supply or the physical demand. This puzzle of showing how financial futures influence physical spot prices is highlighted by Hamilton (2009), who notes that "The key intellectual challenge for such an explanation [of how future prices influence the associated spot commodity] is to reconcile…the price path with what is happening to the physical quantities of petroleum demanded and supplied." Thus, the financial futures market influences the physical spot prices by altering either the real physical demand or supply of crude oil. This study tests whether this influence can be traced through inventories. Smith (2009) advocates that, "The only avenue by which speculative trading might raise spot prices is if it incites participants in the physical market to hold oil off the market – either by amassing large inventories or by shutting in production." This paper tests if crude oil inventories increase (decrease) when the futures spread is positive (negative).

While the relation between futures spreads and inventory is not his primary focus, Singleton (2011) provides preliminary evidence of an active U.S. C&C market by graphing the relationship between

[8] From here on out when we refer to spread, we mean the spread between two crude oil contracts of different maturity. Typically, we will be referring to the spread between the two- and the one-month crude future contracts. The reasoning for this selection is provided in Section 3.2.1. The results of the study hold with other spread specifications also; in particular, the spread between the one-month future and the spot price, as well as the spread between the three- and one-month crude future contracts.

the spread across two- and four-month futures prices and the level of U.S. crude oil inventories, which suggests a tendency throughout the 2004-2009 period for inventories to increase when the futures market is in contango.[9] He notes also that this graphical pattern is even stronger when inventory levels from Cushing or Petroleum Administration for Defense District 2 (PADD2), the district which includes Cushing, are used. However, while he includes inventory changes as a conditioning variable in his formal analysis, he finds that the explanatory power is weak.

Einloth (2009) evaluates the relationship between spreads and inventories in his study of the role of speculation in the 2008 oil price behavior. However, he does not use inventories directly but rather the convenience yield as a proxy for inventories, derived from the prices of Brent crude oil futures. Additionally, in using the pricing of Brent futures to predict U.S. crude oil inventories, Einloth (2009) assumes a frictionless global oil market that, as our results suggest, may not be valid even within the continental U.S. In contrast, we minimize the effect of basis issues in our study by using the West Texas Intermediate (WTI) futures spreads to predict U.S. crude oil inventories, while carefully accounting for international oil flows that link the U.S. market with the global market. Doing so also minimizes the effect of storage measurement errors highlighted by Einloth (2009) and Singleton (2011). Additionally, we include controls that impact inventory levels and prices, such as supply and demand shocks, and other factors that influence inventories directly. We also perform a comprehensive analysis of which futures prices matter and which inventories are impacted, and investigate whether inventory levels adjust immediately to predicted levels or do so with a time lag. We therefore extend the current literature on inventories, spreads and the arbitrage role of inventories.

We find that over the 2004-2011 period crude oil inventories at Cushing were a significant positive function of the spread between the two- and one-month New York Mercantile Exchange West Texas Intermediate (NYMEX WTI) crude oil futures with a lag. We also find that over the 1992-2004 period

---

[9] Singleton's (2011) focus is on explaining returns in crude oil futures markets, which he shows were significantly affected by investor flows (specifically index investors and managed-money accounts) into the oil futures markets around the time of the 2008 oil price spike.

(before the Cushing inventories were reported separately), total U.S. non-Strategic Petroleum Reserve (SPR) inventories and PADD2 inventories were positive functions of lagged spreads. However, over the 2004-2011 period, neither total U.S. non-SPR inventories nor PADD2 inventories are significant functions of the spread once Cushing inventories are removed. None of the other four PADD inventories are significantly related to the spread over either period. Current crude oil inventories appear to be influenced by spreads over the last eight weeks or so. Our interpretation of this finding is that current spreads likely lead to contracts for forward delivery which do not result in a change in actual stock levels until delivery occurs sometime in the future. We observe basically the same results whether examining inventory levels or changes, and these results remain robust when we use different measures of the spread. We further find evidence that total U.S and most individual PADD inventories (but not at Cushing) are a negative (positive) function of the change in current (next week) refinery inputs and a positive (negative) function of the current (next week) imports. These results indicate that storage operators are able to partially anticipate crude oil shortages and surpluses and adjust their inventories accordingly. These findings establish, to our knowledge, the first tangible evidence documented in the literature of a causal link between oil futures and spot markets via inventory changes resulting from arbitrage, and raise several important questions for future research. In particular, our findings suggest that it would be fruitful for researchers looking to understand the impact of financial traders on the spot markets, especially the twin questions of (a) whether financial traders exacerbate or attenuate spot price volatility, and (b) whether they systematically affect the spot oil price level, to study the behavior of the C&C market over time.

We review the theoretical foundations of our study and discuss their empirical implications in the next section. We discuss our data and methodology in Section 3 and present our estimation of the crude oil inventory adjustment lag structure in Section 4. Our main results are presented in Section 5 and our robustness checks are in Section 6. Section 7 concludes.

2. The theoretical link between inventories and the futures-spot spread, and empirical implications

Inventories are connected to the spread through what is known as cash-and-carry (C&C) arbitrage. If the current (time t) futures price for delivery at time t+s, F(t,t+s), exceeds the current spot price, S(t), by more than the cost of storing oil from t to t+s (including transaction costs and net of any convenience yield) plus interest, SC(t, t+s), arbitrageurs can make a riskless profit by buying oil in the spot market for S(t), simultaneously shorting the futures contract at price F(t,t+s), and storing the oil. At time t+s, they can deliver on the futures contract collecting F(t, t+s).[10] Their time t+s profits adjusting for interest costs on the time t expenses are $F(t,t+s) - [S(t)+SC(t,t+s)](1+r)^{s}$.[11] For example, if crude oil spot price is $90, the one month futures price is $100 and the cost of storage is $6; it would make sense to sell the futures contract, purchase spot crude oil and store it for a month and then deliver on their futures contract, at a profit of about $4 per trade. Such arbitrage is profitable and oil inventories would be expected to rise at time t and fall at time t+s when

$$F(t,t+s) > [S(t)+SC(t,t+s)](1+r)^{s}.$$

This issue is important because it is the nexus between oil futures trading and physical oil prices. If we accept that physical energy prices, e.g., gasoline at the pump or oil at the wellhead, are determined by supply and demand, then C&C arbitrage is the mechanism through which futures market speculation could impact physical or spot prices.[12] If futures speculation pushes the futures price up enough to set off the arbitrage described in the previous paragraph, then the demand for oil and the spot price will tend to rise at time t when arbitrageurs buy oil to put into storage, and fall at time t+s when the oil comes out of storage thereby increasing the supply on the spot market.

While we have discussed C&C arbitrage from the point of view of a pure arbitrageur, a similar

---

[10] Due to convergence at maturity, actual delivery on the futures contract is not necessary. Arbitrage profits are approximately the same if the arbitrageur longs the futures contract at time t+s and sells in the spot market.

[11] This specification assumes the storage costs are paid at time t.

[12] Futures prices could also influence long-run supply by impacting drilling activity today or long-run demand by impacting conservation decisions but C&C arbitrage is the main short-run connection.

relationship holds for oil companies, pipelines, and others in the oil industry. When $F(t,t+s) > [S(t)+SC(t,t+s)](1+r)^s$, oil companies and others have an incentive to store and sell oil forward rather than sell in the spot market. Likewise, if $F(t,t+s)$ is far below $[S(t)+SC(t,t+s)](1+r)^s$ they have an incentive to draw down inventories by selling at time t. As noted above, SC(t, t+s) is net of any convenience yield, which is more important for oil firms. Producers, refiners, and marketers hold working inventories as buffers against supply interruptions and fluctuations in demand. When inventory levels are low, they run the risk of a stop-out or shortage. Thus there is an advantage or convenience yield to holding inventory. As inventories are reduced, the risk of a stop-out rises, raising the convenience yield and lowering SC(t,t+s). When inventories increase, the risk of a stop-out falls, lowering the convenience yield and raising SC(t,t+s). Thus, as Einloth (2009) and others point out, SC(t,t+s) varies positively with the inventory level and an ever-increasing difference between the futures price and the cash price is required to induce continued cash-and carry arbitrage when $F(t,t+s) > [S(t)+SC(t,t+s)](1+r)^s$.[13]

Speculative inventory levels should be related to past as well as current futures-spot price spreads. If the time t futures price for delivery at time t+s, F(t,t+s), exceeds the time t futures price for delivery at time t+v, F(t,t+v), where $s > v$, by more than the cost of storage from v to s, SC(t+v, t+s), plus interest, arbitrageurs can make a riskless profit by simultaneously (at time t) longing the t+v futures contact at price F(t, t+v) and shorting the t+s futures contract at price F(t,t+s). At time t+v, they would take delivery on the t+v contract paying F(t, t+v) and store. At time t+s, they would deliver on the t+s contract receiving F(t, t+s). Their time t+s profits adjusting for the interest or opportunity costs of the time t expenses would be $F(t,t+s) – [F(t,t+v)+SC(t+v,t+s)](1+r)^{s-v}$. Thus such arbitrage is profitable and oil inventories would be expected to rise at time t+v and fall at time t+s when

$$F(t,t+s) > [F(t, t+v)+SC(t+v,t+s)](1+r)^{s-v}.$$

Note that in this case, there is no immediate change in inventories. Also, in this case physical

[13] During prolonged contango markets additional crude storage facilities can be constructed which would decrease SC(t,t+s) allowing the futures-spot spread to remain at lower levels in order to achieve profitable C&C arbitrage.

prices tend to be pushed up at future time t+v when the oil is taken off the market and placed in storage and pushed downward at time t+s when the oil comes out of storage and back on the market.

Considerable anecdotal evidence indicates that C&C arbitrage occurs. For instance, several newspaper articles published in 2007 described increasing and decreasing levels of inventory at Cushing, OK, the NYMEX delivery point for the WTI contract, and related the activity to the C&C type arbitrage.[14] The time-series relation between the futures spread and Cushing inventories is graphed in Figure 1. Note that Cushing inventories are positively correlated with the futures spread as predicted by C&C arbitrage. Note also the sharp increase in storage capacity between 2004 and 2011, which some reports tie to building additional capacity for C&C arbitrage.[15]

****** Place Figure 1 about here ******

As noted by the IEA (2008) and others, given the central role that inventories play in the futures price - cash price nexus and the significant interest in the question of how much, if at all, speculation impacts physical oil prices, it is surprising how little research has been done on the relation between the futures-spot spread, $F(t,t+s)- S(t)(1+r)^s$, or for simplicity F(t,t+s)-S(t), and inventories. While some studies have noted simple correlations between inventories and F(t,t+s)-S(t), a careful multivariate approach is needed for several reasons. First, without controlling for other factors that impact inventories, simple correlations do not establish that inventory levels are responding to the futures-spot spread. Suppose, for instance, that demand falls unexpectedly. In that case, inventories would rise and S(t) would tend to fall raising F(t,t+s)-S(t). Thus inventory levels and F(t,t+s)-S(t) would move together but not because inventories are responding to F(t,t+s)-S(t), and their correlation would not constitute evidence that futures speculation impacts cash prices through inventory behavior. Studies that carefully examine how oil futures prices impact spot prices through inventory and production controlling for other changes in supply and

[14] Davis, Ann "Where Has All the Oil Gone?" *Wall Street Journal*, October 6, 2007.

[15] While we do not have direct data on Cushing, OK storage capacity, the amount of crude oil stored in Cushing between April 9, 2004 and April 8, 2011 increased by 259%, from 11,677 to 41,896 thousand barrels.

demand appear warranted. Second, as explained above, current inventory levels and changes should be a function of past, as well as current, futures and spot spreads. Third, most crude oil inventories are held for operational purposes, rather than for speculation or arbitrage, so controlling for factors that influence operational inventory levels should enable better estimates of the impact of the futures spread.

3. Data and Methodology

3.1 Data

In order to estimate the relationship between the spread and crude oil inventories, we obtain weekly ending inventories of crude oil for: 1) U.S., excluding the SPR, the five PADD districts, and Cushing from the Energy Information Administration (EIA) website from 9/11/1992 (4/09/2004 for Cushing) through 7/08/2011. We also obtain weekly data on U.S. oil production levels, imports, refinery inputs, NYMEX WTI future contracts for the first four months, and Cushing WTI spot prices. The variables used are described in Appendix 1; their descriptive statistics are in Table 1, while the correlations between different variables are in Table 2.

****** Place Tables 1 and 2 about here ******

Crude oil is traded on both the spot and the futures market. In the U.S., crude oil futures trade primarily on the NYMEX. The main crude oil futures contract is for the WTI grade of crude oil and it settles at Cushing, OK. While a variety of spot locations are priced, their prices are typically perceived in terms of the basis to the NYMEX WTI crude oil price. A variety of crude oil counterparties, both producers and users, need to buy and sell crude oil physically in the spot market. However, if they need to hedge their exposure forward they need to participate in the futures market. Trading in the WTI crude oil contract ceases on the third business day prior to the 25th calendar day of the prior month. For example, trading in the August contract ceases near the end of July. Thus traders who do not wish to make or take delivery of WTI crude oil at Cushing must reverse their positions prior to this date. If they do not reverse, physical delivery of settled crude oil occurs at Cushing, OK, over the full length of the contract month, i.e., August in our example.

Cushing, OK is a special location for crude oil contracts because physical settlement of the future market transactions occurs there. The other crude oil districts in the U.S., which are the five Petroleum Administration for Defense Districts (PADD's) that the entire U.S. territory is broken into, are equally important especially from the standpoint of product supply and distribution.[16] PADD 1 covers the East coast, PADD 2 the Midwest, including Cushing, PADD 3 the Gulf Coast, PADD 4 the Rocky Mountains and PADD 5 the West Coast.[17] Given that the futures crude oil contract settles at Cushing, OK, the traders involved in C&C arbitrage have an incentive (as discussed below) to locate their storage facilities there. While operational drivers of crude oil inventory are important in all PADDs, the spread and its influence on inventory via C&C arbitrage should be most observable in Cushing.

We have contacted several pipeline and storage operators at Cushing, OK, concerning common institutional arrangements, such as delivery mechanisms and contracts, and speculative strategies. Of the major operators in Cushing, we have interviewed representatives from Plains All American Pipeline, Magellan and Enterprise. All of the above firms lease out storage to customers mainly via longer-term full tank leases or capacity leases. The tank leases are typically done for five year periods. Capacity leases allow several customers to have common stream crude oil in the tank. The main customers for storage leases are refineries, but Exploration and Production (E&P) firms, large physical oil trading firms, as well as trading arms of different banks also lease storage. The operators also said that crude oil deliveries are scheduled months ahead and trading in the spot market occurs in emergency situations. This influenced our choice of the spread for this study.

The data series on crude oil inventory levels exhibit unit roots, which may be due in part to persistent time trends in the data. To avoid issues with unit roots in crude oil inventory levels, we use

---

[16] The PADD's were originally created during World War II for gasoline rationing.

