Speculative Floating Oil*

Andrei Kirilenko   Anna Kruglova

November 22, 2018

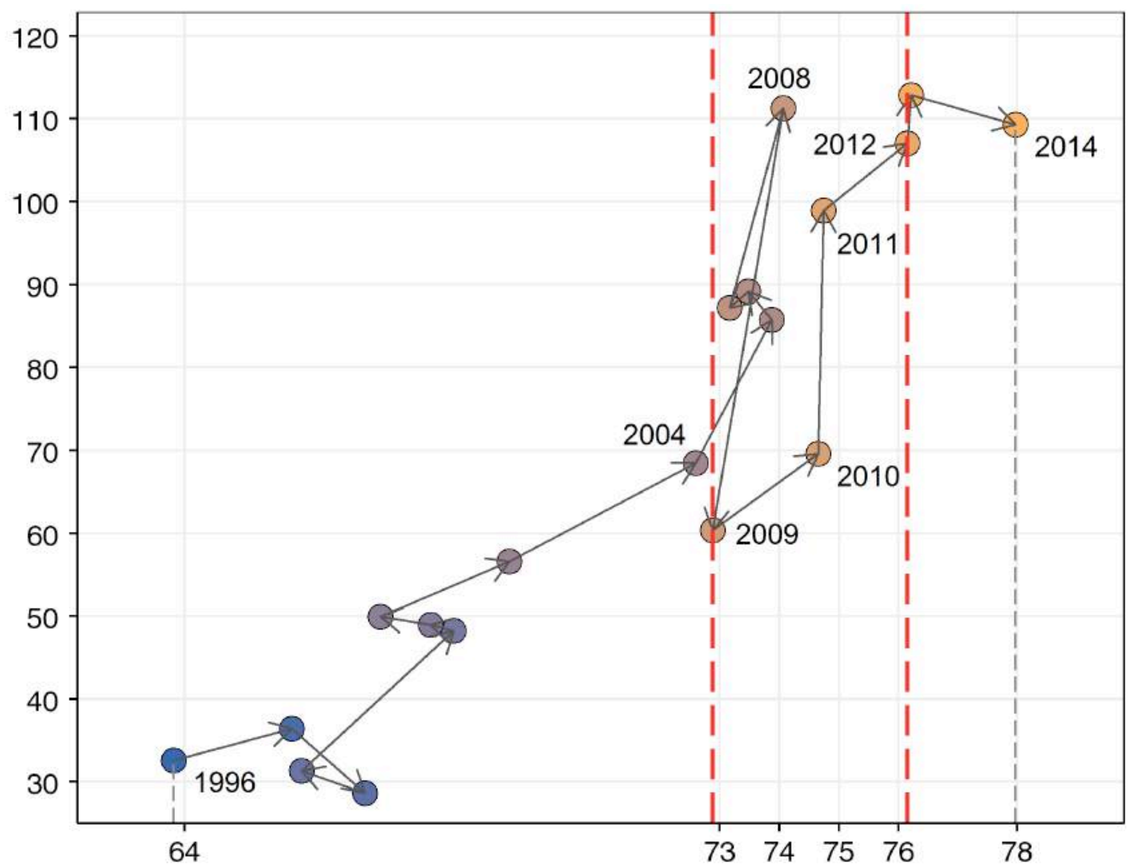
ABSTRACT

We introduce commodity cash-and-carry traders into a limits to arbitrage extension to the speculative storage framework. Carry traders are assumed to have access to a market for off-shore storage technology — shipping vessels that can be chartered to store physical commodity bought in one period for delivery in the next period. We show that in equilibrium, arbitrage activities of cash-and-carry traders are linked with the risk premium in commodity prices and the cost of floating storage. We test empirical predictions of the model using a novel data set with granular information on every tanker that delivered seaborne crude oil into the United States during 2008-2012. Our empirical results are consistent with the predictions of the model and are very strongly present in the data.

*Andrei Kirilenko is with the Imperial College Business School and CEPR and Anna Kruglova is with the Foster School of Business at the University of Washington. We thank Rene Carmona, Rama Cont, Alexander Eydelan, Michael Gordy, James Hamilton, Lutz Kilian, Andrew Papanicolaou, Vikas Raman, Ehud Ronn, Richard Sowers, Thaleia Zariphopoulou, and participants at the 2015 Program on Broad Perspectives and New Directions in Financial Mathematics at the UCLA Institute for Pure and Applied Mathematics, 2016 Energy and Commodity Finance Conference, 2016 Informs Annual Meeting, 2018 Commodity Markets Conference, Duke, Imperial College Business School, University of St. Gallen, University of Lausanne, University of Texas at Austin, the Central Bank of the Russian Federation, and the National Bank of Ukraine for very helpful comments and suggestions. All remaining errors are our own.

Global production of crude oil increased steadily from 64 million barrels per day in 1996 to 73 million barrels per day in 2004. Crude oil prices followed a similar fairly steady upward pattern. During the subsequent five and a half years—from 2004 to mid-2010—global crude oil production remained at about the same level of 73-74 million barrels per day, while prices had a roller coaster ride.

Figure 1: Global Crude Oil Supply and Brent Nearby Futures Prices: 1996–2014

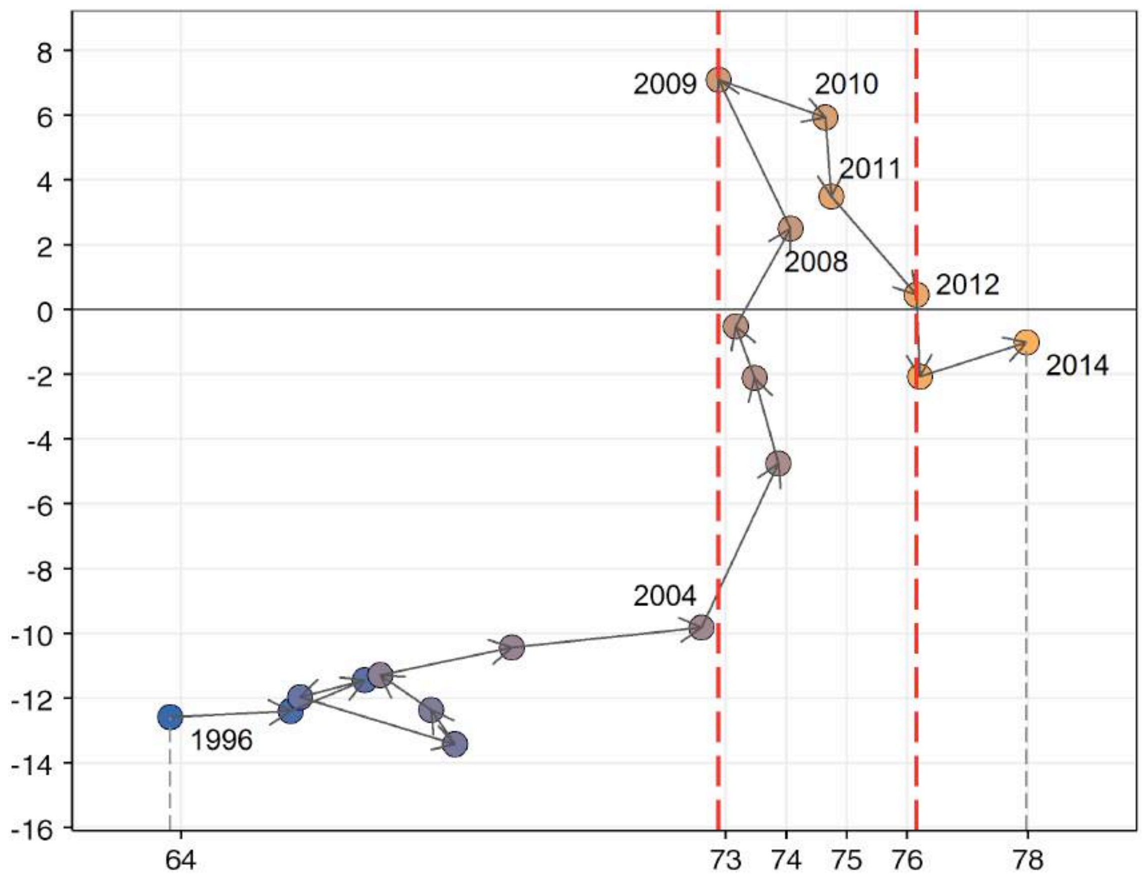


Sources: Data on global crude oil supply is from the U.S. Energy Information Administration. Brent futures prices are from the Intercontinental Exchange. This figure presents global crude oil production and Brent nearby futures prices. Global supply is in millions of barrels per day. Annual average crude oil prices are computed by averaging weekly nearby Brent futures prices.