[17] PADD 3 is home to the U.S. Strategic Petroleum Reserve (SPR) which is a large reserve created for national security purposes. The data used in this study excludes crude used for SPR inventories due to the nature of these reserves.

changes in inventory (first difference) in all our analysis reported here.[18]

3.2 Specification issues

In specifying the model to estimate the relationship between crude oil inventories and spread, we face four issues. First, since there are numerous futures contracts with different maturities, it is important to address the question of which of these contracts should be used to measure the spread -- the nearby contract, the futures contact maturing in two months or in three months, etc. Second, what is the appropriate lag and lag structure? In other words, does a change in the futures-spot spread impact inventory levels quickly or does it take some time? Third, what inventory data should we focus on? Fourth, how should spurious correlation or endogeneity be controlled for? As described below, unexpected shifts in supply and demand should impact both spot oil prices (and hence the futures-spot spread) and oil inventory levels. Thus, if not controlled for, the estimation might pick up this spurious correlation instead of the impact of the spread on inventories. We next discuss our thoughts on and approach to each of these.

3.2.1 Futures-spot and futures-futures spreads

Most of the time, prices of the nearby futures contract and the Cushing WTI spot price are approximately the same except during the roll period; therefore, the spread between them is of little use for our analysis. Spreads between the spot price and the price of any other futures contract, and between prices of different maturity futures, could conceivably set off C&C arbitrage. For example, if the third month futures exceeds the spot by more than storage and transaction costs, that could also set off C&C arbitrage and an increase in inventory levels. Or if the third month futures exceeds the second month futures by more than storage plus transaction costs, arbitrageurs could contract to take delivery in the second month and deliver in the third, so inventories would increase a month in the future. Similarly, if the two-month ahead futures price exceeds the one-month ahead futures price by more than storage plus transaction costs, C&C arbitrage could take place as arbitrageurs contract to take delivery in the next month and deliver a month

[18] We have replicated the analysis reported in this paper using inventory levels data that is detrended and seasonally adjust using a process available from the authors. The results using the detrended levels data, which largely support the conclusions from the change data, are available from the authors.

after next. Hence in our view, the question of the best spread to use is an empirical one.

As it turns out the market is normally in continuous backwardation or contango over the first few months so that the different spreads are highly correlated. For instance, the correlation of the spread between the second month contract and the spot and the spread between the third month contract and the spot is 0.982. The correlation between the third-month-spot spread and the fourth-month-spot spread is 0.995. Hence it makes little difference which spread we use; any one spread tends to pick up the effect of all on inventory levels. Storage operators in Cushing explained during our conversations that only emergency trading is done in the spot market and most crude oil deliveries are scheduled a month ahead. Therefore, we chose to use the spread between the two- and the one-month crude oil future contracts. This spread also has fewer outliers than the two-month ahead to spot spread while the correlation between the two is 0.942. The correlation between the actual spot price and the one-month ahead future contract is 0.9999. Therefore, most of the changes in the spot price are reflected in the one-month futures with the exception of those that are very temporary in nature and are not expected to persist past the current month. Here we report results for the spread defined as the difference between the two- and the one-month crude oil future contracts but the results are virtually the same using the two-month to spot spread, three-month to spot and three-month to one-month spread.

3.2.2 Lag structure

In estimating the relationship between the futures spread and crude oil inventories, one issue is what lags to expect between the spread and inventories and how to specify the lag structure. As explained earlier, we expect today's inventory levels to depend on past spreads since current changes in inventories may be due to contracts signed weeks or months ago. As noted previously, we have also reached out to major storage operators at Cushing to ascertain common delivery arrangements and representative storage and transaction costs.

As we see it, if today's spread is sufficient to set off C&C arbitrage, it may be weeks or days before inventories increase since we only observe actual physical changes in inventories, not the contracts being executed for future delivery. For example, on July 5, 2011 the Cushing spot price was $96.89 and NYMEX futures prices were: $96.89 for August 2011, $97.38 for September; and $97.87 for October 2011. While we are seeking hard data on storage costs, a fairly common rough estimate is about $0.40 a barrel per month which would place the July 5 futures price structure above the breakeven point for profitable C&C arbitrage. July crude oil deliveries were scheduled at the end of June, so if traders want to take non-emergency delivery of crude oil they have to purchase the August contract. If we assume for the moment that storage and transaction costs for the August contract were about $0.40 per month so that C&C arbitrage would be profitable, arbitrageurs might purchase oil for delivery in August and simultaneously short the September futures. In this case inventories would rise over the month of August, meaning anytime between one to eight weeks from today.[19] The October-September spread is also fairly large so another alternative is that today arbitrageurs might long the September contract, while shorting the October contract, and subsequently take delivery on the September contract in September, and make delivery on the October contract in October. In this case, inventories would not rise until September even though the contracts are set in July. Thus, lags of several weeks or even months are quite likely, but beyond that it is hard to say what the lag relationship is. Our interviews with storage operators suggest that in general, the lag could be anywhere from one to nine weeks but do not provide more specificity beyond that. Thus we turn to the data to see what the lags look like.

3.2.3 Choice of inventory locations

While C&C arbitrage is possible utilizing storage at any location, storage at Cushing offers the arbitrageur significant advantages. Suppose that at time t, an arbitrageur buys and stores the oil at location X and shorts the futures contract maturing at time t+s. At time t+s, she transports the oil to Cushing and delivers on the futures contract. In order to be profitable, the futures contract must exceed the location spot

[19] Trading in the WTI crude contract ceases on the third business day prior to the 25th calendar day of the prior month. For example, August 2011 contract would have stop trading and settled physically on July 20, 2011 and then delivery would get scheduled between August 1 and August 31. So, if an August 2011 crude future is purchased on July 5th, it can get delivered anytime between August 1 and August 31. That implies a waiting time between the trade and actual inventory increase in Cushing of four to eight weeks. However, if the August 2011 crude future was purchased on July19, 2011 then the waiting time between the trade and actual inventory increase in Cushing would be between one and a half and six and a half weeks.

price by more than the cost of storage plus the cost of transporting the oil. In addition to arranging storage the arbitrageur must arrange transportation as well.

Instead of delivering the oil to Cushing, the arbitrageur may buy, store, and sell the oil at location X. In this arbitrage, she shorts the futures at the beginning of the arbitrage and longs the same futures contract shortly before trading ceases. As long as the basis or differential between the price at location X and Cushing is constant, this strategy is profitable if and only if the futures-spot or futures-futures spread exceeds the cost of storage (at location X). However, if the basis changes over time, then the profitability is uncertain. In other words, C&C arbitrage utilizing storage and delivery at non-Cushing locations involves a basis risk which is not present if the storage and delivery are at Cushing.[20] For this reason, we focus particular attention on Cushing inventories but also examine the impact of the spread on storage away from Cushing.

3.2.4 Controlling for spurious correlation

A fourth issue in estimating the relationship between the spread and crude oil inventories is how to correct for the spurious correlation caused by unforeseen shifts in supply and demand. For instance, if there is an unforeseen increase in demand, this would tend to lead to a fall in crude oil inventories and at the same time an increase in spot prices, which would mean a fall in the spread. Hence a positive correlation between changes in the spread and changes in inventories would be observed but due to the impact of the demand shift on both prices and inventories - not to C&C arbitrage. Likewise a sudden unforeseen increase in supply would tend to cause a simultaneous increase in crude oil inventories and in the spread. To control for this to the extent possible, we include the changes over the current week in: 1) U.S. crude oil production levels, 2) imports (overall net and by PADD), and 3) refinery inputs (overall and by PADD) between weeks t-1 and t. Consider the change in refinery inputs. The change from the previous week consists of a planned or expected change plus the unplanned or unexpected change. If refinery production increases unexpectedly, this would lead to an unexpected decline in crude oil inventories. Thus to the

[20] On the other hand, storage may be less expensive away from Cushing.

extent part of the change in refinery inputs is unexpected, we expect it to be negatively correlated with the change in crude oil inventories. Similarly, to the extent changes in U.S. crude oil production and imports are unexpected, we expect them to be positively correlated with changes in crude oil inventories. In addition, we include the change in the spot WTI price as an additional variable separate from the spread. If an unexpected change in demand or supply is viewed as temporary, it will tend to impact the spot price but not the futures price. Thus this variable should have a negative coefficient and pick up additional unforeseen shifts in supply and demand which impact both the spread and crude oil inventories. Note that if the shift is seen as permanent so that both spot and futures prices change, there is no spurious correlation problem.

While including current week changes in refinery inputs, imports, and production as independent variables helps control for correlation between changes in the spread and changes in inventory induced by unforeseen shifts in supply and demand, coefficients of these variables must be interpreted with caution. We cannot distinguish between expected and unexpected changes in these variables. By definition, if the data is perfect, the change in inventories this week is equal to the level of imports plus the level of production minus the level of refinery inputs. Since the levels of imports, production, and refinery inputs are by definition equal to the sum of all current and past changes in these variables, there is a small built-in positive correlation between current changes in imports and production and the change in inventories, and a small built-in negative correlation between change in refinery inputs and the change in inventories. Thus positive coefficients for current week changes in production and imports and negative coefficients for refinery inputs need not necessarily indicate an effect of unforeseen changes in supply and demand on inventories.

We also seek to control for other factors that impact desired inventories at time t. Of course, operational inventories are held to bridge any gap or mismatch between supply and demand. Specifically, refineries hold inventories to bridge mismatches between their crude oil supplies and refinery needs. If refinery draws are expected to be larger next week than the combined production and imports supply, then there would be a tendency to hold large current inventories in order to “stock up” for next week. Likewise, if next week's production and imports are expected to be higher than refinery needs, current inventory levels should be smaller. While we cannot observe expected future imports, production, and refinery imports we can observe *ex post* levels and changes. Viewing the actual change as proxies for expected changes, we add lead measures of the changes in refinery inputs, U.S. crude oil production and imports over the week from t to t+1 to our set of independent variables. Note that the expected signs for these lead variables are opposite to those for the current week. We expect a negative coefficient for the current week change in refinery inputs and a positive coefficient for the change next week. We expect positive coefficients for current week's changes in imports and production and negative for the changes next week. The rationale for the current week variables is to pick up the effect of unexpected changes on actual inventories; the rationale for the lead variables is to pick up the effect of expected future changes in these variables on desired inventories.

4. Polynomial distributed lag estimation of the spread on crude oil inventory

We examine the influence of the change in the spread on the change in inventories. Expecting some seasonality in inventory patterns, we control for this with dummy independent variables. Since 52 separate weekly dummy variables are neither feasible nor appropriate, we assume that any seasonality can be captured by a polynomial form. First, we define weekly dummy variables as follows: $w_1 = 1$ if the observation is the first week in January and 0 otherwise, $w_2 = 1$ if the observation is for the second week in January and 0 otherwise, and so forth through $w_{52} = 1$ the last week in December and 0 otherwise. We then specify five dummy variables z_k where z_1 is a zero-degree polynomial of the w_i's, z_2 is a first degree polynomial, z_3 is a second degree polynomial, z_4 is a third degree polynomial, and z_5 a fourth degree polynomial. The graph of storage pattern from the z variables is presented in Figure 2 for Cushing and Figure 3 for the U.S. and the z variables are defined in the following manner:

$$
\begin{aligned}
z_1 &= w_1 + w_2 + w_3 + \cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot + w_{52} \text{ (picked up by the intercept)}\\
z_2 &= w_1 + 2w_2 + 3w_3 + \cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot + 52w_{52}\\
z_3 &= w_1 + 2^2 w_2 + 3^2 w_3 + \cdot\cdot\cdot\cdot\cdot\cdot\cdot + 52^2 w_{52}\\
z_4 &= w_1 + 2^3 w_2 + 3^3 w_3 + \cdot\cdot\cdot\cdot\cdot\cdot\cdot + 52^3 w_{52}\\
z_5 &= w_1 + 2^4 w_2 + 3^4 w_3 + \cdot\cdot\cdot\cdot\cdot\cdot\cdot + 52^4 w_{52}
\end{aligned}
$$

We estimate the impact of the change in spreads up to twelve weeks ago on the change in current inventory using the following model:

$$\Delta STOCK_{i,t} = \beta_{0,i} + \sum_{k=1}^{4} \beta_{i,k}\, Z_{k,t} + \sum_{j=5}^{16} \beta_{i,j}\, \Delta SPREAD_{j,t} + \sum_{j=17}^{J} \beta_{i,j}\, \Delta Y_{i,j,t} \tag{1}$$

where $\Delta \mathrm{STOCK}_{i,t}$ is the change in inventories between weeks t-1 and t at one of the following locations i: U.S., Cushing, PADD1-PADD5, as well as U.S. without Cushing and PADD 2 without Cushing. $Z_{k,t}$ variables control for seasonality and are created as described above, $\Delta \mathrm{SPREAD}_{j,t}$ is the change in the spread between the two- and one-month WTI crude oil future contracts and its lags going back twelve weeks, and $\Delta \mathrm{Y}_{j,i,t}$ represents other possible factors j impacting inventories in region i, including current and lead changes in imports, production, and refinery inputs. In addition, but not shown in Equation (1), autoregressive and moving average lagged error terms are included as needed to remove autocorrelation in the residuals.

As discussed above, in the absence of measurement error, the change in inventories for the U.S. (including SPR) would equal imports plus production minus refinery inputs. Thus including flow *levels* of these variables would result in built-in correlation. Instead of flow levels, we include *changes* in import, production, and refinery input flows in $\Delta \mathrm{Y}_{j,i,t}$ to capture changes in inventories. These variables smooth out temporary mismatches between (a) crude oil additions plus production and imports, and (b) crude oil withdrawals minus the refinery intake. Not only current but also coming week changes in these variables are included. Next week, or lead, variables are utilized as proxies for expected changes in crude oil additions and withdrawals. We anticipate that the change in inventories should be a positive function of the changes in imports and U.S. production over the current week and a negative function of the changes in these two variables over the coming week. The change in inventories should be a negative function of the change in refinery inputs this week and a positive function of the change in refinery inputs next week.

In order to impose some structure on the spread coefficients and improve the efficiency with which they are estimated, we condense the twelve lagged spread variables in Equation (1) to four. Accordingly, we estimate a polynomial distributed lag (PDL) model in which the coefficients of the lagged spreads follow a fourth degree polynomial. PDL_j is a j-1 degree polynomial of the twelve lagged spread. In other words, PDL_1 is a zero degree polynomial, PDL_2 a first degree polynomial, etc. Details of the structure are available in Appendix 2. Therefore, our final model is:

$$\Delta STOCK_{i,t} = \beta_{0,i} + \sum_{k=1}^{4} \beta_{i,k}\, Z_{k,t} + \sum_{j=5}^{8} \beta_{i,j}\, \Delta PDL_{j,t} + \sum_{j=9}^{J} \beta_{i,j}\, \Delta Y_{i,j,t} \qquad (2)$$

5. Results

The weekly change in Cushing inventory is a positive function of the change in the current and past spreads but changes in other PADDs' inventory are not. Changes in PADD 2 inventories, which include Cushing, are a significant function of change in the spread when the inventory figures include Cushing, but not when Cushing inventories are excluded. Similarly, changes in overall U.S. (non-SPR) inventories are a positive function of spread 1992-2004 but not for the 2004-2011 period. The implication is that, at least since 2004, C&C arbitrage is largely concentrated at Cushing.