Starting in about mid–2010, global oil supply began to rise largely due to the use of new on-shore production technologies such as slant drilling and hydraulic fracturing. In 2010-2011, global oil supply reached 75 million barrels per day while prices rose from 70 to 100 dollars per barrel. In 2012, global crude oil production increased to 76 million barrels per day while prices fluctuated around 110 dollars per barrel.

During 2008-2012, the term structure of Brent futures prices has stayed on average highly upward sloping, a.k.a. in contango, and often very highly upward sloping. Periods when the term structure was very highly upward sloping were called the times of supercontango.

Figure 2: Global Crude Oil Supply and the Slope of the Term Structure of Brent Futures Prices: 1996–2014



Sources: Global crude oil supply data is from the U.S. Energy Information Administration. Brent futures prices are from the Intercontinental Exchange. This figure presents global crude oil supply and the slope of the term structure of Brent futures prices. Global production is in millions of barrels per day. The slope of the Brent futures term structure is in dollars per thousand barrels calculated as the difference between the futures contracts three months out and the nearby. Annual averages are computed from daily term structure prices.

A period of contango, let alone supercontango, is likely to be noticed by speculators – those, who are interested in buying one hundred thousand barrels of crude oil at, say, 50 dollars per barrel now, storing it, and then selling it at 55 dollars per barrel two to three months later.

Canonical speculative storage models of Deaton and Laroque (1992, 1996) capture this logic when deriving an equilibrium relationship between prices, inventories, supply and demand for a storable commodity like crude oil. In the models, total supply consists of new production in a given period (assumed for technological reasons to be exogenous) and inventory that was deliberately put into storage in the previous period. Total demand for a commodity in a given period consists of stochastic consumer demand and demand for inventory to be put into storage until the next period. The amount of storage reflects a speculative decision based on the current and expected future prices and storage costs.

In equilibrium, current price is predicted to be equal to the expected discounted future price minus the cost of storage. If the the expected discounted future price is higher than the current price and the premium exceeds the cost of storage, the speculators will have a strong motive to buy a commodity at the current price, put it in storage, and sell it at the higher price in the future.

Crucially, however, the reverse is not true. Namely, the current price can be much higher than the expected discounted future price and there is little speculators can do about it because speculative storage cannot be negative — supply cannot be borrowed from future periods for current consumption. The canonical commodity price formation framework with speculative storage is thus able to match a crucial empirical fact of commodity prices — "long periods of stagnant prices interrupted by sharp upward spikes" (Deaton and Laroque (2003)).

The framework, however, fails to match other aspects of the commodity price process, most critically high autocorrelation of prices to the point of nonstationarity and autoregressive heteroscedasticity combined with structural breaks. Moreover, the models of Deaton and Laroque (1992, 1996) and their extensions assume "free entry into the storage sector as well as risk neutrality, implying that the actions of arbitrageurs will raise or lower the current price until it is at a level which renders the strategy unprofitable in expectation" (Dvir and Rogoff (2014)). Assumption of risk neutrality also implies that the canonical framework has little to say about futures prices as agents do not have a reason to hedge commodity price risks.

There is, however, a growing theoretical and empirical literature which argues that because arbitrageurs are capital constrained, arbitrage activities are subject to frictions and limitations that manifest themselves in both market prices and positions of arbitrageurs (see, for example, Shleifer and Vishny (1997)). Moreover, if arbitrage activities require specialized knowledge, they are delegated to specialized funds operated by asset managers, who extract the entire surplus from these activities subject to leverage constraints — adding to market frictions in the process (see, for example, Berk and Green (2004)).

Acharya, Lochstoer and Ramadorai (2013) develop a two-period equilibrium model of commodity markets that includes frictions due to limits to arbitrage. The model consists of commodity consumers, commodity producers, and asset managers. In each period, commodity consumers demand a certain amount of commodity, e.g., crude oil, in the physical (cash) market. Competitive producing firms supply the physical market with an inelastic supply of the commodity save for an amount that they choose to store as inventory and make available in the next period. Risk averse producers also have access to the futures market where they can hedge their natural long position in the physical commodity by taking a short futures position. The producers' hedging demand in the futures market is accommodated by specialized, capital-constrained commodity asset managers, who provide the long side of the futures trade in return for appropriate compensation and only up to a limited size. The authors show that frictions and limitations imposed on the producers and commodity asset managers help explain how hedging activities in the futures market translate into equilibrium commodity prices as a function of speculative inventory. By assumption, arbitrageurs in Acharya, Lochstoer and Ramadorai (2013) trade only in the futures market, but not in the cash market. Thus, in order to link cash and futures market prices with speculative inventory decisions, the authors rely on the concept of a convenience yield — an assumed reason to hold inventory for ease of future access or "an embedded timing option" (Routledge, Seppi and Chester Spatt (2000)).

In this paper, we propose an extension to the equilibrium model of Acharya, Lochstoer and Ramadorai (2013) by adding commodity cash-and-carry traders—risk averse arbitrageurs who possess specialized knowledge and technology to arbitrage between the physical (cash) and futures markets—and test implications of the model by using a unique, highly disaggregated data for crude oil imports into the United States.

Cash-and-carry traders are assumed to have access to a market for *off-shore* floating storage technology — shipping vessels that can be used to store physical crude oil in one period and deliver it in the subsequent period.[1] The market for off-shore floating storage technology is assumed to be competitive and driven solely by the demand and supply of vessels suitable for storage and transportation of crude oil. It is further assumed that the floating storage market is open only to carry traders at a cost proportional to the amount stored as floating inventory. In addition to having access to costly floating storage technology, carry traders possess appropriate knowledge to trade in the crude oil cash and futures markets. Carry traders are assumed to be risk-averse and constrained by the size of their unhedged arbitrage position.

In equilibrium, we preserve the canonical results of Deaton and Larogue, i.e. "long periods of stagnant prices interrupted by sharp upward spikes." In addition, we show that in equilibrium, during the periods of sharp upward spikes in prices in the physical market, futures prices are also in contango or even supercontango. During these periods, equilibrium speculative off-shore inventory of commodity carry traders is positive. In contrast, when the term structure is slightly upward sloping, flat or downward sloping, equilibrium inventory of commodity carry traders is at zero.

Empirically, we should expect to see a component of inventories associated with

[1] On the use of floating storage to speculate in the market for crude oil see, for example, Atkins (2016).

activities of cash-and-carry traders to begin rising around 2008 until about mid–2010 (when contango is increasing) and then to decline from the second half of 2010 until the end of 2012 (when contango is decreasing). The amounts put into off-shore speculative storage should be of relatively limited size due to frictions and limitations of executing cash-and-carry arbitrage.

Instead of relying on aggregate inventory data as in Kilian and Murphy (2013) and Kilian and Lee (2014), we are able to use granular data on every tanker that delivered seaborne crude oil into the United States during 10/01/2008–12/31/2012 to build a proxy for the speculative floating inventory component. As predicted by theory, we find that speculative floating inventory is strongly positively related to the slope of Brent futures prices and negatively related to the costs of using vessels for use as speculative storage.

We observe that speculative floating inventory imported into the U.S. increased during 2008–August 2010 when the global supply response was constrained and then decreased to zero during the second half of 2010 and into 2011 when the U.S. was increasing its domestic production of crude oil even though the term structure of (globally determined) Brent futures prices remained on average in contango. We also find that price volatility was lower when the speculative inventory was rising during 2008—mid-2010.

The remainder of the paper is as follows. Section 1 presents the model. Section 2 presents considerations for our empirical strategy to examine the predictions of the model by using disaggregated U.S. imports data for crude oil. Section 3 presents derivation of the time series for floating speculative inventory. Section 4 presents our empirical analysis. Section 5 concludes.

I. An Equilibrium Model of Commodity Prices with Speculative Storage and Cash-and-Carry Traders

We develop an equilibrium model of commodity prices in the presence of speculative storage and limits to arbitrage. Our model extends the equilibrium model of Acharya, Lochstoer and Ramadorai (2013) by adding cash-and-carry traders - arbitrageurs who possess specialized knowledge and technology to arbitrage between the physical and the futures markets for crude oil.

There are four types of agents in the two-period equilibrium model: consumers of crude oil, producers of crude oil, commodity fund managers, and cash-and-carry traders. The first three types of agents are modeled the exact same way as in Acharya, Lochstoer and Ramadorai (2013). A description of these agents just sufficient for our purposes is below. For ease of exposition, wherever possible we preserve the original notation.

A. Consumers

Consumers maximize the following objective function:

$$u\left(C_0, Q_0\right) + \beta E_0 u\left(C_1, Q_1\right), \tag{1}$$

where Q_t is the total quantity of crude oil supplied and C_t is consumption of other goods, respectively in period $t = 0, 1$.

$$u\left(C_t, Q_t\right) = \frac{1}{1-\rho}\left(\left(C_t^{\frac{\epsilon-1}{\epsilon}} + \omega Q_t^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}\right)^{1-\rho} \tag{2}$$

is assumed to be a constant elasticity of substitution utility function with the elasticity of substitution $\epsilon > 0$, the relative risk aversion $\rho > 0$, and the share parameter $\omega > 0$.

The solution to the first-order optimality condition implies the inverse demand function of the form

$$S_t = \omega\left(\frac{C_t}{Q_t}\right)^{\frac{1}{\epsilon}}, \tag{3}$$

where S_t is the price of crude oil in period $t = 0, 1$.

The marginal rate of substitution between periods 0 and 1 in the Euler equation is given by

$$\Lambda = \beta\left(\frac{C_0}{C_1}\right)^{-\rho}\left(\frac{1+\omega(\frac{Q_1}{C_1})^{\frac{\epsilon-1}{\epsilon}}}{1+\omega(\frac{Q_0}{C_0})^{\frac{\epsilon-1}{\epsilon}}}\right)^{\frac{\frac{1}{\rho}-\epsilon}{\frac{\epsilon-1}{\rho}}}. \tag{4}$$

For parsimony, demand for other goods, C_t is assumed to be an exogenous random variable with $E(\ln C_t) = \mu$ and $Var(\ln C_t) = \sigma^2$. This assumption together with the functional form of the inverse demand function implies the variance of the spot price is

$$\sigma_s^2 = \omega^2 Q_1^{-\frac{2}{\epsilon}}\left(e^{\frac{\sigma^2}{\epsilon^2}} - 1\right)e^{\frac{2\mu}{\sigma}+\frac{\sigma^2}{\epsilon^2}} = \kappa Q_1^{-\frac{2}{\epsilon}}, \tag{5}$$

where $\kappa > 0$ is a constant.

B. Producers

There is an infinite number of production firms with a mass normalized to unity. Production firms are operated by production managers who have access to three technologies. Firstly, production managers operate a specialized production technology that exogenously generates a deterministic output of g_0 barrels of crude oil in period zero and g_1 barrels in period one.

Secondly, production managers have access to *on-shore* storage technology for crude oil. This storage technology is available to all production managers on the same terms as they collectively own it. Namely, for i barrels of crude oil put into on-shore storage facility at time zero, a production manager receives $i(1-\delta)$ barrels of crude oil in period one, where $0 < \delta < 1$ denotes depreciation due to storage in physical terms (barrels).

Thirdly, production managers have access to the crude oil futures market where they can hedge against fluctuations in spot crude oil prices.

Production managers are assumed to be risk averse maximizers of the value of their firms (the firms, in turn, are fully owned by consumers) subject to the variance of next period earnings. To do so, in period zero, a production manager sells $g_0 - i$ barrels of crude oil in the physical market at the period-zero spot price of S_0 dollars per barrel. In addition, the production manager sells h_p barrels worth of futures contracts in the futures market at the price of F dollars per barrel (F is known in period zero). In period one, the production manager sells $i(1-\delta)+g_1$ barrels of crude oil at the period-one spot price of S_1 dollars per barrel and cash settles the short futures position, h_p.

The objective function of a representative production manager is formally described as follows:

$$\begin{aligned} \max_{i,h_p} & S_0\left(g_0 - i\right) + h_p F + E\left\{\Lambda S_1\left(i(1-\delta) + g_1 - h_p\right)\right\} \\ & -\frac{\gamma_p}{2} Var\left\{S_1\left(i(1-\delta)+g_1\right) + h_p\left(F - S_1\right)\right\} \\ s.t. \quad & i \geq 0, \end{aligned} \tag{6}$$

where γ_p denotes a representative production manager's risk aversion coefficient.

The first-order condition with respect to i gives the optimal rule for on-shore inventory:

$$i^\star(1-\delta) = \frac{(1-\delta)E\left\{\Lambda S_1\right\} - S_0 + \lambda_i}{(1-\delta)\gamma_p\sigma_s^2} - g_1 + h_p, \tag{7}$$

where λ_i is the Lagrange multiplier on the on-shore inventory nonnegativity constraint and σ_s^2 is the variance of spot crude oil prices.

Optimal on-shore inventory rises with an increase in the amount hedged by the producer in the futures market and falls with an increase in the risk aversion coefficient of the producer, spot price volatility and production in the next period. In the event of an on-shore inventory stock-out, $\lambda_i > 0$ and the current spot price S_0 can be higher than expected discounted future spot price as in the speculative storage framework of Deaton and Laroque (1992, 1996).

The first-order condition with respect to h_p gives the optimal rule for hedging demand:

$$h_p^\star = \frac{F - E\left\{\Lambda S_1\right\}}{\gamma_p\sigma_s^2} + i^\star(1-\delta) + g_1. \tag{8}$$

Notably, if $F = E\{\Lambda S_1\}$ indicating that the futures and spot crude oil markets are free of frictions and limitations, then it is optimal for a representative production manager to be fully hedged, i.e., to set $h_p^\star = i^\star(1-\delta) + g_1$. In contrast, if $E\{\Lambda S_1\} > F$, then it is optimal for a representative production manager to be less than fully hedged, i.e., to demand a smaller short open position in the futures markets.

C. Commodity fund managers

Commodity fund managers are risk-averse long-only speculative investors who possess a specialized knowledge to invest in spot and futures markets for crude oil; they do not have capacity to invest in oil producing firms nor have access to the physical crude oil storage technology of any kind. Their optimization decisions are constrained by the variance of their net speculative position. In period zero, a commodity fund manager goes long h_s barrels of futures contracts in the futures market at a price of F dollars per barrel. In period one, the commodity fund manager cash settles the entire long position at the period-one spot price of S_1 dollars per barrel.

The objective function of a representative commodity fund manager is described as follows:

$$\max_{h_s} h_s\left(E\{\Lambda S_1\} - F\right) - \frac{\gamma_s}{2} Var\left(h_s\left(S_1 - F\right)\right), \tag{9}$$

where γ_s denotes a representative commodity fund manager's risk aversion coefficient.

The first-order condition with respect to h_s gives the optimal rule for the optimal long speculative position in the futures market:

$$h_s^\star = \frac{E\{\Lambda S_1\} - F}{\gamma_s \sigma_s^2}. \tag{10}$$

If $E\{\Lambda S_1\} > F$, commodity fund managers are optimally willing to provide a greater long open interest.

D. Cash-and-carry traders

Cash-and-carry traders are a new type of agent that we introduce into the model of Acharya, Lochstoer and Ramadorai (2013). These traders possess a specialized knowledge to arbitrage between the cash (physical) and futures markets for crude oil. In the cash market, they are assumed to have access to *off-shore* costly storage technology — shipping vessels that can be used to store physical crude oil in period zero for delivery in period one. The floating storage market is open only to carry traders at a cost proportional to the amount stored as floating inventory. The market for off-shore floating storage technology is assumed to be competitive and driven solely by the demand and supply of vessels suitable for storage and transportation of crude oil. To that end, it is assumed that between period zero and one, the market for floating storage is unaffected by any frictions, limitations, quantities and prices in the physical or futures markets for crude oil. Carry traders are further assumed to not have access to the on-shore storage facilities owned and operated by the producers. Lastly, in addition to having access to costly floating storage technology, carry traders possess appropriate knowledge to trade in the cash and futures markets for crude oil.