5.1 Cushing results

Estimation results for Equation (2) are presented in Tables 3, 4, 5 and 6 for weekly changes in inventories at Cushing, U.S. non-SPR, PADD 2, and PADDs 1, 3, 4, 5, respectively. The Cushing equation is estimated from 4/16/2004, when the Cushing data is first available, to 7/8/2011. Equations for other areas are estimated over various periods including 1992-2011. ΔPDL coefficients shown in Panel A of the tables are for the polynomial variables calculated from the spread changes over the current and twelve past weeks. Their joint significance is tested with the Wald p-values shown at the bottom of Panel A of the tables. Panel B of the tables shows the implied coefficients for the twelve lagged spread differences calculated from the ΔPDL coefficients in the top of the tables. Thus, in Table 3 Panel B, the coefficient for the current spread is 282.03, the coefficient for last week's spread is 324.29, and the coefficient for the week before that is 333.20. Above we argued that due to delivery arrangements, there may be a considerable delay between the arbitrage trades in response to the spread and the actual change in crude oil inventories since we observe actual inventories but not contracts for future delivery. The estimated coefficient pattern in the bottom part of Table 3 is certainly consistent with this. Indeed, coefficients for the spreads two, three and four weeks ago are actually higher than the coefficient for the current week.

****** Place Tables 3, 4, 5 & 6 about here ******

The lag pattern shown in Panel B of Table 3 shows that we observe positive inventory changes through at least the first eight weeks. The largest increases are observed over the first four weeks or so. Then the increases start declining. Cushing results in Table 3 imply that a one- time $0.10 increase in the spread leads to a positive change in inventories of about 28,203 barrels the first week, approximately 153,241 barrels after four weeks, and 211,142 barrels after nine weeks.

Cushing data is only available from 2004-2011 and the Cushing results reported in Table 3 pertain to this period. Model 1 evaluates actual changes in Cushing inventories, while Model 2 pertains to a modified version where Cushing inventory changes are winsorized at the 1% level to control for outliers. Cushing inventory changes are mainly explained by the spread. No operational variables, other than U.S. oil production, significantly influence Cushing inventory changes. The removal of the PDL spread lags from the regressions drops the adjusted R^2 for Cushing models by 55% from 15% to 7%. However, the same removal for overall U.S. (non-SPR), PADD 1, 3, 4, and 5, and PADD2 non-Cushing inventories decreases the adjusted R^2 only slightly, if at all. The overall results are similar when the change in Cushing inventories is winsorized. Table 3 also shows evidence of a seasonal pattern in Cushing inventories as reflected in the z variable coefficients.

The graph of the Cushing seasonal pattern in inventories as implied by the z variables is presented in Figure 2. The pattern of inventory levels shows a short period of crude oil withdrawal from storage at the beginning of the year (weeks 1-3) after which inventory additions begin (week 4) and last through May (week 20), at which point withdrawals resume and continue through autumn, with additions to storage resuming in late autumn (week 40) and persisting through the end of the year. Figure 2(a) also shows that inventory levels in Cushing have grown significantly over 2004-2011 as more storage facilities have been built. The average annual inventory increase is shown by the difference in the starting and ending points of the inventory "level" series. Figure 2(b) adjusts for this capacity increase by normalizing the graph scale in Figure 2(a) to start and end at the same level, and shows the seasonal pattern on this adjusted scale.

****** Place Figure 2 about here ******

5.2 Results for total U.S. above-ground, on-shore storage.

Results for the U.S. (non-SPR) change in inventories are presented in Table 4. In Model 1, we present estimations of Equation (2) for the full 1992-2011 period and, in Model 1-winsorized, we use the same data as in Model 1 but winsorize the dependent variable—the change in U.S. inventories--at the 1% level. In Table 4 Model 2 we estimate the equation for the 1992-2004 period, over which separate Cushing storage data is unavailable. In Model 3, Equation (2) is estimated over the 2004-2011 period, including Cushing storage inventories and in Model 4, over 2004-2011, excluding Cushing. The changes in the spread are significant in explaining U.S. inventory changes in Models 1-2 for the 1992-2011 and 1992-2004 periods respectively, but not in Models 3-4 which cover the 2004-2011 period. However, over 2004-2011, the change in spread is the main significant explanatory variable for Cushing inventory changes, as shown in Table 3. This indicates that the seeming relationship between the U.S. (non-SPR) change in inventories and the spread is largely driven by Cushing inventories.

As shown in Table 4, the spread is not significant in explaining overall U.S. (non-SPR) inventories over 2004-2011, whether or not Cushing inventories are included in the total. Consistent with this, as reported in Table 6, none of the PADD 1, 3, 4 or 5 district inventories appear influenced by the spread. Combined with the findings from Table 3 that Cushing inventories are a significant positive function of the spread over 2004-2011, this indicates to us that most C&C arbitrage in recent years has been confined to Cushing. However, it bears noting that the inventory figures include above-ground on-shore inventories only. We cannot rule out C&C arbitrage involving off-shore storage in tankers, or producers opting to leave oil in the ground in response to the futures spread.

The insignificant result in Model 3 of Table 4 could be due to the largely operational role of inventories outside Cushing since the operating parameters are very significant). Model 3 of Table 4 reports U.S. results over 2004-2011 with Cushing included but shows that the spread is not significant in explaining inventory. In contrast, Model 3 of Table 5 reports PADD2 results over 2004-2011 with Cushing included and shows that spread is significant. The implication here is that when Cushing is included in overall U.S. numbers, it loses power to explain overall U.S. inventory changes especially since the explanatory power of operating variables in all other PADDs increases.

While we found little impact of the operational variables on inventories at Cushing, they are important in explaining total U.S. inventories. Moreover, the explanatory power of operational variables grows overtime, as the R^2 over 2004-2011 (69%) is much higher than the R^2 over 1992-2004 (21%). As expected, we find a significant positive relation between inventory changes and changes in both imports and U.S oil production over the current week and a negative relation between the change in inventories and the change in refinery inputs. Also, as expected, the signs are reversed for changes in imports, U.S. production, and refinery inputs over the coming week, though the coefficient for the coming change in U.S. production is not significant. The results for the lead variables indicate that if refinery inputs are expected to be higher (lower) next week than imports and production, then storage operators tend to increase (decrease) current inventories in anticipation. There is also evidence of a seasonal pattern as reflected in the z variable coefficients.

The graph of the U.S. (non-SPR) seasonal pattern in inventories as implied by the z variables is presented in Figure 3. In Figure 3(a) we presents the changes and levels of U.S. crude oil inventories as predicted by the z pattern over the 1992-2011 period; in Figure 3(b) over the 1992-2004 period; and in Figure 3(c) over the 2004-2011 period. The pattern of crude oil inventory levels shows that crude oil additions to storage start in early autumn (around weeks 36-38), followed by a quick period of crude oil withdrawals over the last couple of weeks of the year, after which additions resume at the start of the new year and continue through May (around week 20). Then in May crude oil starts being withdrawn as inventories decrease. Crude oil withdrawals last over the summer and early autumn months, till the end of September (around week 35). Figure 3(b) shows that overall U.S. crude oil storage capacity did not increase significantly over 1992-2004, as the line graphing inventory levels converges to its starting point. This differs from Figure 3(c) which shows that capacity increased over the 2004-2011 period. This capacity increase is partially attributed to Cushing, as can be seen from Figure 2.[21]

****** Place Figure 3 about here ******

5.3 PADD 2 results

Results for PADD 2 crude oil inventories are presented in Table 5. We present PADD2 results separately from other PADD district because it physically includes Cushing. In Model 1, we present Equation (2) estimations for the full 1992-2011 period and, in Model 1-winsorized, we use the same data as in Model 1 but winsorize the dependent variable—the change in PADD 2 inventories--at the 1% level. In Model 2 we estimate the equation for the 1992-2004 period, over which separate Cushing storage data is unavailable. In Model 3, Equation (2) is estimated over the 2004-2011 period, including Cushing storage inventories and in Model 4, over 2004-2011, excluding Cushing. The changes in the spread are significant in explaining inventory changes in Models 1-3 but not in Model 4 which excludes Cushing over the 2004-2011 period. So, unlike total U.S. (non-SPR) and PADD 1, 3, 4, and 5 inventories, PADD 2 inventories are a significant function of the spread over 2004-2011 until we exclude Cushing. Furthermore, if we remove all spread lags from the list of independent regression variables, the adjusted R^2 declines for PADD 2 inventory changes when Cushing is included from 17% to 12%, but it remains at 5% when Cushing is excluded from PADD 2 inventories. This shows that Cushing inventories are driving the explanatory power of spread for PADD 2 inventories. We conclude that there is little evidence of above-ground C&C

---

[21] While we do not have capacity data, we can partly infer capacity from levels since level will not exceed capacity. The overall U.S. (non-SPR) weekly crude levels did not change a lot (unlike Cushing) – from around 333,494 (October 30, 1992) to 355,456 (July 8, 2011) thousand barrels. In Cushing there was over a 250% increase in levels, which is unlikely to have been accomplished without capacity increases.

arbitrage in PADD 2 outside of Cushing.

Changes in PADD 2 inventories are partially explained by current, but not lead, operational variables. PADD 2 inventories are a positive function of current week imports and production and a negative function of refinery inputs. This is in line with the results for overall U.S. inventories, where the operational variables also explain inventory changes, but different from the results for Cushing, where the operational variables are insignificant. Evidence of a seasonal pattern, as reflected by the z variables, also exists in PADD 2 as in all the other areas.

5.4 Results for other PADD districts.

Results for the other individual PADD districts, excluding PADD 2, are presented in Table 6. There is no evidence that changes in the spread impact inventory changes in these regions, as the PDL lags of the spread are neither jointly nor individually significant. However, inventory changes in most PADD districts are a positive function of recent changes in imports and a negative function of imports over the coming week. They are a negative function of recent changes in refinery inputs and a positive function of the change in refinery inputs over the coming week.

Overall, our results from Tables 3-6 suggest that over the 2004-2011 period the spread is significant in explaining inventories mainly at Cushing indicating that above-ground on-shore C&C arbitrage in the U.S. currently takes place largely in Cushing.

6. Robustness Checks

In this section we carry out check of robustness of the results reported in the previous section to alternative specifications. In results that are available from the authors, we have repeated our analysis where the dependent variables are levels, rather than the changes, in inventories. Results largely parallel those in section 5. Several alternative specifications for Cushing are presented in Table 7 and for the total U.S. (no-SPR) in Table 8, as well as for both in Tables 9-11. These alternatives include adding lagged control variables, evaluating the percentage instead of the barrel change in inventories and examining the role of cross-PADD imports. We also estimate our models without the PDL structure for the spreads and using month dummies instead of the z weekly variables to adjust for seasonality. Finally, we present our results using two different measures of the spread, instead of the original difference between the two- and one-month NYMEX WTI futures, we use the difference between the two-month future and the spot WTI prices, as well as the difference between the tree- and two-month NYMEX WTI futures.

We include the percentage change in inventories as an alternative specification to partially control for the large growth in capacity, especially at Cushing, over the evaluation period. We also consider the possibility that Cushing inventories might be influenced by the inflows and outflows of crude oil from other PADDs. While imports are not significant in explaining the Cushing inventory changes, our imports measure does not contain imports from other PADDs. After studying the correlations for storage and imports for different PADDs and consulting EIA data we establish that the largest exchange of crude oil occurs between PADD 2 and PADD 3. We therefore include changes in PADD 3 storage and imports as controls in Cushing regressions. These alterations do not change the significant role of the spread in explaining Cushing inventories.

We add lagged control variables for several reasons, of which the first is to reduce the influence of asynchronous data reporting for inventories and the above mentioned controls. In this case, we expect the signs on the lagged coefficients to be the same as for the ones for the current week—negative for refinery inputs and positive for production and imports. The second reason to add lagged control terms deals with "reaction adjustment" in the following manner: if refinery inputs were higher than expected last week, this week there could be a tendency to restock inventory and if production and imports were higher than expected last week, this week there could be a tendency to "drain down" inventory. In this case, we would expect the lagged control coefficients to be of a different sign than the ones for the current week—positive for lagged refinery inputs and negative for lagged production and imports. However, like the current week imports, production, and refinery input variables, the lags of these variables may be correlated with the change in inventories due to their correlation with the levels of these variables. This additional specification again does not alter our main results.

We also we want to explore whether our PDL structure for the twelve spread lags is responsible for our results, i.e., to see if the main results are robust to including the twelve individual spreads. The same issue applies to using the four seasonal z variables. We are also interested in whether these specifications result in more efficient estimates of the spread and seasonal patterns.

Finally, we want to evaluate the relationship between crude oil inventories and the spread using different measures of the spread to ensure that our results are not driven my our chosen spread measure—the difference between the two- and the one-month NYMEX WTI futures. We present two alternative spread measures—the difference between the two-month future and the spot WTI prices, as well as the difference between the three- and one-month WTI futures. The correlation between different spread levels is high, over 94%. But we use spread changes, not levels, in our models and while the correlation between our main spread measure, the difference between the two- and one-month WTI futures, and one of the alternatives, the difference between the three- and one-month WTI futures, remains high at 96%. The correlation between our main spread measure and another alternative, the difference between the two-month futures and the spot WTI prices, is only 64%.

Alternative models for Cushing over the 2004-2011 period are presented in Table 7. In Model 1, we use an alternative dependent variable—instead of the barrel change we use the percentage change in inventories. In Model 2, we add lagged control variables for crude oil refinery inputs, production and imports. In Model 3, we add controls for possible transfers of crude oil from PADD3 into PADD2, where Cushing is located. Cushing changes in inventory remain a positive significant function of the spread in all specifications in Table 7.

****** Place Table 7 about here ******

Model 1 results in Table 7, in which the percentage change in inventories replaces the barrel change as the dependent variable, echo the Table 3 results in that changes in the Cushing inventory are again a significant positive function of the spread. The change in production is the only operational control variable that explains Cushing inventory changes. There is again evidence of a seasonal pattern as reflected in the z variable coefficients.

The implied coefficient pattern for the twelve lags of the change in spread in Table 7, Model 1, Panel B is similar to that of Table 3 Panel B, as it again shows that spread influences inventories with a lag. Table 7 Panel B shows that inventory increases over about the first eight weeks. A one-time $0.10 increase in the spread leads to 1.25% increase in Cushing inventories over the first week, approximately 6.23% increase over four weeks and 8.34% increase after nine weeks.

Model 2 in Table 7 adds lagged control variables. Cushing inventory changes are still a significant positive function of the change in spread. Again none of the control variables, except the current production changes and the lagged refinery inputs change, are significant in explaining Cushing inventory changes. The lagged refinery inputs change is positively related with current Cushing inventory changes. This is inconsistent with our non-synchronous data interpretation but in line with but with “reaction adjustment” interpretation—if the refinery draws were higher than expected last week then there would be a tendency to restock this week.

Results in Model 3 in Table 7, in which PADD 3 imports and inventories are included as explanatory variables for Cushing inventories, are interesting. Changes in Cushing inventories are still a significant positive function of the spread. But also changes in PADD 3 imports are significant in explaining Cushing inventory changes, while PADD 3 inventory changes are not. No other control variables, not even production, which exhibited explanatory power in all other models, are significant in explaining Cushing inventory changes. PADD 3 imports may matter because a part of crude oil coming into PADD 3, which includes the largest U.S. crude oil port in Louisiana, might eventually be destined for PADD 2. However, that crude oil is temporarily stored in PADD 3 before transfer to PADD 2, so is not originally reported as a PADD 2 import in the data. These results suggest that a future study that takes imports between different PADDs into account could be interesting.

Alternative models for the changes in overall U.S. (non-SPR) inventories are presented in Table 8. Models 1 and 3 evaluate the full sample period from 1992-2011 and Models 2 and 4 evaluate the 2004-2011 period. In Models 1 and 2 we present estimations using an alternative dependent variable. Instead of the change in barrels we use the percentage change in inventories. In Models 3 and 4 we add lagged control variables for crude oil refinery inputs, production and imports. Results are consistent with our main results that total U.S. (non-SPR) inventory changes are a positive function of the change in spread over the full sample period from 1992-2011 but not over the 2004-2011 timeframe. Again the results indicate that since 2004 most on-shore and above-ground C&C arbitrage apparently occurs at Cushing.

****** Place Table 8 about here ******

Models 1 and 2 in Table 8 where we use the percentage change in total U.S. (non-SPR) inventories as a dependent variable are very similar to Models 1 and 3 in Table 4 where the dependent variable is the barrel change. Otherwise the model specifications for models in Tables 8 and 4 are the same. Consistent with the previous results, the percentage change in total U.S. (non-SPR) inventories is a positive function of the spread in Model 1 from 1992-2011, but the spread is insignificant in Model 2 from 2004-2011. Also consistent with the Table 4 results, operational controls and seasonal variables remain significant in explaining U.S. (non-SPR) inventories in both time periods and in both Models 1 and 2 in Table 8.

The results from Models 3 and 4 in Table 8 in which we add lagged control variables for refinery inputs, production and imports are consistent with the results from Models 1 and 2 from Table 8, as well as models in Table 4. As before, the change in total U.S. (non-SPR) inventories is a positive function of the change in spread over the 1992-2011 period, but not over the 2004-2011 period.

On the other hand, all operating variables--current, lag, and most lead changes are significant in explaining total U.S. (non-SPR) inventory changes. We did not include lag control variable changes past one week back because we reasoned that longer dated lags would bring our models to an approximate tautology.[22]

In general, Tables 7 and 8 indicate that results dealing with the influence of the spread on Cushing and total U.S. crude oil inventories remain consistent with alternative specifications. Spread mainly explains Cushing inventory changes, while operational variables largely explain inventory changes in other

[22] With perfect data, the change in inventories equals imports plus production minus refinery inputs for that week. So, a regression of inventory changes on import, production, and refinery input levels would be estimating a tautology. Since lagged inventory changes going back many weeks proxy for the current inventory levels, models that include long-dated lags of imports, production and refinery inputs estimate an approximate tautology. An inclusion of just one lag increases the adjusted R^2 by 9% between Models 1 and 3 and by and 15% between Models 2 and 4 in Table 8.

areas. The influence of spread on inventories occurs with a lag. The implied coefficient pattern for the twelve lagged spreads calculated from the ΔPDL coefficients in the models again confirms that when the spread rises inventories in both Cushing and the overall U.S. increase most, not in the current week, but two to four weeks out. In other words spreads two to four weeks out have a stronger influence on the change in inventories than the current spreads.

Table 9 presents regressions for both Cushing and the U.S. using the twelve individual spread change lags instead of imposing the PDL structure on the lagged spreads and monthly dummy variables instead of using weekly dummy variables structured in a polynomial form. In Model 1, we estimate the relations for Cushing from 2004-2011, in Model 2, for the U.S. from 1992-2011 and, in Model 3, for the U.S. from 2004-2011. Again our results hold using parsimonious specifications. Cushing inventory changes are a positive function of the change in spread, as confirmed by the Wald test (p-value of 0.000) for the joint significance of all twelve spread lags. U.S inventory changes are a positive function of the change in spread over 1992-2011, but not over 2004-2011. As compared with the PDL specification, the coefficient pattern for the lagged spread changes jumps around. Also, the standard errors of the coefficients of the lagged spread terms (*not shown but available upon request*) for the non-PDL models are considerably higher, so the PDL specification does (as expected) provide more efficient estimations. Standard errors are also higher when monthly dummies, instead of the polynomial specification of weekly dummies, are used to adjust for seasonality. Therefore, models with polynomial specifications have several advantages—they are more efficient and they provide us with a smoother coefficient pattern for seasonality and for the influence of the change in spread on inventory changes.

****** Insert Table 9 about here ******

Tables 10 and 11 present regressions for both Cushing and the U.S. using alternative measures of the spread. In Model 1, we estimate the relations for Cushing from 2004-2011, in Model 2, for the U.S. from 1992-2011 and, in Model 3, for the U.S. from 2004-2011. The main measure of the spread used in the paper is the difference between the two- and one-month NYMEX WTI futures and the reasoning for this choice is described in section 4.1. In Table 11 we follow Equation (2) but regress crude oil inventory changes on the spread defined as the difference between the two-month futures and the spot WTI prices, while in Table 11 the spread is defined as the difference between the three- and one-month NYMEX WTI futures. Again our results hold using different measures of spread. Cushing inventory changes are a positive function of the different spread measures over 2004-2011. U.S. inventory changes are a positive function of different spread measures over 1992-2011, but not over 2004-2011. None of the operational variables, except for production changes, are significant in explaining Cushing inventory changes. On the other hand, all current and lead operational variables, except lead production, are significant in explaining U.S. inventory changes. The patterns of spread coefficients available in Panel B are similar to all our other models as they show that spreads influence crude inventories with a lag.

****** Place Table 10 and 11 about here ******

7. Conclusions

From regressions of crude oil inventory changes on current and lagged spreads, we find that crude oil inventories at Cushing are a strong positive function of current and lagged futures spreads. We find that current crude oil inventories are influenced not only by current spreads but by spreads over the last eight weeks. Indeed we find that current inventories are a stronger function of spreads several weeks ago than of current spreads. Our interpretation of this is that current spreads likely lead to contracts for forward delivery which do not result in a change in actual stock levels until delivery occurs sometime in the future. For instance, if in July the price of the September futures contract exceeds the price of the August contract by more than the cost of storage, an arbitrageur may long the August futures contract and short the September contract. He would then take delivery on the August contract and make delivery on the September contract so we would observe inventories rising in August and falling in September due to the July spread. Thus, the influence of the spread on storage is not immediate. We find no convincing evidence that the futures spread materially impacts inventories outside Cushing. Total U.S. inventories and PADD 2 inventories over the 1992-2004 period, i.e. before Cushing inventories were reported separately, are significant positive functions of the spread. However, when Cushing inventories are subtracted from

PADD 2 figures over the 2004-2011 period, the spread terms are insignificant. Similarly, the spread is insignificant in explaining total U.S. inventory changes over 2004-2011. The spread variables are also insignificant in regressions for all the other PADD districts.

It has been hypothesized that the futures spread impacts inventories held by refiners, pipelines, and other oil companies as well as arbitrageurs. In other words, if the spot price is far enough above the futures price, refiners and pipelines will draw down current inventories (thus risking a stock-out) and replenish later at lower prices, and they will build up inventories when the spot price is low relative to the futures price. The fact that we find little evidence that inventories outside Cushing are impacted by the futures spread indicates that if this activity occurs at all it is too small for us to detect with aggregate data. However, our inventory data includes above ground, on-shore inventories only. We cannot rule out the possibility that the futures spread may impact producers' decisions whether to pump the oil or leave it in the ground. Nor can we shed light on the impact of the spread on tanker storage in the Gulf or elsewhere.

We further find evidence that total U.S. and most PADD district inventories, except for Cushing, are partially explained by operational variables, particularly current and future changes in imports, U.S. oil production, and refinery inputs. These results indicate that refiners and storage operators increase inventories when they foresee future refinery needs exceeding future imports plus U.S. crude oil production, and reduce inventories when they foresee a surplus of imports and domestic production over refinery needs. We do not find these variables having much impact on Cushing inventories, further indicating that inventories there are mostly held for arbitrage (and possibly speculative) purposes.

To our knowledge, this study provides the first comprehensive analysis of the causal relationships between oil spot prices, futures prices and inventories, including the C&C arbitrage relation between oil inventories and the futures spread, and how inventories are impacted by contango versus normal backwardation in oil futures prices. Our findings establish, to our knowledge, the first tangible evidence documented in the literature of a causal link between oil futures and spot markets via inventory changes resulting from arbitrage. These findings provide an important foundation for future research on the impact of financial traders on the spot markets, especially the twin questions of whether financial traders (a) exacerbate or attenuate spot price volatility, and (b) whether they systematically affect the spot oil price level.

Appendix 1 -- Variable Description

The data comes from the *EIA WEEKLY PETROLEUM STATUS REPORTS* or is created based on variables from the report *http://www.eia.gov/oil_gas/petroleum/data_publications/weekly_petroleum_status_report/wpsr.html.*

Below are the descriptions of variable levels. In our analysis we mainly use changes in these variables. Δ(variable name) means a change in that variable over the last week calculated via first difference. (variable name)(+1) denotes a lead, or the next period value, and (variable name)(-1) denotes a lag, for that specific variable. Years from which the data is available and used in the analysis, as well as, units of measure are also presented. The data is weekly and ends on 7/8/20011.

Variable	**Description**	**From**	**Units**
ΔPDL_1-ΔPDL_4	Polynomial distributed lag (PDL) change in spread lags between the two- and one-month WTI crude futures for spread2_1. Robustness tests include the PDL change in spread lags between the two month future and spot WTI for spread2_spot (Table 10) and between three- and one-month WTI crude futures for spread 3_1 (Table 11).	9/11/1992	
Spread**X_Y**	The spread between the X and Y month out future NYMEX WTI crude contract (X,Y = 1,2,3,4). Typically, spread2_1, which is the spread between the two- and one-month WTI crude futures.	9/11/1992	Dollars per Barrel
ΔSpread2_1	The change in spread between the two- and the one-month WTI crude futures.	9/11/1992	Dollars per Barrel
Spot_WTI	Cushing, OK NYMEX WTI crude oil spot price FOB	9/11/1992	Dollars per Barrel
Future_WTI_**X**	Cushing, OK WTI NYMEX crude oil future contract X (X = 1,2,3,4)	9/11/1992	Dollars per Barrel
Stock_US	Weekly U.S. crude oil inventories (stocks) excluding the Strategic Petroleum Reserve (SPR)	9/11/1992	Thousand Barrels
Stock_X	Weekly crude oil inventories (stocks) excluding SPR for PADD_**X** (X=1, 2, 3, 4, or 5)	9/11/1992	Thousand Barrels
Stock_Cushing	Weekly Cushing, OK crude oil inventories (stocks) excluding SPR	4/9/2004	Thousand Barrels
Stock_US_non Cushing	Weekly U.S. crude oil inventories (stocks) excluding SPR and excluding Cushing	4/9/2004	Thousand Barrels
Stock_2_non Cushing	Weekly crude oil inventories (stocks) for PADD2 excluding SPR and excluding Cushing	4/9/2004	Thousand Barrels
Prod_US	Weekly U.S. field production of crude oil	9/11/1992	Thousand Barrels/Day
Imports_US	Weekly Net Inflows (Imports Excluding SPR - Exports)	9/11/1992	Thousand Barrels/Day
Imports**X**	Weekly crude oil imports excluding SPR for PADD_**X** (X=1, 2, 3, 4, or 5)	9/11/1992	Thousand Barrels/Day
RefinerInputUS	Weekly U.S. refiner net input of crude oil (balance between crude oil supply and disposition)	9/11/1992	Thousand Barrels/Day
RefinerInput**X**	Weekly refiner net input of crude oil for PADD_**X** (X=1, 2, 3, 4, or 5)	9/11/1992	Thousand Barrels/Day
Jan-Nov	Monthly dummies	9/11/1992	1 or 0
Z1-Z5	Weekly dummies (Z1 is 1, Z2 is level, Z3 is squared, Z4 is cubed, Z5 is to the forth power) that remove seasonality in inventory	9/11/1992	

PADD 1 -- East Coast; PADD 2 -- Midwest; PADD 3 -- Gulf Coast;

PADD 4 -- Rocky Mountains; PADD 5 -- West Coast.

Appendix 2 –Polynomial Distributed Lag (PDL) Model

An example of how we set up a fourth degree PDL model is shown below based on the distributed lag model using twelve lags of X, which in our main model is ΔSPREAD:

$$Y_t = \lambda + \beta_0 X_t + \beta_1 X_{t-1} + \cdot \cdot \cdot \cdot + \beta_{12} X_{t-12} \tag{A.1}$$

where β_i can be approximated by a fourth degree polynomial:

$$\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \alpha_3 i^3 + \alpha_4 i^4 \qquad \text{where } i = 1 \text{ to } 12$$

$$\begin{aligned}
\beta_0 &= \alpha_0 \\
\beta_1 &= \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 \\
\beta_2 &= \alpha_0 + \alpha_1 2 + \alpha_2 2^2 + \alpha_3 2^3 + \alpha_4 2^4 \\
\beta_3 &= \alpha_0 + \alpha_1 3 + \alpha_2 3^2 + \alpha_3 3^3 + \alpha_4 3^4 \\
\vdots \quad & \quad \vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
\beta_{12} &= \alpha_0 + \alpha_1 12 + \alpha_2 12^2 + \alpha_3 12^3 + \alpha_4 12^4
\end{aligned} \tag{A.2}$$

Substituting β_i from Equation (A.2) into the distributed lag Equation (A.1), and transforming it yields:

$$\begin{aligned}
Y_t = \lambda &+ \alpha_0 X_t + (\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) X_{t-1} + (\alpha_0 + \alpha_1 2 + \alpha_2 2^2 + \alpha_3 2^3 + \alpha_4 2^4) X_{t-2} \\
&+ (\alpha_0 + \alpha_1 3 + \alpha_2 3^2 + \alpha_3 3^3 + \alpha_4 3^4) X_{t-3} + \cdots + (\alpha_0 + \alpha_1 12 + \alpha_2 12^2 + \alpha_3 12^3 + \alpha_4 12^4) X_{t-12}
\end{aligned}$$

which can be rewritten as:

$$\begin{aligned}
Y_t = \lambda &+ \alpha_0 (X_t + X_{t-1} + X_{t-2} + \cdots + X_{t-12}) + \alpha_1 (X_{t-1} + 2X_{t-2} + 3X_{t-3} + \cdots + 12X_{t-12}) + \cdots \\
&+ \alpha_4 (X_{t-1} + 2^4 X_{t-2} + 3^4 X_{t-3} + \cdots + 12^4 X_{t-12})
\end{aligned}$$

or

$$Y_t = \lambda + \alpha_0 Z_{0t} + \alpha_1 Z_{1t} + \alpha_2 Z_{2t} + \alpha_3 Z_{3t} + \alpha_4 Z_{4t}$$

where the Z variables are constructed using 12 lags of X:

$$\begin{aligned}
Z_{0t} &= X_t + X_{t-1} + X_{t-3} + \cdots + X_{t-12} \\
Z_{1t} &= X_{t-1} + 2X_{t-2} + 3X_{t-3} + \cdots + 12X_{t-12} \\
Z_{2t} &= X_{t-1} + 2^2 X_{t-2} + 3^2 X_{t-3} + \cdots + 12^2 X_{t-12} \\
Z_{3t} &= X_{t-1} + 2^3 X_{t-2} + 3^3 X_{t-3} + \cdots + 12^3 X_{t-12} \\
Z_{4t} &= X_{t-1} + 2^4 X_{t-2} + 3^4 X_{t-3} + \cdots + 12^4 X_{t-12}
\end{aligned}$$

References

Acharya, V., Lochstoer, L., and Ramadorai, T., 2011, Limits to Arbitrage and Hedging: Evidence from Commodity Markets, National Bureau of Economic Research Working Paper 16875.