Carry traders are risk averse. Their optimization decisions are constrained by the variance of the value of their speculative position. In period zero, a carry trader buys y barrels of crude oil in the at the period-zero spot price of S_0 dollars per barrel and puts the entire inventory into floating storage at a cost R_0 dollars per barrel. The carry trader also sells short h_c barrels worth of futures contracts at the price of F dollars per barrel. In period one, the floating inventory is delivered and the carry trader sells the physical inventory and cash settles the short position at the period-one spot price of S_1 dollars per barrel.

The objective function of a representative carry trader is described as follows:

$$\max_{y,h_c} -yS_0 - yR_0 + h_cF + E\left\{\Lambda S_1\right\}(y - h_c) - \frac{\gamma_c}{2} Var\left\{S_1 (y - h_c)\right\} \tag{11}$$
$$s.t. \quad y \geq 0,$$

where γ_c denotes a representative carry trader's risk aversion coefficient.

The first-order condition with respect to y gives the optimal rule for the optimal floating inventory:

$$y^\star = \frac{E\left\{\Lambda S_1\right\} - S_0 - R_0 + \lambda_y}{\gamma_c \sigma_s^2} + h_c, \tag{12}$$

where λ_y is the Lagrange multiplier on the off-shore inventory nonnegativity constraint.

Optimal off-shore inventory rises with an increase in the amount hedged by the carry trader in the futures market and falls with the cost of floating storage, an increase in the risk aversion of the carry trader and spot price volatility. In the event of an off-shore inventory stock-out, $\lambda_y > 0$ and the current spot price S_0 can be higher than expected discounted future spot price.

The first-order condition with respect to h_c gives the optimal rule for a carry trader's optimal (short) arbitrage position in the futures market:

$$h_c^\star = \frac{F - E\left\{\Lambda S_1\right\}}{\gamma_c \sigma_s^2} + y^\star. \tag{13}$$

Note that if $E\left\{\Lambda S_1\right\} > F$, then it is optimal for a carry trader to demand a smaller short position. This is in the same direction as the producer (who is also short futures) and in the opposite direction from the asset manager (who is providing the long side).

E. Equilibrium

The market for physical crude oil clears the same was as in Acharya, Lochstoer and Ramadorai (2013) with an addition of the speculative floating oil. Recall that unlike on-shore storage technology that results in a loss of physical oil of size $0 < \delta < 1$ between periods zero and one, floating storage is costly, but does not result in any intertemporal loss of physical oil. Accordingly, in equilibrium, supply in period one,

$$Q_1 = Y^\star + I_0^\star(1-\delta) + G_1, \tag{14}$$

where $Y^\star$ is the aggregate optimal off-shore inventory, $I_0^\star$ the aggregate on-shore inventory carried over from period one, and G_1 is the aggregate production in period one.

Futures market clears in accordance under the zero net supply condition

$$h_p^\star + h_c^\star = h_s^\star, \tag{15}$$

with long-only positions established by commodity fund managers having to be exactly equal to the sum of short positions demanded by the commodity producers and the carry traders. Note that both commodity carry traders and commodity producers demand a short position in the futures markets against the limited capacity of capital-constrained commodity asset managers to provide the long side. This introduces additional limits to hedging for commodity producing firms and translates into equilibrium prices.

Substituting Equations (8), (10), and (13) into Equation (15), and then substituting Equation (14) for the market clearing optimal supply results in

$$E\left\{\Lambda S_1\right\} - F = \gamma\sigma_s^2 Q_1\left(Y^\star, I_0^\star\right), \tag{16}$$

where $\gamma > 0$ is defined such that $\frac{1}{\gamma} = \frac{1}{\gamma_p} + \frac{1}{\gamma_s} + \frac{1}{\gamma_c}$. Further substituting Equation (5) for σ_s^2 results in

$$E\left\{\Lambda S_1\right\} - F = \gamma\kappa Q_1^{1-\frac{2}{\epsilon}}\left(Y^\star, I_0^\star\right), \tag{17}$$

Substituting out the futures price using the equilibrium arbitrage condition (19) and the period one equilinrium supply (14) results in

$$E\left\{\Lambda S_1\left(Y^\star, I_0^\star\right)\right\} - S_0\left(Y^\star, I_0^\star\right) - R_0 + \lambda_y\left(Y^\star\right) = \gamma\kappa\left(Y^\star + I_0^\star(1-\delta) + G_1\right)^{1-\frac{2}{\epsilon}}, \tag{18}$$

an equation that combines spot prices and aggregate supply as functions of equilibrium on-shore and off-shore inventory.

F. Model predictions and other considerations

Combining the two first-order conditions for the cash-and-carry trader given by Equations (12) and (13) results in an equilibrium arbitrage condition of the form:

$$\frac{F - S_0 - R_0 + \lambda_y\left(y^\star\right)}{\gamma_c\sigma_s^2} = 0. \tag{19}$$

Equation (19) gives predictions about the relationship between the slope of the term structure of futures prices, $F - S_0$, speculative floating inventory, $y^\star$, and the cost of floating storage, R_0.

Intuitively, when $F > S_0 + R_0$, i.e., the (deferred) futures price is higher than the current spot (or nearby futures) price and the cost of off-shore storage, it is optimal to set $y^\star > 0$ and profit from the carry trade. In contrast, when $F \leq S_0 + R_0$, it is optimal to set $y^\star = 0$ (and $\lambda_y > 0$). From this intuition, a change in λ_y is negatively related to a change in $y^\star$, and there is a non-linearity or a regime change at the level of the "floating inventory stockout" when $y^\star = 0$.

Model predictions reflect the assumption that producers operate a specialized production technology that exogenously generates g_0 barrels of crude oil in period zero and g_1 barrels in period one. This is the same assumption as in Deaton and Laroque (1992, 1996). Dvir and Rogoff (2009, 2014) extend the canonical speculative storage framework by allowing for endogenous supply response. The authors derive equilibrium responses to higher current oil prices under two stylized regimes: constrained supply and unconstrained supply. They show that if the supply of oil in the current period is technologically or otherwise unconstrained, then, in (a rational expectations) equilibrium, suppliers respond by selling oil in the current period until the temporarily elevated current price declines to the level where current price and the expected future price minus storage costs equalize. In this regime, inventories held in speculative storage decline; possibly all the way to zero.

However, if for technological, regulatory or other reasons, the supply of oil in the current period is constrained while the demand for oil is expected to remain elevated, the equilibrium dynamics of prices and inventories follows a different pattern. In the constrained supply regime, in equilibrium, both (already high) current price and expected future price will increase because as the authors put it "rising prices due to rising demand can be seen as a process which is likely to continue." All in all, the expected future price minus storage costs rises to the level of the elevated current price because the supply response is not expected to be forthcoming. Equilibrium inventories held in speculative storage also increase.

While we maintain the exogenous supply assumption for commodity producers as in Deaton and Laroque (1992, 1996), the presence of cash-and-carry traders effectively introduces a mechanism to adjust oil supply between adjacent periods, albeit of limited capacity. It is, thus, straightforward to extend our empirical predictions regarding floating storage to regimes with constrained and unconstrained supply of Dvir and Rogoff (2009, 2014) as long as we can pinpoint these regimes in the data.