Einloth, J., 2009, Speculation and Recent Volatility in the Price of Oil, Working Paper, Division of Insurance and Research, Federal Deposit Insurance Corporation.

Etula, E., 2010, Broker-Dealer Risk Appetite and Commodity Returns, Federal Reserve Bank of New York Staff Report No. 406.

Hamilton J., 2009, Causes and Consequences of the Oil Shock of 2007-08, *Brookings Paper on Economic Activity*, Spring, 215-261.

Interagency Task Force on Commodity Markets (ITFCM), 2008, Interim Report on Crude Oil.

International Energy Agency (IEA), 2008, Medium Term Oil Market Report.

International Monetary Fund (IMF), 2008, World Economic Outlook: Housing and the Business Cycle.

Irwin, S., Sanders, D., Merrin, R., 2009, Devil or Angel: The Role of Speculation in the Recent Commodity Price Boom (and Bust), *Journal of Agricultural and Applied Economics* 41, 377–391.

Kaufmann, R. K., Ullman, B., 2009, Oil Prices, Speculation, and Fundamentals: Interpreting Causal Relations among Spot and Futures Prices, *Energy Economics* 31, 550–558.

Kilian, L., Murphy, D., 2010, The Role of Inventories and Speculative Trading in the Global Market for Crude Oil, Working Paper, Center for Economic Policy Research and University of Michigan at Ann Arbor.

Krugman, P., 2008, Speculation and Signatures, NYTimes.com, June 24.

Krugman, P. 2009, Oil speculation, NYTimes.com, July 8.

Masters, M. W., 2008, Testimony before the Committee on Homeland Security and Governmental Affairs, U.S. Senate, Washington DC, May 20.

Organization for Economic Co-operation and Development (OECD) Working Party on Agricultural Policies and Markets, 2010, Speculation and Financial Fund Activity: Draft Report Annex I.

Parsons, J. E., 2010, Black Gold & Fool's Gold: Speculation in the Oil Futures Market, *Economia* 10, 81-116.

Phillips, P., Yu, J., 2010, Dating the Timeline of Financial Bubbles, Working Paper, Yale University and Singapore Management University.

Shleifer, A., and Vishny, R., 1997, The Limits of Arbitrage, *Journal of Finance* 52, 35-55.

Singleton, K. 2011, The 2008 Boom/Bust in Oil Prices, Working Paper, Stanford University.

Smith J., 2009, World Oil: Market or Mayhem? *Journal of Economic Perspectives* 23, 245-164.

Sornette, D., Woodard, R., and Zhou, W., 2009, The 2006-2008 Oil Bubble: Evidence of Speculation, and Prediction, *Physica A* 388, 1571–1576.

U.S. Senate Permanent Subcommittee on Investigations, 2006, The Role of Market Speculation in Rising Oil and Gas Prices: A Need to Put the Cop Back on the Beat, Staff Report, Committee on Homeland Security and Governmental Affairs, United States Senate.

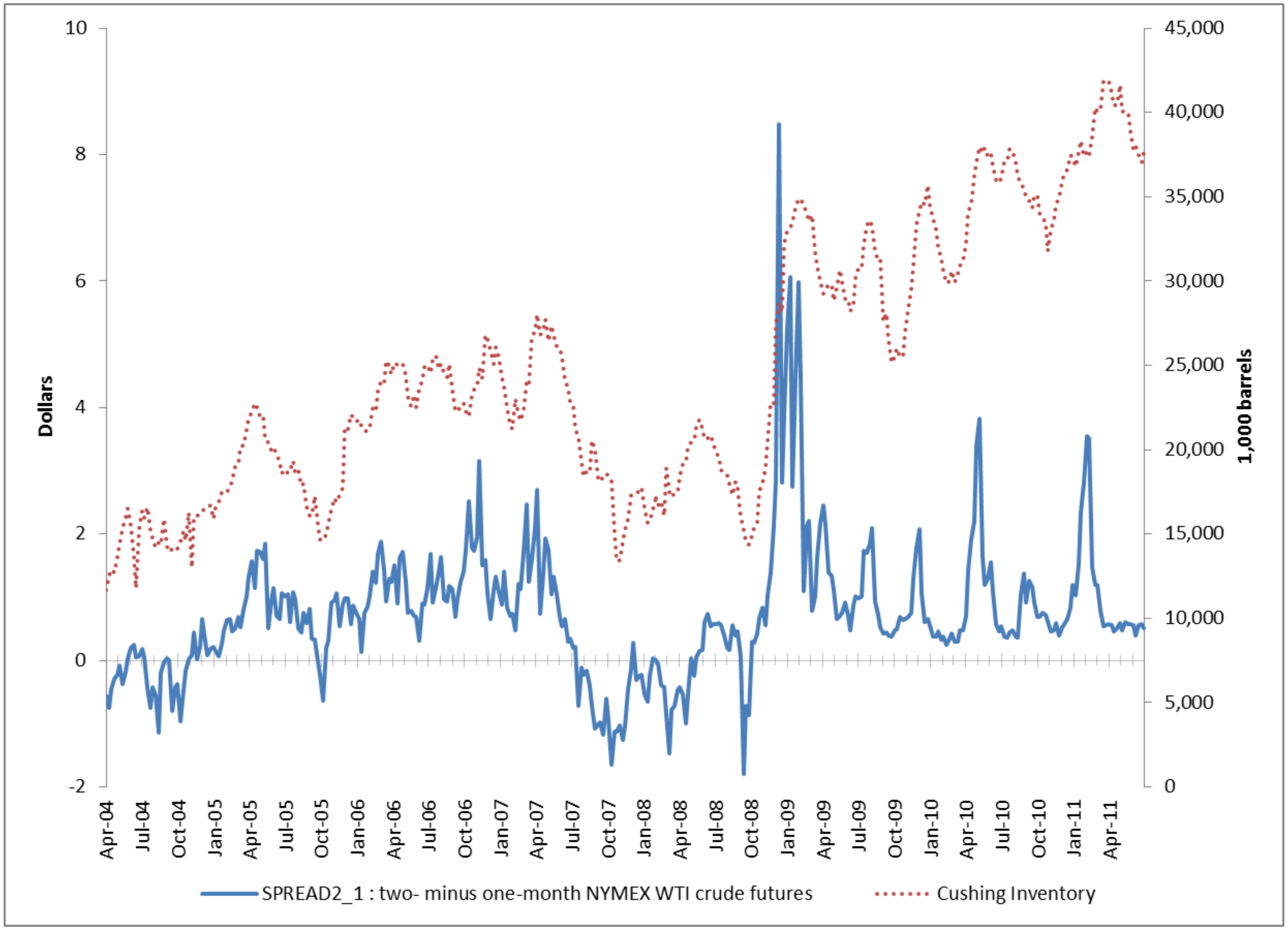

Figure 1—Cushing Crude Inventory and the Spread Between the Two- and One-Month NYMEX WTI Crude Futures.
This figure plots crude oil inventories in Cushing, OK and the spread between the two- and one-month WTI crude futures between April 9, 2004 and July 8, 2011.

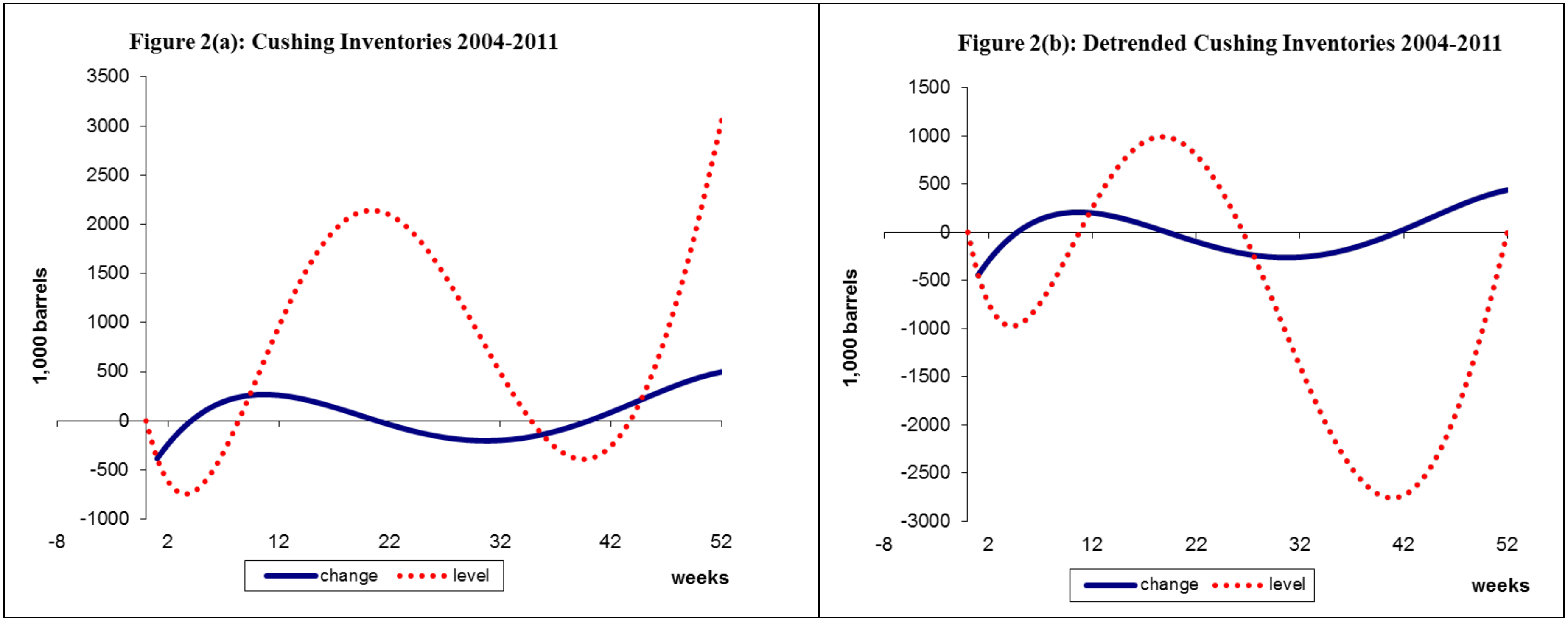

Figure 2 – Cushing Crude Oil Inventory Levels and Changes as Predicted by the Seasonal Adjustment from the Z Variables
This figure plots seasonal pattern at Cushing, OK crude oil inventories as implied by the weekly seasonal Z variables. Variable definitions are in Appendix 1. Both inventory levels and changes are presented over the 2004-2011 period. Figure 2(a) shows unadjusted inventory levels in Cushing. There has been a significant increase in capacity in Cushing, OK over 2004-2011as shown by the difference in the starting and ending points of the inventory “level” series. Figure 2(b) adjusts for this capacity increase by normalizing the graph scale in Figure 2(a) to start and end at the same level, and shows the seasonal pattern on this adjusted scale.

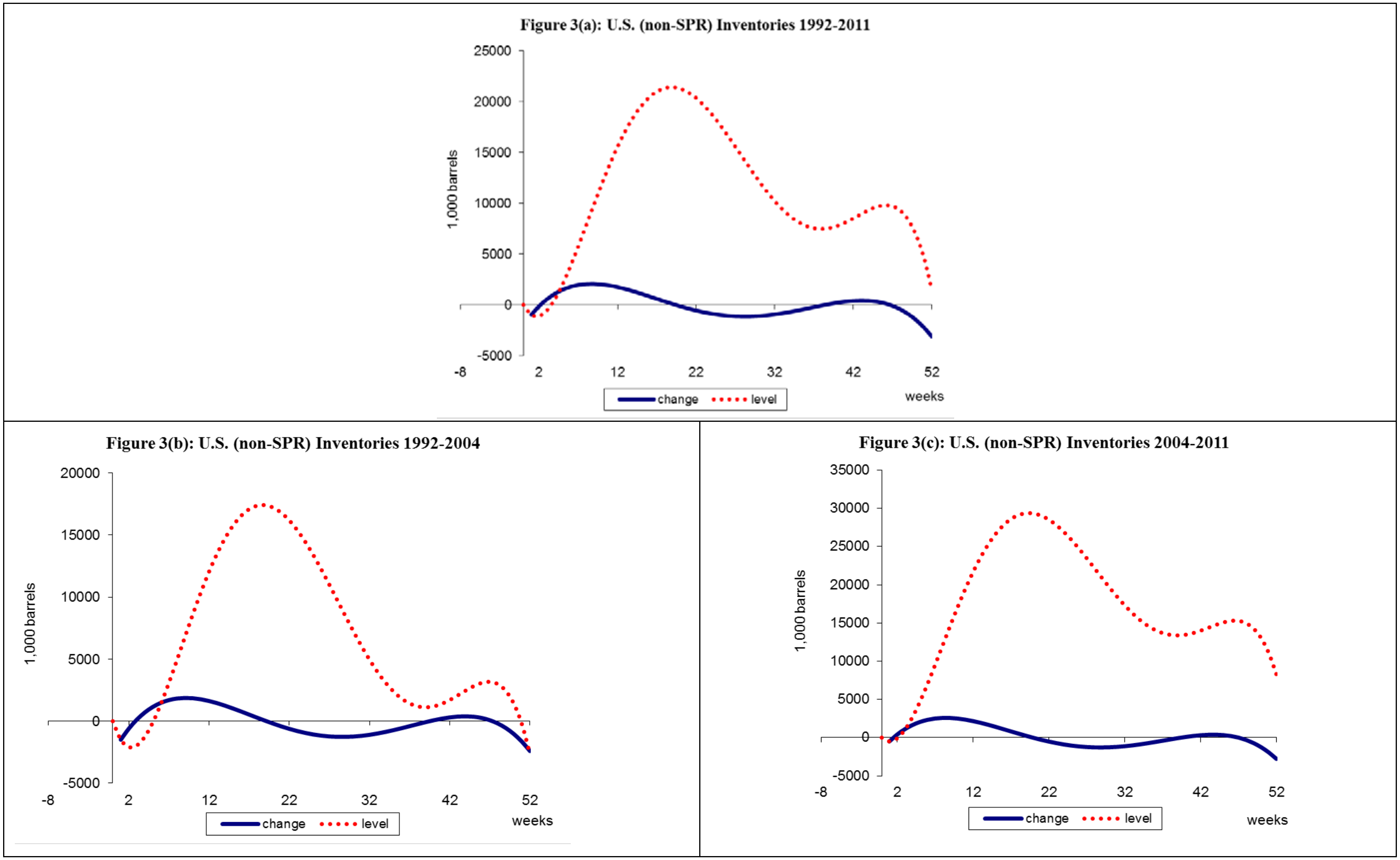

Figure 3 – U.S. (non-SPR) Crude Oil Inventory Levels and Changes as Predicted by the Seasonal Adjustment from the Z Variables
This figure plots the U.S. (non-SPR) seasonal pattern in inventories as implied by the weekly seasonal Z variables. Variable definitions are available in Appendix 1. Both inventory levels and changes are presented. Figure 3(a) covers the 1992-2011 period; Figure 3(b) the 1992-2004 period; and Figure 3(c) the 2004-2011 period. Figure 3(b) shows that overall U.S. crude oil storage capacity did not increase significantly over 1992-2004, as the line graphing inventory levels converges to its starting point. This differs from Figure 3(c) which shows that capacity increased over the 2004-2011 period. This capacity increase is partially attributed to Cushing, as can be seen from Figure 2.