II. Considerations for the empirical strategy

A. Practical considerations

Understanding the practicalities of setting up a cash-and-carry trade in crude oil is useful for considerations for our empirical strategy. In practice, a speculative cash-and-carry trade that links physical and futures markets for crude oil is set up very similarly to the way it is described in the theoretical model.[2]

Consider the following simplified cash-and-carry trade between the Arabian Gulf and the U.S. Gulf. At time t, a carry trader buys Y_t barrels of crude oil in the Arabian Gulf at the price S_t dollars per barrel. The trader also charters a vessel at the daily cost R_t dollars per barrel in order to ship Y_t barrels to the U.S. Gulf. The vessel is expected to be at sea (both en route and as floating storage) for T days. Knowing that, the trader sells $Y_t \frac{S_t+R_tT}{F_t^{t+T}}$ barrels worth of crude oil futures contracts, where F_t^{t+T} is the time t price of the futures contract that matures in $t+T$. After T days, Y_t barrels of crude oil arrive in the U.S. Gulf. Of this amount, $Y_t \frac{S_t+R_tT}{F_t^{t+T}}$ barrels are delivered at the expiration of the futures contract to settle the short futures position. If $F_t^{t+T} > S_t + R_tT$, then there is $Y_t(1 - \frac{S_t+R_tT}{F_t^{t+T}})$ additional barrels to sell at F_t^{t+T} per barrel, earning $F_t^{t+T}Y_t(1 - \frac{S_t+R_tT}{F_t^{t+T}})$ dollars in profits or $Y_t(F_t^{t+T} - S_t - R_tT)$. The carry trader will continue buying, storing and shipping multiples of Y_t if $F_t^{t+T} - S_t - R_tT > 0$ subject to leverage or financing constraints, as well as additional frictions. Additional frictions include, for example, costs per vessel over and above the costs of shipping.

For illustrative purposes, consider the following numerical example closely based on actual data.

> Y_t is 2 million barrels, the capacity of a Very Large Cargo Carrier (VLCC). F_t^t, the price of the nearby Brent futures contract, which we assume to be approximately equal to S_t, is 57 dollars per barrel. T is 180 days. This includes the estimated time at sea for the route from Ras Tanura, Saudi Arabia to Houston, TX at an average speed of 13.5 knots. This time also includes loading time and unloading times (using Aframax vessels), adverse weather conditions adjustments, laytime, and possible use of the VLCC as floating storage. F_t^{t+180}, the price of the Brent futures contract expiring in 180 days is 66 dollars. Daily time charter of the VLCC, R_t, is 47,000 dollars per day. Futures contracts with delivery in 180 days sold short, $Y_t \frac{F_t^t+R_tT}{F_t^{t+T}}$ is 1,855,455 barrels or 1,856 contracts of Brent futures contracts (per contract specifications, the size each contact is 1,000 barrels). VLCC charter costs add up to 4.23 dollars per barrel. Additional costs include monthly VLCC hull cleaning, two days steaming to remove growth and idle bunkering costs, as well as cargo insurance, lease of Aframax vessels for unloading, possible

[2] For cash-and-carry trades, see also Knittel and Pindyck (2013) and Frankel (2014).

demurrage charges, margin and financing costs and trading fees. Based on the industry estimates, these costs add up to about 4.4 million dollars or 2.19 dollars per barrel. Under these assumptions, total profits per VLCC amount to about 5.16 million dollars or 2.58 dollars per barrel. If the price of the six months out futures contracts drops from 66 to 63.42 dollars per barrel, a cash-and-carry speculator becomes indifferent between engaging in the trade or not.

B. Regression setup

Suppose that at time t, we are able to observe a proxy for the aggregate speculative floating inventory Y_t. Based on the implications of the equilibrium model as specified in Equation (19), we can set up the following empirical specification written in a form of a linear regression

$$F_t^{t+T} - F_t^t = \beta_1 Y_t + \beta_2 T R_t + \epsilon_t, \tag{20}$$

where $F_t^{t+T} - F_t^t$ is the slope of the term structure of futures prices, and R_t is the daily cost of chartering a vessel. The variables $F_t^{t+T} - F_t^t$ and R_t, which are expressed in dollars per barrel, are market determined and observable. The coefficients β_1 and β_2 are expected to be positive. The residual is denoted by ϵ_t.

The empirical specification should also account for a possibility of two separate regimes associated as per Dvir and Rogoff (2009, 2014) with constrained and unconstrained overall supply. If the supply is unconstrained, the amount of speculative floating inventory might be quite small even though the slope of the term structure may be positive and steep and the costs of shipping may be relatively low. If the supply is constrained, however, an increase in the amount of the speculative floating inventory should be positively associated with the slope of the term structure of crude oil futures prices for given costs of floating storage.

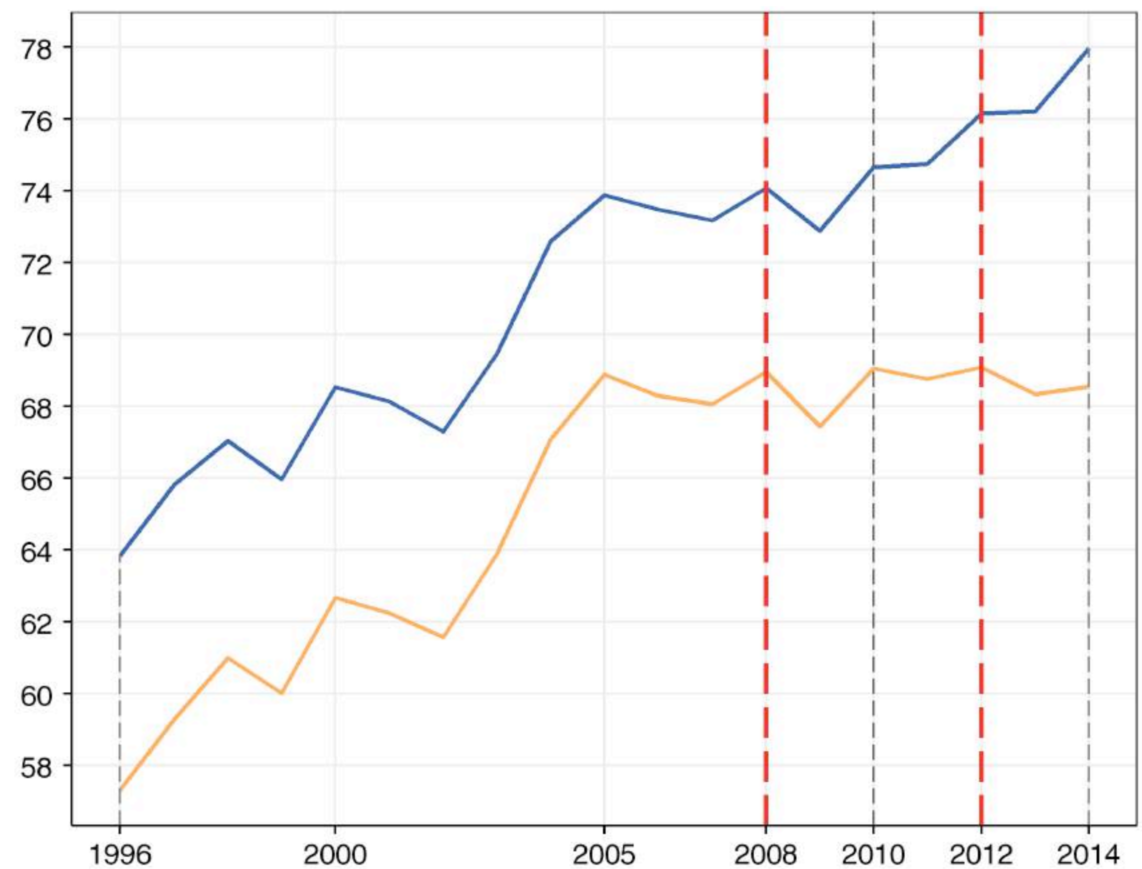
C. Nuanced data considerations

As mentioned before, it is very difficult to empirically single out inventory associated with speculative storage activity out of total global inventories. As noted by Kilian and Murphy (2013) and Kilian and Lee (2014), data on global crude oil inventories is not publicly available; the two studies construct proxies for global crude oil inventories using publicly available data from the U.S. Energy Information Administration (EIA) and the Energy Intelligence Group, respectively. Furthermore, the component of global crude oil inventories associated with speculative activity is also not publicly observable. It is even more difficult to extract speculative floating inventory, Y_t out of the publicly available data.

In this paper, we use granular data on every tanker that delivered seaborne crude oil into the United States during 2008-2012 to derive a proxy for the speculative floating inventory component. Before we describe the data, however, we establish a number of nuanced data considerations that will help us formulate our empirical analysis.

Firstly, increase in global oil supply from mid–2010 on is primarily associated with the surge in oil production in one country – the United States – due to the use of new on-shore technologies such as slant drilling and hydraulic fracturing. Figure 3 illustrates that after a prolonged period of decline, U.S. oil production has been on the steady rise starting in mid–2010, while the supply response from the rest of the world remained flat.

Figure 3: Global Crude Oil Production With and Without the U.S.: 1996–2014

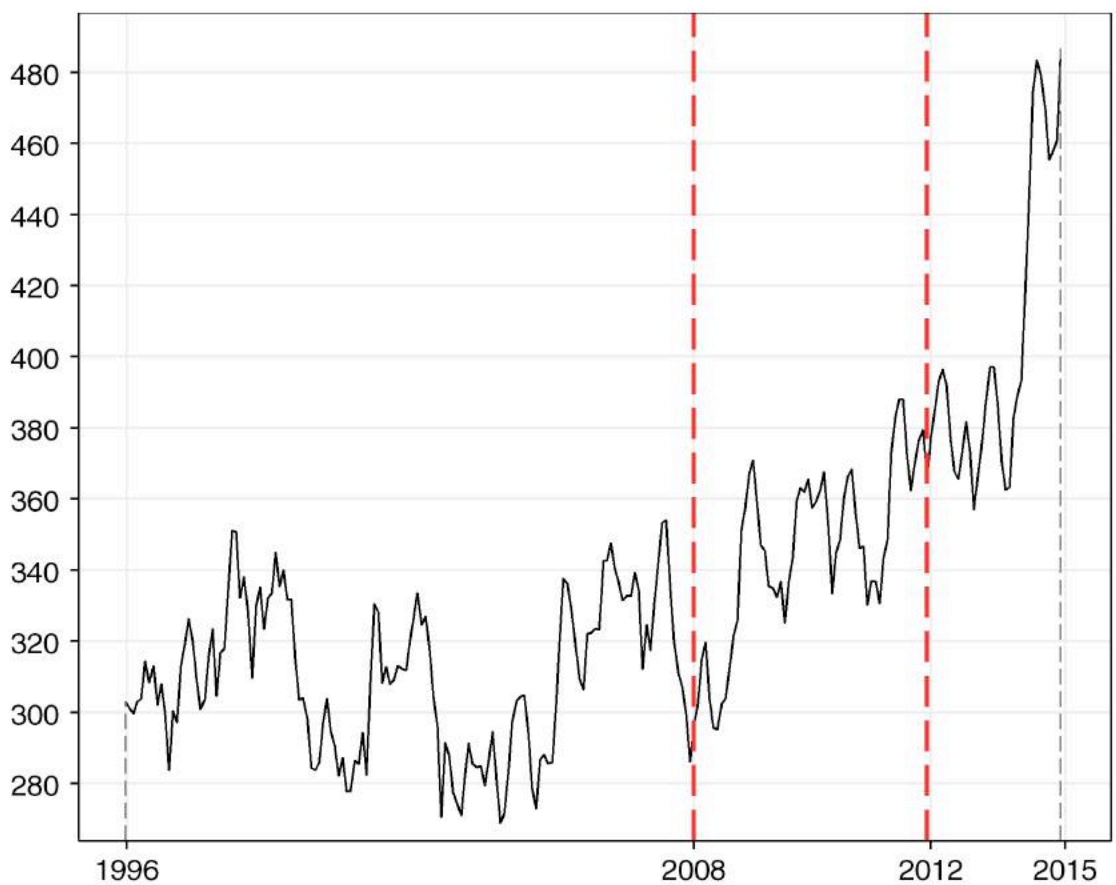


Sources: U.S. Energy Information Administration. The top line plots the total global production of crude oil in million of barrels per day averaged lever a year. The bottom line plots global production minus production of crude oil in the United States. The difference between the two lines is crude oil production in the U.S. in million barrels per day averaged over a year.

If practically the entire increase in the global oil supply can be associated with the increase in on-shore oil production in the U.S., then the change in the supply regime from constrained to unconstrained can be traced to the U.S.

Secondly, due to the export ban that has been in effect between 1975 and 2015, U.S. producers could not export any of its produced crude oil prior to 2015. Thus, while the global oil supply remained constrained, the U.S. oil supply became unconstrained. However, because the additionally produced crude oil could not be exported, it needed to be first stored in the available U.S. storage facilities and then transported to and processed by U.S. refineries for both domestic consumption and export (as there was no ban on exports of refined oil products from the U.S.). Consequently, as Figure 4 illustrates, available inland storage facilities experienced a steady increase in inventories.

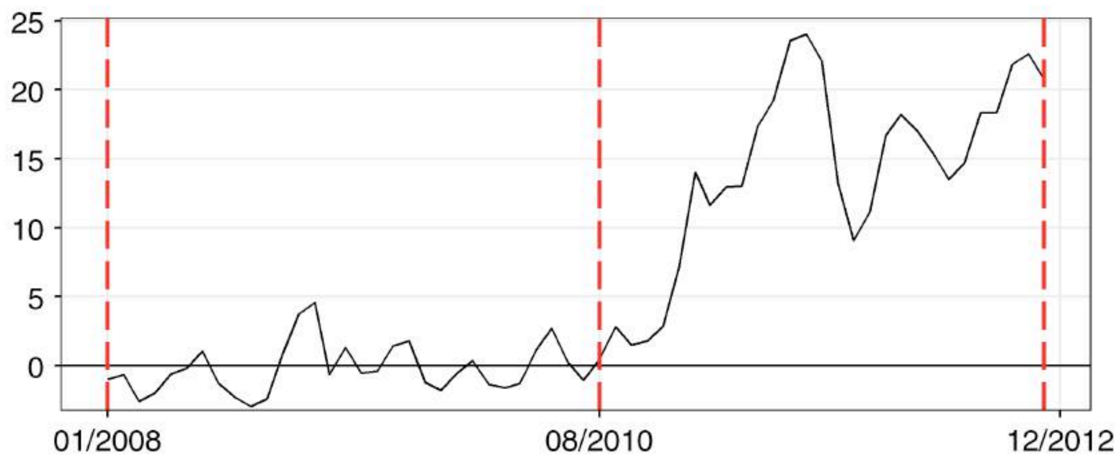
Figure 4: U.S. Commercial Crude Oil Inventories: 1996–2015



Sources: U.S. Energy Information Administration. The solid line plots U.S. end of the month commercial crude oil inventories in millions of barrels. The horizontal dashed line plots capacity utilization of U.S. on-shore storage facilities (in percent). For technological reasons, 85 percent capacity utilization means that on-shore storage facilities are full.

Thirdly, as the supply of inland crude oil in the U.S. became unconstrained while global seaborne supply remained constrained, a persistent positive spread developed between global seaborne benchmark Brent and U.S. inland benchmark, West Texas Intermediate (WTI). According to Figure 5, the Brent–WTI spread became persistently positive after August 2010.

Figure 5: Brent–WTI Spread: 2008–2012



Sources: Intercontinental Exchange and New York Merchantile Exchange. The line plots the difference between the average monthly price of the Brent nearby futures contract and the average monthly price of the WTI nearby futures contract.

At times, a barrel of global Brent crude traded for more than 20 dollars more than a barrel of U.S. inland WTI crude. Yet, in spite of the glut of inland crude oil, for a variety of technological reasons, inland crude could not be readily transported to or used in a large number of refineries in the U.S., especially along the Eastern seaboard.[3]