Table 1-- Description of Variables in Levels and Changes (1st Difference)
The number, mean, median and standard deviation (and autocorrelation for changes) for levels and weekly changes (1st difference) of the variables used in the analysis are reported. Years for which the data is available and used in the analysis are also presented. Variable definitions are in Appendix 1.

			LEVELS			**CHANGES (1st Difference)**				
Variables	N	Years	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	AC 1st order	AC p-value
STOCK_US	983	1992-2011	319,097.20	320,634.00	23,355.21	27.56	132.50	3,912.40	0.045	0.161
STOCK_CUSHING	379	2004-2011	24,741.52	23,157.00	7,866.80	68.71	56.00	979.94	0.100	0.051
STOCK_US_nonCUSHING	379	2004-2011	304,012.90	304,584.00	17,244.58	90.24	317.00	3,365.87	0.351	0.000
STOCK2	983	1992-2011	70,007.28	68,376.00	10,320.64	33.49	8.50	1,287.79	0.085	0.007
STOCK2_nonCUSHING	379	2004-2011	49,980.12	48,074.00	5,370.63	36.87	68.00	1,102.75	-0.054	0.293
STOCK1	983	1992-2011	14,650.79	14,764.00	1,523.02	-3.19	40.50	1,115.03	-0.362	0.000
STOCK3	983	1992-2011	162,503.00	162,172.00	13,808.29	9.98	29.50	3,198.50	0.008	0.791
STOCK4	983	1992-2011	12,989.03	12,680.00	1,609.29	3.73	-1.50	311.82	-0.085	0.007
STOCK5	983	1992-2011	58,947.08	56,518.00	7,943.24	-16.46	-1.00	2,030.22	-0.281	0.000
SPREAD2_SPOTWTI	983	1992-2011	0.19	0.11	1.08	0.00	0.00	0.67	-0.344	0.000
SPREAD2_1	983	1992-2011	0.16	0.10	0.92	0.00	0.01	0.47	-0.261	0.000
SPOT_WTI	983	1992-2011	41.42	28.86	27.92	0.08	0.10	2.57	-0.022	0.487
PROD_US	983	1992-2011	5,850.92	5,808.00	615.73	-1.40	1.00	107.50	-0.201	0.000
REFINERINPUTUS	983	1992-2011	14,708.67	14,793.00	752.25	1.38	15.00	282.95	-0.001	0.978
REFINERINPUT1	983	1992-2011	1,446.55	1,477.00	181.78	-0.06	-1.00	72.65	-0.101	0.002
REFINERINPUT2	983	1992-2011	3,253.51	3,260.00	163.98	-0.02	2.00	94.39	-0.064	0.046
REFINERINPUT3	983	1992-2011	6,976.75	7,076.00	518.09	1.45	9.00	231.14	-0.053	0.097
REFINERINPUT4	983	1992-2011	508.52	511.00	48.41	0.08	0.00	22.53	-0.151	0.000
REFINERINPUT5	983	1992-2011	2,523.34	2,538.00	148.75	-0.08	2.00	85.57	-0.160	0.000
IMPORTS_US	983	1992-2011	8,685.16	8,876.00	1,307.88	2.79	-20.00	779.04	-0.534	0.000
IMPORTS1	983	1992-2011	1,421.93	1,415.00	272.57	-0.20	-6.00	346.94	-0.568	0.000
IMPORTS2	983	1992-2011	975.69	979.00	208.23	0.54	4.00	160.72	-0.501	0.000
IMPORTS3	983	1992-2011	5,400.42	5,502.00	802.97	0.92	-21.00	708.78	-0.488	0.000
IMPORTS4	983	1992-2011	203.83	205.00	87.65	0.19	0.00	51.94	-0.573	0.000
IMPORTS5	983	1992-2011	747.10	742.00	370.68	1.27	0.00	253.44	-0.588	0.000

Table 2--Correlations between variables

Correlations between weekly variable changes (1st difference) are reported. Variable definitions are in Appendix 1. Data involving Cushing covers the 2004-2011 period, data for all other variables is for 1992-2011

	Δ(STOCK_US)	Δ(STOCK_CUSHING)	Δ(STOCK_US_nonCUSHING)	Δ(STOCK2)	Δ(STOCK2_nonCUSHING)	Δ(STOCK1)	Δ(STOCK3)	Δ(STOCK4)	Δ(STOCK5)	Δ(SPREAD 2_1)	Δ(SPOT_WTI)	Δ(PROD_US)	Δ(REFINERINPUTUS)	Δ(REFINERINPUT1)	Δ(REFINERINPUT2)	Δ(REFINERINPUT3)	Δ(REFINERINPUT4)	Δ(REFINERINPUT5)	Δ(IMPORTS_NET_NOSPR)	Δ(IMPORTS1)	Δ(IMPORTS2)	Δ(IMPORTS3)	Δ(IMPORTS4)	Δ(IMPORTS5)
Δ(STOCK_US)	**1**																							
Δ(STOCK_CUSHING)	.17	**1**																						
Δ(STOCK_US_nonCUSHING)	.96	-.12	**1**																					
Δ(STOCK2)	.28	.59	.12	**1**																				
Δ(STOCK2_nonCUSHING)	.20	-.17	.25	.70	**1**																			
Δ(STOCK1)	.17	-.03	.18	.04	.07	**1**																		
Δ(STOCK3)	.77	-.04	.79	-.14	-.13	-.12	**1**																	
Δ(STOCK4)	.09	-.13	.12	-.06	.04	.03	.05	**1**																
Δ(STOCK5)	.32	-.03	.33	.01	.03	-.13	-.08	-.07	**1**															
Δ(SPREAD 2_1)	.04	.16	-.01	.13	.02	.01	-.02	.02	-.01	**1**														
Δ(SPOT_WTI)	-.02	.04	-.03	.02	-.01	.06	.01	.00	-.11	-.30	**1**													
Δ(PROD_US)	.17	.09	.14	.15	.09	-.01	.03	.07	.18	.05	-.07	**1**												
Δ(REFINERINPUTUS)	.08	.09	.05	.09	.02	-.01	.01	-.01	.09	.09	-.02	.45	**1**											
Δ(REFINERINPUT1)	.02	.09	-.01	.03	-.05	-.03	.03	-.05	-.02	-.02	-.05	.00	.16	**1**										
Δ(REFINERINPUT2)	-.05	.03	-.06	-.10	-.15	.05	-.04	-.02	.03	.05	-.04	.11	.36	.09	**1**									
Δ(REFINERINPUT3)	.12	.11	.09	.14	.07	.02	.01	.01	.12	.11	-.01	.46	.89	-.10	.06	**1**								
Δ(REFINERINPUT4)	.00	-.14	.05	-.10	.01	-.10	.07	-.19	.07	-.01	.01	-.01	.13	.14	.09	.01	**1**							
Δ(REFINERINPUT5)	-.08	-.09	-.06	-.04	.03	-.09	-.01	.04	-.09	-.07	.05	.02	.24	-.07	.04	-.02	.01	**1**						
Δ(IMPORTS_NET_NOSPR)	.52	.12	.48	.13	.05	.18	.41	-.07	.11	.01	.03	.09	.31	.07	.10	.29	.06	.02	**1**					
Δ(IMPORTS1)	.09	.05	.07	.03	-.01	.35	-.05	-.05	.04	.02	.06	-.09	-.05	.11	.10	-.09	-.02	-.07	.20	**1**				
Δ(IMPORTS2)	.08	.09	.06	.15	.10	-.06	.03	-.03	.04	-.01	.03	-.10	.03	-.07	.13	.01	.05	-.02	.19	-.05	**1**			
Δ(IMPORTS3)	.42	.10	.39	.09	.02	.04	.45	-.01	-.06	-.02	.05	.14	.29	.01	.02	.30	.06	.04	.76	-.23	-.04	**1**		
Δ(IMPORTS4)	.06	.05	.05	.02	-.02	.09	.02	.06	.00	.03	.04	.03	.03	-.02	-.02	.03	.10	.03	.04	.04	-.03	.00	**1**	
Δ(IMPORTS5)	.09	-.05	.11	-.03	.01	-.06	.00	-.08	.31	.04	-.11	.03	.10	.05	.00	.09	-.03	.03	.24	-.13	-.01	-.15	-.15	**1**

Table 3--Polynomial distributed lag (PDL) estimation of impact of the futures spread changes on Cushing crude inventory changes

The dependent variable, ΔSTOCK, is the weekly change Cushing crude inventories. Variable definitions are in Appendix 1. Any (+1) variables indicate a lead for that specific variable. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. Data is weekly. Model 1 estimates over the 4/16/2004 - 7/08/2011 period, has 377 observations.

Panel A.	**Model 1 Δ Cushing stock 2004-2011**		**Model 1 winsorized Δ Cushing stock 2004-2011**	
	Coeff.	p-value	Coeff.	p-value
C	-556.14	0.022	-588.92	0.008
Δ(SPOT_WTI)	23.98	0.159	20.11	0.183
Δ(REFINERINPUT2)	-0.22	0.703	-0.08	0.883
Δ(REFINERINPUT2)(+1)	-0.25	0.673	-0.21	0.715
Δ(PROD_US)	0.55	0.039	0.59	0.021
Δ(PROD_US)(+1)	-0.28	0.270	-0.26	0.307
Δ(IMPORTS2)	0.43	0.212	0.37	0.266
Δ(IMPORTS2)(+1)	-0.29	0.433	-0.30	0.387
Z2	185.98	0.002	190.82	0.001
Z3	-13.45	0.003	-13.80	0.002
Z4	0.33	0.011	0.34	0.008
Z5	0.00	0.036	0.00	0.025
ΔPDL_1	167.89	0.000	162.41	0.001
ΔPDL_2	-59.45	0.001	-63.13	0.001
ΔPDL_3	-0.12	0.952	0.30	0.878
ΔPDL_4	1.10	0.070	1.33	0.050
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.000**		**0.000**
Adjusted R-squared	15%		16%	

Panel B.	**Δ SPREAD2_1**		**Δ SPREAD2_1**	
Lags	Coeff.	p-value	Coeff.	p-value
0	282.03	0.000	265.64	0.000
1	324.29	0.000	319.85	0.000
2	333.20	0.000	334.89	0.000
3	315.40	0.000	318.71	0.000
4	277.50	0.000	279.27	0.000
5	226.12	0.000	224.52	0.000
6	167.89	0.000	162.42	0.000
7	109.42	0.023	100.91	0.028
8	57.33	0.246	47.95	0.307
9	18.26	0.698	11.49	0.799
10	-1.20	0.977	-0.50	0.990
11	5.60	0.900	19.92	0.692
12	45.26	0.574	80.70	0.386
Sum of Lags	2161.08	**0.000**	2165.76	**0.000**

Table 4--Polynomial distributed lag (PDL) estimation of the impact of the futures spread changes on the U.S. non-SPR crude inventory changes

The dependent variable, ΔSTOCK, is the weekly change in U.S. non-SPR crude inventories. Variable definitions are in Appendix 1. Lagged autoregressive error terms are included when needed. Any (+1) variables indicate a lead for that specific variable. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. Data is weekly. Model 1 estimates over 12/11/1992 - 7/8/2011, has 969 observations and is presented with original and 1% winsorized data; Model 2, over 12/11/1992 - 4/9/2004, has 591 observations; Model 3, over 4/23/2004 - 7/8/2011, has 376 observations; Model 4, excludes Cushing from U.S. inventories for 4/23/2004 to 7/8/2011, has 377 observations.

Panel A.	**Model 1 Δ U.S. stock 1992-2011**		**Model 1 winsorized Δ U.S. stock 1992-2011**	
	Coeff.	p-value	Coeff.	p-value
C	-1984.29	0.000	-1847.15	0.000
Δ(SPOT_WTI)	-29.23	0.420	-26.32	0.470
Δ(REFINERINPUTUS)	-1.03	0.017	-0.94	0.021
Δ(REFINERINPUTUS)(+1)	1.88	0.000	1.89	0.000
Δ(PROD_US)	4.21	0.000	4.18	0.000
Δ(PROD_US)(+1)	0.61	0.521	0.65	0.494
Δ(IMPORTS_US)	1.37	0.000	1.33	0.000
Δ(IMPORTS_US)(+1)	-0.90	0.000	-0.88	0.000
Z2	1111.54	0.000	1078.12	0.000
Z3	-96.63	0.000	-94.53	0.000
Z4	2.81	0.000	2.76	0.000
Z5	-0.03	0.000	-0.03	0.000
ΔPDL_1	454.26	0.021	444.08	0.023
ΔPDL_2	-68.21	0.152	-63.73	0.182
ΔPDL_3	-7.48	0.274	-6.96	0.304
ΔPDL_4	1.38	0.406	1.27	0.448
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.017**		**0.018**
Adjusted R-squared	28%		28%	

Panel B.	**Δ SPREAD2_1**		**Δ SPREAD2_1**	
Lags	Coeff.	p-value	Coeff.	p-value
0	295.49	0.157	301.47	0.147
1	435.43	0.002	429.92	0.002
2	518.91	0.001	506.33	0.001
3	554.24	0.002	538.32	0.002
4	549.70	0.004	533.53	0.005
5	513.61	0.009	499.58	0.010
6	454.26	0.021	444.08	0.023
7	379.96	0.054	374.67	0.057
8	299.00	0.123	298.96	0.124
9	219.70	0.224	224.59	0.217
10	150.34	0.327	159.18	0.302
11	99.24	0.438	110.35	0.386
12	74.69	0.670	85.73	0.621
Sum of Lags	4544.57	**0.005**	4506.70	**0.005**

Table 4 (*continued*)--Polynomial distributed lag (PDL) estimation of the impact of the futures spread changes on the U.S. non-SPR crude inventory changes