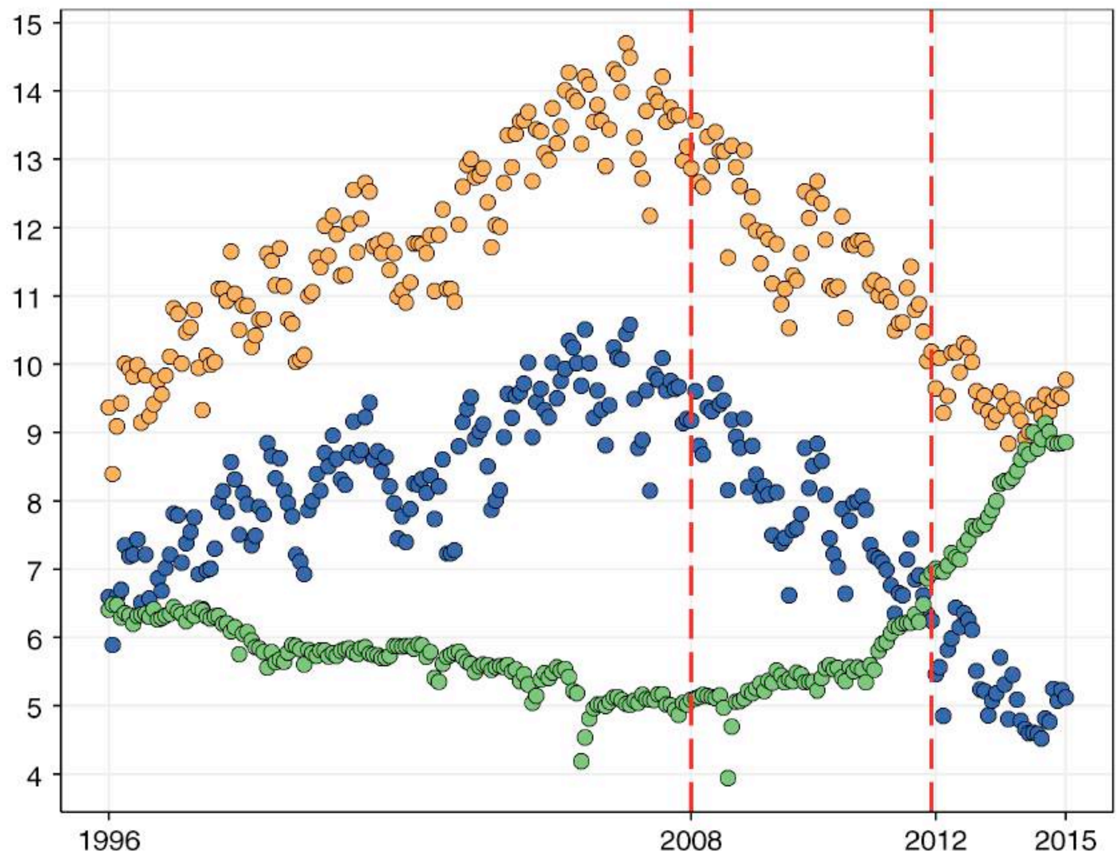
While these technological reasons were being partially worked out, imports of seaborne crude into the United States remained above U.S. domestic production of crude oil until the end of 2012. Trends in U.S. domestic production and seaborne imports into the U.S. are illustrated in Figure 6.

These observations suggest an empirical strategy that can be employed to examine the predictions of speculative storage framework under limits to arbitrage by using U.S. imports data for crude oil.

According to speculative storage theory, restricted supply should be associated with a build-up of speculative inventories and unrestricted supply should be associated with a decline in speculative inventories. Supply of crude oil in the U.S. became unrestricted

[3]We thank Vikas Raman for bringing the Brent–WTI spread as an indicator of a possible structural break in the crude oil market.

Figure 6: U.S. Crude Oil Imports and Production: 1996–2014



Sources: U.S. Energy Information Administration. The top series is average monthly crude oil imports (in million barrels per day), including Strategic Petroleum Reserve. The middle series is average monthly crude oil imports (in million barrels per day), including Strategic Petroleum Reserve, but excluding Canada and Mexico. The bottom series is average monthly crude oil production (in million barrels per day).

starting in about August 2010, while it remained restricted globally because of the ban on the export of crude oil from the U.S. Theory predicts that we should see speculative U.S. inventories increase from 2008 until about August 2010 and then decline. However, because crude oil could not be exported from the U.S. at the time, additionally produced oil was being added to the available U.S. on-shore storage facilities, quickly filling them up. As a result, total U.S. inventories were rising toward total available storage capacity from 2008. At the same time, while the U.S. was producing more domestically, it was still consuming more than it was producing which created a need to import additional crude oil primarily by sea while the term structure of futures prices stayed in contango.

It is possible then that seaborne imports into the U.S.–while on a declining trend–can contain a component that can identified as speculative floating inventory.

III. Floating Inventory Data

A. Bill of Lading

Our source data consists of 200,930 individual Bills of Lading (BOL) for all seaborne imports of crude oil and energy products into the United States during 2008–2012 made available to us by DataMyne, a data aggregator and analytics company.

A Bill of Lading (BOL) is "a document that establishes the written evidence of a contract for the carriage and delivery of goods sent by sea for a certain freight."[4] It serves as evidence of the right to entry into the U.S.[5] For illustrative purposes, a stylized Bill of Lading is presented in Figure 7.[6]

<Insert Figure 7>

Analogously to customs declarations that must be submitted by individuals entering the United States, Bills of Lading must be mandatorily submitted by all cargo carriers entering U.S. ports to the U.S. Customs and Border Protection (CBP) — the federal law enforcement agency of the United States Department of Homeland Security with the mandate to enforce laws and regulations related to the collection of customs duties and the crossing of U.S. borders. The enforcement mandate of the CBP ensures that BOL source data is accurate as counteparties are obligated to accurately report the name of the transportation company, the vessel, and the route. In addition, as a BOL is typically linked to a Letter of Credit and associated payments, counterparties are motivated to report the date of arrival in a U.S. port as close to the verified date of arrival as possible so they get paid faster.

Not all of the BOL information, however, is fully standardized. As the responsibility for filing BOL and assigning classification codes is delegated to transportation companies, product description codes and counterparty codes could vary. We utilize internal standardization protocols for company names, products and destinations implemented by DataMyne, which first collected source BOL data for each arriving vessel from the CBP under the U.S. Freedom of Information Act and then standardized them for subsequent analytics.

Furthermore, some important information is not required to be reported in a BOL. This information consists of the date of a vessel's departure and whether some or all of the

[4]Mason v. Lickbarrow, 1 H. Bl. 359.

[5]Trade Act of 2002, U.S. Customs and Border Protection

[6]Figures and tables on the data and empirical results are at the end of the paper.

cargo is being used as floating storage. As described below, we will employ statistical learning tools to classify some vessels to a speculative floating storage category. The BOLs also do not contain information on the price or value of the cargo.

To check the completeness of the granular BOL data, we calculated monthly statistics for the volume of seaborne oil imports and compared them to aggregate seaborne oil import statistics reported by the U.S. Energy Information Administration (EIA) — all imports minus imports from Canada and Mexico which have a land border with the U.S. We found that statistics calculated from the granular BOL source data nearly perfectly match statistics reported by the EIA. This is important as the EIA data is compiled from summary reports by importers and refineries, while the BOL data is compiled from reports by transportation companies for individual vessels entering U.S. ports. The difference between the two series is about 3 percent (BOL series is smaller than the EIA series) or approximately five to six days of seaborne import flow and could be associated with a time lag between dates in BOLs and custom declarations, as well as, differences in methodologies for identifying the country of origin and sea/land routes.

B. Data Transformation—Nodes

We represent information on U.S. seaborne import shipments imbedded in source BOL data as a directed dynamic acyclic graph $G_t = (V_t, E_t)$ consisting of a set of notes, V_t, and edges, E_t, at time t, from 01/012008 to 12/31/2012, where t denotes one day.

Specifically, V_t consists of shippers and receivers at time t. The entire 2008–2012 data sample allows us construct V — a finite set of size $n \in \mathbb{N}$ that represents all n companies shipping and receiving shipments during the sample period. We construct the reference set V with the unique numbered vertices as 1, 2,. . . , n such that $V_t \in V$ for each t. E_t set represents instances of Bills of Lading or transactions between shipping u_t^i and receiving v_t^j companies, where transactions, edges, are indicated by the ordered pair (u_t^i, v_t^j) with $u_t, v_t \in\ V$.

To help us better identify receiving nodes, we have further augmented information associated with each receiving node with the data about the U.S. port of arrival for each particular vessel. Thus, vertices for US buyers (consignees) v_t^j are defined by a pair of fields "Company ID" and "U.S. Port of Arrival", while vertices for seller (shipper) u_t^i are defined only by "Company ID."

For illustrative purposes, Figure 8 presents a network representation of BOL shipments during 2011.

<Insert Figure 8>

Following the initial representation, we further transform the data by adding information on ownership structure and a coarse industry classification of the owner — P for producer and T for trader.

Figure 9 presents a network representation of BOL shipments during 2011 with this additional information.