Panel A.	Model 2		Model 3		Model 4	
	Δ U.S. stock		Δ U.S. stock		Δ U.S. non-Cushing	
	1992-2004		2004-2011		2004-2011	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
C	-2558.21	0.000	-1557.91	0.066	-1330.06	0.067
Δ(SPOT_WTI)	-64.10	0.568	0.46	0.983	-34.02	0.262
Δ(REFINERINPUTUS)	-1.70	0.033	-1.30	0.000	-1.30	0.000
Δ(REFINERINPUTUS)(+1)	1.77	0.017	2.23	0.000	2.21	0.000
Δ(PROD_US)	4.02	0.092	3.02	0.000	3.00	0.000
Δ(PROD_US)(+1)	-0.22	0.909	-0.32	0.669	0.20	0.788
Δ(IMPORTS_US)	1.16	0.000	1.88	0.000	1.81	0.000
Δ(IMPORTS_US)(+1)	-0.71	0.004	-1.72	0.000	-1.54	0.000
Z2	1164.76	0.000	1160.14	0.000	1025.31	0.000
Z3	-97.58	0.000	-102.29	0.000	-90.76	0.000
Z4	2.77	0.000	2.95	0.000	2.64	0.000
Z5	-0.03	0.000	-0.03	0.000	-0.02	0.000
ΔPDL_1	1147.91	0.001	142.75	0.547	60.72	0.803
ΔPDL_2	-85.05	0.398	-54.83	0.324	-12.94	0.823
ΔPDL_3	-20.09	0.112	-5.31	0.393	-6.60	0.326
ΔPDL_4	-0.28	0.937	1.77	0.211	0.82	0.625
AR(1)	-0.15184	0.000	0.55464	0.000	0.47128	0.000
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.001**		**0.584**		**0.3721**
Adjusted R-squared	21%		69%		61%	

Panel B.	Δ SPREAD2_1		Δ SPREAD2_1		Δ SPREAD2_1	
Lags	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
0	995.87	0.046	-102.14	0.410	-277.19	0.061
1	1106.14	0.001	62.69	0.687	-142.57	0.364
2	1184.69	0.000	163.73	0.420	-45.85	0.828
3	1229.84	0.000	211.62	0.359	17.89	0.942
4	1239.89	0.000	217.00	0.367	53.61	0.834
5	1213.15	0.000	190.49	0.428	66.24	0.793
6	1147.91	0.001	142.75	0.547	60.72	0.802
7	1042.49	0.002	84.39	0.719	42.00	0.858
8	895.18	0.008	26.06	0.911	15.02	0.948
9	704.30	0.036	-21.62	0.923	-15.28	0.945
10	468.14	0.137	-48.00	0.815	-43.96	0.829
11	185.01	0.555	-42.44	0.804	-66.08	0.720
12	-146.78	0.739	5.67	0.969	-76.70	0.700
Sum of Lags	11265.80	**0.000**	890.20	**0.667**	-412.14	**0.848**

Table 5--Polynomial distributed lag (PDL) estimation of the impact of the futures spread changes on PADD2 crude inventory changes

The dependent variable, ΔSTOCK, is the weekly change in PADD2 crude inventories. Variable definitions are in Appendix 1. Any (+1) variables indicate a lead for that specific variable. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. Data is weekly. Model 1 estimates over 12/11/1992 - 7/8/2011, has 968 observations and is presented with original and 1% winsorized data; Model 2, over 12/11/1992 - 4/9/2004, has 591 observations; Model 3, over 4/16/2004 - 7/8/2011, has 377 observations; Model 4, excludes Cushing from PADD2 inventories from 4/16/2004 to 7/8/2011, has 377 observations.

Panel A.	**Model 1**		**Model 1 winsorized**	
	Δ PADD2 stock		**Δ PADD2 stock**	
	1992-2011		**1992-2011**	
	Coeff.	p-value	Coeff.	p-value
C	-1020.25	0.000	-1028.80	0.000
Δ(SPOT_WTI)	16.82	0.219	14.43	0.276
Δ(REFINERINPUT2)	-1.88	0.000	-1.81	0.000
Δ(REFINERINPUT2)(+1)	0.38	0.360	0.35	0.388
Δ(PROD_US)	0.92	0.021	0.90	0.021
Δ(PROD_US)(+1)	0.58	0.095	0.56	0.103
Δ(IMPORTS2)	0.87	0.002	0.85	0.002
Δ(IMPORTS2)(+1)	-0.51	0.099	-0.51	0.087
Z2	344.50	0.000	341.92	0.000
Z3	-25.92	0.000	-25.62	0.000
Z4	0.68	0.000	0.67	0.000
Z5	-0.01	0.000	-0.01	0.000
ΔPDL_1	289.48	0.000	278.10	0.000
ΔPDL_2	-51.74	0.012	-48.89	0.012
ΔPDL_3	-2.59	0.228	-2.53	0.211
ΔPDL_4	0.72	0.289	0.70	0.269
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.000**		**0.000**
Adjusted R-squared	13%		13%	

Panel B.	**Δ SPREAD2_1**		**Δ SPREAD2_1**	
Lags	Coeff.	p-value	Coeff.	p-value
0	351.06	0.000	330.11	0.000
1	393.38	0.000	372.35	0.000
2	408.90	0.000	388.67	0.000
3	401.94	0.000	383.23	0.000
4	376.84	0.000	360.20	0.000
5	337.91	0.000	323.77	0.000
6	289.48	0.000	278.10	0.000
7	235.88	0.000	227.38	0.000
8	181.42	0.002	175.76	0.002
9	130.44	0.029	127.44	0.028
10	87.25	0.128	86.57	0.120
11	56.18	0.316	57.35	0.298
12	41.55	0.589	43.93	0.564
Sum of Lags	3292.21	**0.000**	3154.85	**0.000**

Table 5 (*continued*)--Polynomial distributed lag (PDL) estimation of the impact of the futures spread changes on PADD2 crude inventory changes

Panel A.	**Model 2**		**Model 3**		**Model 4**	
	Δ PADD2 stock		**Δ PADD2 stock**		**Δ PADD2 non-Cushing**	
	1992-2004		**2004-2011**		**2004-2011**	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
C	-850.00	0.000	-1073.35	0.002	-517.21	0.072
Δ(SPOT_WTI)	-15.93	0.735	19.03	0.212	-4.95	0.726
Δ(REFINERINPUT2)	-1.80	0.000	-2.23	0.001	-2.01	0.001
Δ(REFINERINPUT2)(+1)	0.64	0.225	-0.11	0.868	0.14	0.799
Δ(PROD_US)	0.06	0.897	1.49	0.001	0.95	0.015
Δ(PROD_US)(+1)	0.94	0.156	0.17	0.610	0.45	0.125
Δ(IMPORTS2)	0.71	0.047	1.23	0.004	0.80	0.026
Δ(IMPORTS2)(+1)	-0.47	0.238	-0.55	0.221	-0.26	0.499
Z2	309.33	0.000	383.50	0.000	197.52	0.008
Z3	-24.00	0.000	-28.71	0.000	-15.26	0.008
Z4	0.64	0.000	0.74	0.001	0.41	0.016
Z5	-0.01	0.000	-0.01	0.003	0.00	0.030
ΔPDL_1	602.68	0.000	205.46	0.002	37.57	0.497
ΔPDL_2	-109.90	0.000	-31.27	0.194	28.18	0.154
ΔPDL_3	-5.82	0.223	-2.14	0.429	-2.02	0.426
ΔPDL_4	2.89	0.008	0.03	0.970	-1.07	0.116
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.000**		**0.001**		**0.3416**
Adjusted R-squared	11%		17%		5%	

Panel B	**Δ SPREAD2_1**		**Δ SPREAD2_1**		**Δ SPREAD2_1**	
Lags	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
0	428.66	0.023	309.69	0.005	27.66	0.784
1	645.63	0.000	304.65	0.000	-19.63	0.780
2	764.31	0.000	294.44	0.000	-38.76	0.558
3	802.02	0.000	279.24	0.000	-36.16	0.595
4	776.10	0.000	259.22	0.000	-18.28	0.782
5	703.87	0.000	234.57	0.001	8.45	0.889
6	602.68	0.000	205.46	0.002	37.57	0.497
7	489.84	0.000	172.08	0.013	62.67	0.240
8	382.70	0.001	134.61	0.064	77.28	0.160
9	298.58	0.007	93.23	0.205	74.97	0.190
10	254.81	0.017	48.11	0.484	49.31	0.395
11	268.73	0.020	-0.56	0.993	-6.15	0.922
12	357.67	0.027	-52.60	0.504	-97.85	0.261
Sum of Lags	6775.59	**0.000**	2282.15	**0.000**	121.08	**0.837**

Table 6--Polynomial distributed lag (PDL) estimation of the impact of the futures spread changes on PADD 1,3,4 and 5 crude inventory changes

The dependent variable, ΔSTOCK, is the weekly change PADD 1,3,4 and 5 crude inventories. Variable definitions are in Appendix 1. Lagged autoregressive and moving average error terms are included when needed. Any (+1) variables indicate a lead for that specific variable. X denotes 1,3,4, and 5 for the PADD 1,3,4 and 5 equations respectively. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. The weekly data is from 12/11/1992 to 7/08/2011 and has 968 observations.

Panel A.	**Δ PADD1 stock**		**Δ PADD3 stock**		**Δ PADD4 stock**		**Δ PADD5 stock**	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
C	72.09	0.374	-650.44	0.289	-114.41	0.013	-520.45	0.064
Δ(SPOT_WTI)	1.48	0.803	-13.22	0.702	-0.05	0.989	-26.68	0.156
Δ(REFINERINPUT_X)	-0.71	0.109	-0.78	0.070	-2.00	0.000	-2.42	0.001
Δ(REFINERINPUT_X)(+1)	1.59	0.000	1.70	0.001	1.80	0.000	0.59	0.391
Δ(PROD_US)	0.18	0.547	0.95	0.384	0.10	0.222	1.71	0.003
Δ(PROD_US)(+1)	0.13	0.588	1.16	0.212	0.12	0.143	-1.05	0.102
Δ(IMPORTS_X)	0.60	0.000	1.22	0.000	0.20	0.413	1.47	0.000
Δ(IMPORTS_X)(+1)	-0.62	0.000	-0.91	0.000	-0.16	0.456	-0.67	0.015
Z2	7.19	0.723	506.15	0.000	52.65	0.000	253.93	0.000
Z3	-0.66	0.658	-45.68	0.000	-4.43	0.000	-22.61	0.000
Z4	0.01	0.740	1.37	0.000	0.13	0.000	0.66	0.000
Z5	0.00	0.793	-0.01	0.000	0.00	0.000	-0.01	0.000
ΔPDL_1	-4.30	0.830	152.87	0.352	-13.05	0.450	60.97	0.251
ΔPDL_2	16.59	0.154	0.75	0.986	-3.62	0.492	-25.18	0.245
ΔPDL_3	0.38	0.776	-3.93	0.512	0.65	0.347	-3.26	0.182
ΔPDL_4	-0.76	0.108	0.08	0.958	0.13	0.478	0.81	0.288
AR(1)	0.33	0.000			-0.12	0.002	-0.30	0.000
AR(2)	0.07	0.036						
MA(1)	-0.921	0.000						
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.548**		**0.926**		**0.784**		**0.539**
Adjusted R-squared	34%		23%		7%		15%	

Table 7 --Alternative specification models of the impact of the futures spread changes on Cushing crude inventory changes. Alternatives include--using % Cushing inventory change as a dependent variable, adding lagged and cross-PADD controls.

The dependent variable, ΔSTOCK, is the weekly change in Cushing inventories. In Model 1, it is the percentage change * 100, while in Models 2 and 3 it is the barrel change in inventories. Variable definitions are in Appendix 1. Any (+ 1) variable indicates a lead for that specific variable, while (-1) is a lag. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. Data is weekly. Models 1-3 estimate over 2004-2011, have 377 observations. Model 1 uses percentage instead of barrel ΔSTOCK, Model 2 adds lagged variables, Model 3 adds cross-PADD controls.

Panel A.	**Model 1 % Δ Cushing stock 2004-2011**		**Model2-with lags Δ Cushing stock 2004-2011**		**Model 3-with PADD3 Δ Cushing stock 2004-2011**	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
C	-270.08	0.005	-515.15	0.036	-532.50	0.026
Δ(SPOT_WTI)	12.36	0.095	22.91	0.168	21.55	0.200
Δ(REFINERINPUT_2)(-1)			1.36	0.006		
Δ(REFINERINPUT_2)	-0.15	0.618	-0.03	0.961	-0.26	0.651
Δ(REFINERINPUT_2)(+1)	-0.17	0.577	-0.03	0.954	-0.17	0.774
Δ(PROD_US)(-1)			0.53	0.064		
Δ(PROD_US)	0.33	0.014	0.68	0.006	0.37	0.128
Δ(PROD_US)(+1)	-0.13	0.363	-0.20	0.385	-0.40	0.118
Δ(IMPORTS_2)(-1)			-0.34	0.393		
Δ(IMPORTS_2)	0.25	0.150	0.21	0.573	0.41	0.223
Δ(IMPORTS_2)(+1)	-0.09	0.556	-0.45	0.202	-0.43	0.225
Δ(STOCKS_3)					-0.03	0.121
Δ(IMPORTS_3)					0.26	0.001
Z2	91.97	0.000	178.13	0.004	194.89	0.001
Z3	-6.64	0.001	-13.24	0.005	-14.59	0.002
Z4	0.17	0.006	0.34	0.011	0.37	0.006
Z5	0.00	0.023	0.00	0.029	0.00	0.019
ΔPDL_1	62.55	0.001	164.07	0.001	173.22	0.000
ΔPDL_2	-23.22	0.001	-59.93	0.000	-59.11	0.000
ΔPDL_3	0.11	0.906	0.06	0.977	-0.18	0.924
ΔPDL_4	0.37	0.117	1.14	0.060	1.06	0.077
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.000**		**0.000**		**0.000**
Adjusted R-squared	12%		17%		17%	

Panel B.	**Δ SPREAD2_1**		**Δ SPREAD2_1**		**Δ SPREAD2_1**	
Lags	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
0	125.64	0.000	279.69	0.000	292.69	0.000
1	134.97	0.000	322.77	0.000	331.91	0.000
2	133.40	0.000	331.80	0.000	339.01	0.000
3	123.14	0.000	313.61	0.000	320.34	0.000
4	106.44	0.000	275.04	0.000	282.25	0.000
5	85.50	0.000	222.92	0.000	231.09	0.000
6	62.55	0.001	164.07	0.000	173.22	0.000
7	39.81	0.040	105.34	0.033	114.98	0.015
8	19.50	0.316	53.56	0.298	62.73	0.197
9	3.85	0.833	15.54	0.752	22.81	0.621
10	-4.92	0.760	-1.86	0.965	1.58	0.968
11	-4.59	0.811	8.18	0.848	5.39	0.901
12	7.07	0.840	52.49	0.494	40.58	0.607
Sum of Lags	832.35	**0.000**	2143.13	**0.000**	2218.59	**0.000**

Table 8--Alternative specification models of the impact of the futures spread changes on U.S. crude inventory changes. Alternatives include--using % U.S. inventory change as a dependent variable and adding lagged controls.

The dependent variable ΔSTOCK, is the weekly change in U.S. non-SPR inventories. In Models 1 and 2, it is the percentage change *100, while in Models 3 and 4 it is the barrel change in inventories. Variable definitions are in Appendix 1. Lagged autoregressive and moving average error terms are included when needed. Any (+ 1) variable indicates a lead for that specific variable, while (-1) is a lag. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. Data is weekly. Models 1 and 3 estimate over 1992-2011 and have 969 observations; Models 2 and 4 over 2004-2011 and have 377 observations.

Panel A.	**Model 1**		**Model 2**		**Model3-with lags**		**Model4-with lags**	
	% Δ U.S. stock		**% Δ U.S. stock**		**Δ U.S. stock**		**Δ U.S. stock**	
	1992-2011		**2004-2011**		**1992-2011**		**2004-2011**	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
C	-61.454	0.000	-71.069	0.003	-2363.15	0.000	-2247.78	0.003
Δ(SPOT_WTI)	-1.032	0.366	-0.026	0.966	-25.61	0.451	8.44	0.648
Δ(REFINERINPUTUS)(-1)					-1.45	0.000	-0.55	0.092
Δ(REFINERINPUTUS)	-0.031	0.024	-0.040	0.000	-1.61	0.000	-2.16	0.000
Δ(REFINERINPUTUS)(+1)	0.062	0.000	0.065	0.000	1.45	0.000	1.80	0.000
Δ(PROD_US)(-1)					3.77	0.000	2.93	0.000
Δ(PROD_US)	0.142	0.000	0.088	0.000	5.11	0.000	5.03	0.000
Δ(PROD_US)(+1)	0.021	0.502	-0.018	0.409	0.60	0.559	0.28	0.669
Δ(IMPORTS_US)(-1)					0.90	0.000	1.49	0.000
Δ(IMPORTS_US)	0.044	0.000	0.057	0.000	2.05	0.000	3.21	0.000
Δ(IMPORTS_US)(+1)	-0.028	0.000	-0.055	0.000	-0.57	0.003	-1.11	0.000
Z2	34.913	0.000	41.034	0.000	1172.76	0.000	1275.53	0.000
Z3	-3.040	0.000	-3.519	0.000	-98.63	0.000	-108.28	0.000
Z4	0.089	0.000	0.101	0.000	2.81	0.000	3.08	0.000
Z5	-0.001	0.000	-0.001	0.000	-0.03	0.000	-0.03	0.000
ΔPDL_1	14.342	0.014	9.582	0.127	472.80	0.011	238.67	0.278
ΔPDL_2	-2.054	0.166	-1.600	0.331	-67.18	0.155	-84.50	0.084
ΔPDL_3	-0.254	0.220	-0.302	0.060	-9.18	0.146	-6.79	0.220
ΔPDL_4	0.041	0.428	0.061	0.141	1.24	0.447	2.16	0.084
AR(1)			0.609	0.000			0.57	0.000
AR(2)			-0.004	0.948				
AR(3)			-0.191	0.001				
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.017**		**0.097**		**0.010**		**0.304**
Adjusted R-squared	28%		70%		31%		76%	

Panel B.	**Δ SPREAD2_1**		**Δ SPREAD2_1**		**Δ SPREAD2_1**		**Δ SPREAD2_1**	
Lags	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
0	8.69	0.186	-4.92	0.120	278.25	0.157	34.78	0.709
1	13.15	0.003	2.37	0.599	424.62	0.002	221.49	0.085
2	15.88	0.001	7.23	0.222	515.51	0.000	329.83	0.055
3	17.12	0.002	10.01	0.131	558.35	0.001	372.76	0.060
4	17.11	0.003	11.08	0.101	560.56	0.001	363.24	0.085
5	16.10	0.006	10.82	0.099	529.57	0.004	314.23	0.148
6	14.34	0.014	9.58	0.127	472.81	0.011	238.67	0.282
7	12.08	0.037	7.74	0.205	397.68	0.038	149.54	0.509
8	9.55	0.094	5.66	0.349	311.63	0.106	59.78	0.794
9	7.00	0.189	3.72	0.531	222.08	0.223	-17.65	0.936
10	4.69	0.306	2.27	0.680	136.45	0.375	-69.79	0.726
11	2.85	0.464	1.69	0.719	62.16	0.609	-83.68	0.595
12	1.73	0.747	2.34	0.560	6.64	0.967	-46.37	0.685
Sum of Lags	140.29	**0.003**	69.60	**0.211**	4476.30	**0.004**	1866.82	**0.340**

Table 9--Alternative specification models of the impact of futures spread changes on Cushing and U.S. (non-SPR) crude inventory changes. Alternatives include using monthly seasonal dummies and non-PDL spread structure.

The dependent variable, ΔSTOCK, is the weekly change in inventories for Cusing in Model 1 and for U.S (non-SPR) in Models 2 and 3. These models do not use ΔPDL spread terms and seasonal Z variables, but use actual spread lags and monthly dummy variables. Variable definitions are in Appendix 1. Lagged autoregressive and moving average error terms are included when needed. Any (+ 1) variables indicate a lead for that specific variable. X denotes PADD 2 data in Model 1 and U.S. data in Models 2 and 3. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. Models 1 evaluates Cushing inventory changes from 2004-2001 has 377 observations; Model 2 U.S. inventory changes from 1992-2011 has 969 observations; Model 3 U.S. inventory changes from 2004-2011, has 377 observations.

	Model 1-no PDL or Zs Δ Cushing stock 2004-2011		**Model 2-no PDL or Zs Δ U.S. stock 1992-2011**		**Model 3-no PDL or Zs Δ U.S. stock 2004-2011**	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
C	490.84	0.004	-1990.45	0.000	-2020.14	0.010
Δ(SPOT_WTI)	4.86	0.727	-45.39	0.118	3.13	0.876
Δ(REFINERINPUT_X)	-0.15	0.817	-1.26	0.005	-1.50	0.000
Δ(REFINERINPUT_X)(+1)	-0.41	0.517	1.72	0.000	2.15	0.000
Δ(PROD_US)	0.40	0.139	3.90	0.000	2.77	0.000
Δ(PROD_US)(+1)	-0.31	0.275	0.21	0.825	-0.56	0.460
Δ(IMPORTS_X)	0.36	0.322	1.31	0.000	1.83	0.000
Δ(IMPORTS_X)(+1)	-0.31	0.405	-0.96	0.000	-1.82	0.000
JAN	-649.41	0.001	2539.72	0.000	3099.97	0.001
FEB	-565.73	0.012	2716.80	0.000	3359.83	0.000
MAR	-31.41	0.890	4205.47	0.000	4017.87	0.000
APR	-231.04	0.249	3486.20	0.000	3828.18	0.000
MAY	-477.90	0.035	2142.67	0.000	2295.30	0.016
JUN	-771.39	0.003	926.52	0.089	982.58	0.284
JUL	-297.25	0.269	1019.91	0.048	1201.85	0.186
AUG	-762.99	0.001	1059.09	0.047	1219.87	0.204
SEP	-934.94	0.000	690.61	0.236	698.94	0.450
OCT	-467.93	0.044	3273.74	0.000	3096.60	0.003
NOV	33.93	0.879	2187.38	0.000	2396.10	0.010
Δ Spread2(-1)	90.92	0.328	419.65	0.017	147.96	0.376
Δ Spread2(-2)	245.56	0.000	342.70	0.114	40.73	0.863
Δ Spread2(-3)	285.23	0.000	130.71	0.591	-20.67	0.943
Δ Spread2(-4)	164.99	0.022	419.62	0.068	176.90	0.480
Δ Spread2(-5)	176.31	0.001	478.21	0.055	483.24	0.030
Δ Spread2(-6)	97.54	0.097	882.58	0.001	754.19	0.008
Δ Spread2(-7)	176.81	0.009	565.98	0.069	521.06	0.090
Δ Spread2(-8)	95.09	0.177	240.18	0.395	168.80	0.600
Δ Spread2(-9)	6.05	0.907	189.11	0.446	-90.19	0.705
Δ Spread2(-10)	-28.03	0.536	22.72	0.899	-62.01	0.739
Δ Spread2(-11)	-55.49	0.298	-65.46	0.729	-59.92	0.741
Δ Spread2(-12)	77.16	0.143	220.94	0.318	82.74	0.609
AR(1)					0.58	0.000
AR(2)					-0.02	0.732
AR(3)					-0.13	0.049
Joint Wald Test Δ SPRD2		**0.000**		**0.039**		**0.252**
Adjusted R-squared	15%		28%		70%	

Table 10 --Alternative specification models of the impact of the futures spread changes on Cushing and total U.S. (non-SPR) inventory changes. Alternative spread definition is the difference between the two-month future and the spot WTI crude price.

The dependent variable, ΔSTOCK, is the weekly change in inventories for Cushing in Model 1 and for U.S (non-SPR) in Models 2 and 3. Variable definitions are in Appendix 1. Lagged autoregressive and moving average error terms are included when needed. Any (+ 1) variables indicate a lead for that specific variable. X denotes PADD 2 data in Model 1 and U.S. data in Models 2 and 3. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. Data is weekly. Models 1 evaluates over 2004-2001, has 377 observations; Model 2, over 1992-2011, has 969 observations; Model 3, over 2004-2011, has 377 observations.

Panel A.	**Model 1**		**Model2**		**Model 3**	
	Δ Cushing stock		**Δ U.S. stock**		**Δ U.S. stock**	
	2004-2011		**1992-2011**		**2004-2011**	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
C	-584.66	0.018	-1985.58	0.000	-1516.11	0.073
Δ(SPOT_WTI)	12.74	0.419	-34.42	0.347	3.46	0.876
Δ(REFINERINPUT_X)	-0.20	0.739	-1.03	0.017	-1.29	0.000
Δ(REFINERINPUT_X)(+1)	-0.27	0.661	1.88	0.000	2.25	0.000
Δ(PROD_US)	0.55	0.040	4.19	0.000	2.97	0.000
Δ(PROD_US)(+1)	-0.28	0.279	0.57	0.538	-0.44	0.553
Δ(IMPORTS_X)	0.44	0.203	1.37	0.000	1.89	0.000
Δ(IMPORTS_X)(+1)	-0.27	0.464	-0.90	0.000	-1.71	0.000
Z2	194.54	0.002	1107.82	0.000	1158.45	0.000
Z3	-14.05	0.003	-96.24	0.000	-102.24	0.000
Z4	0.35	0.011	2.80	0.000	2.95	0.000
Z5	0.00	0.036	-0.03	0.000	-0.03	0.000
ΔPDL_1	143.89	0.001	395.21	0.020	94.03	0.633
ΔPDL_2	-55.68	0.001	-57.10	0.111	-70.47	0.081
ΔPDL_3	-1.63	0.228	-7.28	0.142	-4.03	0.436
ΔPDL_4	1.43	0.009	1.24	0.275	1.64	0.122
AR(1)					0.55	0.000
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.000**		**0.039**		**0.287**
Adjusted R-squared	13%		28%		69%	

Panel B.	**Δ SPREAD2_SPOT**		**Δ SPREAD2_SPOT**		**Δ SPREAD2_SPOT**	
Lags	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
0	111.46	0.043	208.10	0.119	18.07	0.862
1	203.41	0.000	343.83	0.003	140.95	0.254
2	249.33	0.000	427.83	0.002	206.64	0.200
3	257.79	0.000	467.53	0.003	224.97	0.227
4	237.34	0.000	470.37	0.005	205.76	0.298
5	196.52	0.000	443.79	0.009	158.83	0.426
6	143.89	0.000	395.21	0.020	94.03	0.633
7	88.01	0.042	332.07	0.047	21.18	0.912
8	37.41	0.422	261.80	0.107	-49.90	0.788
9	0.66	0.989	191.85	0.205	-109.37	0.528
10	-13.70	0.731	129.64	0.323	-147.40	0.327
11	2.89	0.927	82.61	0.445	-154.17	0.175
12	58.98	0.236	58.20	0.637	-119.84	0.118
Sum of Lags	1574.00	**0.000**	3812.82	**0.010**	489.74	**0.772**

Table 11 --Alternative specification models of the impact of the futures spread changes on Cushing and total U.S. (non-SPR) inventory changes. Alternative spread definition is the difference between the three- and one-month WTI crude futures.

The dependent variable, ΔSTOCK, is the weekly change in inventories for Cusing in Model 1 and for U.S (non-SPR) in Models 2 and 3. Variable definitions are in Appendix 1. Lagged autoregressive and moving average error terms are included when needed. Any (+ 1) variables indicate a lead for that specific variable. X denotes PADD 2 data in Model 1 and U.S. data in Models 2 and 3. The regression is run via OLS (Ordinary Least Squares) with the Newey-West adjustment. Data is weekly. Models 1 estimates over 2004-2001, has 377 observations; Model 2, over 1992-2011, has 969 observations; Model 3, over 2004-2011, has 377 observations.

Panel A.	**Model 1**		**Model2**		**Model 3**	
	Δ Cushing stock		**Δ U.S. stock**		**Δ U.S. stock**	
	2004-2011		**1992-2011**		**2004-2011**	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
C	-554.72	0.029	-2001.48	0.000	-1656.81	0.053
Δ(SPOT_WTI)	24.64	0.167	-25.83	0.499	1.24	0.955
Δ(REFINERINPUT_X)	-0.21	0.724	-1.02	0.018	-1.30	0.000
Δ(REFINERINPUT_X)(+1)	-0.27	0.650	1.88	0.000	2.23	0.000
Δ(PROD_US)	0.55	0.037	4.20	0.000	3.02	0.000
Δ(PROD_US)(+1)	-0.27	0.271	0.61	0.514	-0.32	0.662
Δ(IMPORTS_X)	0.43	0.208	1.37	0.000	1.89	0.000
Δ(IMPORTS_X)(+1)	-0.27	0.452	-0.90	0.000	-1.71	0.000
Z2	185.02	0.003	1114.38	0.000	1174.50	0.000
Z3	-13.36	0.005	-96.90	0.000	-103.20	0.000
Z4	0.33	0.014	2.82	0.000	2.98	0.000
Z5	0.00	0.041	-0.03	0.000	-0.03	0.000
ΔPDL_1	105.49	0.001	326.70	0.008	146.23	0.340
ΔPDL_2	-38.47	0.003	-43.97	0.184	-22.10	0.554
ΔPDL_3	-0.19	0.886	-5.04	0.260	-4.27	0.313
ΔPDL_4	0.71	0.110	0.93	0.423	0.83	0.402
AR(1)					0.55	0.000
Joint Wald Test ΔPDL_1-ΔPDL_4		**0.000**		**0.008**		**0.795**
Adjusted R-squared	15%		28%		69%	

Panel B.	**Δ SPREAD3_1**		**Δ SPREAD3_1**		**Δ SPREAD3_1**	
Lags	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
0	175.92	0.003	207.33	0.157	-53.42	0.587
1	204.23	0.000	303.79	0.002	46.67	0.657
2	210.84	0.000	362.15	0.000	113.42	0.391
3	200.00	0.000	388.02	0.001	151.79	0.310
4	175.98	0.000	387.00	0.001	166.74	0.284
5	143.06	0.000	364.69	0.003	163.24	0.293
6	105.49	0.001	326.70	0.008	146.23	0.340
7	67.54	0.042	278.63	0.025	120.69	0.428
8	33.49	0.346	226.08	0.069	91.57	0.546
9	7.59	0.828	174.65	0.137	63.83	0.665
10	-5.89	0.845	129.96	0.197	42.44	0.754
11	-2.68	0.925	97.59	0.257	32.34	0.779
12	21.48	0.665	83.17	0.498	38.51	0.711
Sum of Lags	1337.03	**0.000**	3329.74	**0.001**	1124.05	**0.403**