<Insert Figure 9>

In the figure, all subsidiaries of a company are represented by one node — the parent. For example, the node "Exxon Mobil" represents all 35 subsidiaries and companies working under the Exxon Mobil umbrella in the market. This transformation leads to a decrease in the number of nodes from 348 to 174 and a reduction in the number of edges from 1318 to 623. It modifies the type of a graph from directed acyclic to directed and allows for self-loops. Self-loops are shipments from a foreign to a domestic division (child) of the same parent company. Note that we keep industry classification at the level of children companies, while connections among companies at the level of parent companies. We do not use an industry attribute for parent companies due to high degree of vertical integration in the oil industry, where a parent is often a company managing upstream, downstream, logistic and/or trading children divisions (companies).

C. Data Transformation—Edges

Having transformed the nodes or companies, we now proceed to transform the edges that follow very similar patterns into a single edge using the intuitive concept of a "trading agreement." According to a "trading agreement", a buyer and a seller agree that within a calendar year, the seller will deliver a certain total quantity of a certain blend of crude oil to the buyer at a certain U.S. port with a certain periodicity. For example, during 2012, A agrees to deliver to B a total of 120 metric tons of light sweet crude oil to a specific port in the U.S. Gulf by delivering 10 metric tons of the specified crude oil every month.

We do not observe actual long-term and short-term trading agreements between importers and exporters, but we believe that the arrival of each shipment that we do observe (at time $t+T$) is associated with entering into such an agreement at some prior date t. The fields we use to learn about latent trading agreements are the unique ID for a US buyer (consignee) with US port of arrival, $u^{pi} \in V^p$, the unique ID for an international seller (shipper) $v^{pj} \in V^p$, the product (custom code group HS 2709 crude oil), the quantity (in metric tons) $\{X_1, ..., X_{n_t}\}$, and the date of arrival t.

Trading agreements that we construct do not contain prices, because BOL data does not include prices. However, each trading agreements can be further associated with over 240 different brands and blends of crude oil within the product code HS 2709. Thus, we can use available price data for each brand and blend of crude oil at each date in each location to put a price on each arrival. Then, we can calculate average price for each trading agreement and validate this information against customs data provided by the US Census.

D. Data Categories

Having transformed the data on both nodes and edges, we proceed to characterize patterns in the data for the following categories: PPEdges, PPTPPTLoops, TPPTTTEgdes, and TTLoops. PPEdges represent shipments from producers to other producers. PPTPPTLoops represent shipments from a producer to itself, from a trader owned by a parent company to a producer owned by the same parent company or from a producer owned by a parent company to a trader owned by the same parent company. TPPTTTEgdes represent shipments from a producer to a trader, from a trader to a producer or from a trader to another trader. TTLoops represent shipments from a trader to itself.

Figure 10 presents monthly shipments in millions of barrels per day.

<Insert Figure 10>

PPEdges are blue, PPTPPTLoops are orange, TPPTTTEgdes are green, and TTLoops are red.

E. Speculative Floating Oil

We define Speculative Floating Oil (SFO), an empirical proxy for Y_t, as shipments from a trader to itself: TTLoops. SFO is a fraction of crude oil and oil products imported into the U.S. by sea during 2008-2012. Figure 11 presents monthly shipments of SFO during 2008-2012 in millions of barrels per day.

<Insert Figure 11>

Table I presents summary statistics for monthly time series of SFO in levels, in differences, in logs, and in log-differences. Statistically, SFO monthly series in levels, Y_t is well approximated by a non-stationary ARIMA(1,1,1) process while the SFO series in differences, ΔY_t is well approximated by a stationary ARIMA(1,0,1) process. The process ΔY_t, is presented in Figure 12 below.

<Insert Figure 12>

Visually, Figure 12 indicates that the volatility of ΔY_t is lower after about mid-2010. In order to formally empirically test for a change in regime of the ΔY_t process, we assume that it follows an autoregressive process of order 1 and test for the instability in variance. Figure 13 shows the time series of Wald test statistic for an unknown regime change point. The the Wald test statistic peaks at September 2010. These results are also presented in Table II.

<Insert Figure 13>

The presence of the two regimes in speculative floating oil before and after September 2010 is an empirical regularity of the SFO series that we have found without conditioning on world oil prices, US production or global demand. Yet, it is worth recalling that as we have previously mentioned, starting in August 2010, a persistent positive spread began to develop between global seaborne benchmark Brent and U.S. inland benchmark, WTI just as surge in oil production in the U.S. due to new technologies became the dominant reason behind the increase in global oil supply.

IV. Empirical analysis

Our model predicts that the slope of Brent futures prices should be strongly positively related to speculative floating inventory after accounting for charter (freight) costs.

A. Data on oil futures and freight costs and timing considerations

Source data for the slope of the Brent term structure — log twelve month deferred minus the log nearby — in dollars per barrel are from the Intercontinental Exchange. To create a time series of futures prices, the roll date was set for the first day of each month. No price adjustments were made to eliminate artificial jumps in the prices of consecutive futures contracts. Brent futures contract is specified for 1000 barrels; to get to the dollars per barrel specification, we divided the series by 1000. Monthly series are computed by averaging daily data over each month.

Source data for the daily cost of chartering a vessel for the route between the Arabian Gulf and the U.S. Gulf is from the Platts Oilgram Price Report. These are monthly flat spot dirty tanker rates for a route Arab Gulf to the U.S. Gulf Coast for a 270,000 metric tons tanker. By Platts methodology, contracted as "the average rate for routes Ras Tanura - LOOP (via Quoin island, L&B via: Cape), Mina al-Ahmadi - Houston via Cape (via Quoin island), Kharg Island - Corpus Christi (via: Quoin Island, L&B via: Cape)." Monthly series are computed by averaging over weekly published source data.

Since we only observe when loaded vessels enter U.S. ports, we need to make an assumption to align the date when the oil arrives in the U.S., i.e., the time $t + T$, with the date t when the associated carry trade might have been put in. We assume that $T = 180$ days. This way, when we observe that on January 1, 2008, a vessel loaded with what we classify as speculative floating oil enters the U.S., we will statistically relate it to the data for the term structure of Brent crude oil futures prices and for the costs of chartering a vessel as of July 1, 2007, i.e., lagged six months, which is when we believe the carry trade was put in.

B. Baseline regression

To check the prediction of theory, we specify a baseline regression of the slope of the term structure of futures prices on the changes in speculative floating oil controlled for the costs of chartering a vessel.

$$L^6(F_t^{t+12} - F_t) = \alpha + \beta_1 \Delta Y_t + \beta_2 t L^6 R_t + \rho \Delta Y_{t-1} + \gamma Y_{t-1} + \epsilon_t,$$

where ΔY_t denotes changes in monthly shipments of speculative floating oil, $(F_t^{t+12} - F_t)$ is the slope of Brent term structure (log twelve month deferred futures minus log nearby) in dollars per barrel, R_t is the daily cost of chartering a vessel for the route between the Arabian Gulf and the U.S. Gulf, the relative time trend t is as a fraction of T=180 days, and L^6 denotes a six-months lag operator. The regression specification accounts for the autoregressive empirical properties of the ΔY_t process. The regression specification also accounts for the fact that ΔY_t are arithmetic rather logarithmic differences by including the previous period price Y_{t-1}.

The coefficients of the baseline regression are as follows (standard errors are given in parenthesis below the coefficients):

$$\begin{aligned} L^6(F_t^{t+12} - F_t) &= \underset{(0.002)}{-0.008} + \underset{(0.007)}{0.033} \Delta Y_t + \underset{(0.227)}{0.276} L^6(R_t) \\ &\quad \underset{(0.007)}{-0.015} \Delta Y_{t-1} + \underset{(0.007)}{0.063} Y_{t-1} + \epsilon_t \end{aligned}$$

The goodness–of–fit statistics of the regression are: Multiple $R^2 = 0.67$; Adjusted $R^2 = 0.64$; F-statistic = 26.31 on 4 and 53 DF; p-value = 0.0000. The length of the times series is 58 months. The coefficient on the slope of the speculative floating oil is positive and statistically significant at the one percent level and the coefficient on the shipping cost is positive but statistically insignificant.

C. Regression with regime change

In order to account for the regime change in the SFO series due a change in the supply regime from constrained to unconstrained, we adjust the regression specifications as follows:

$$\begin{aligned} L^6(F_t^{t+12} - F_t) &= D_t^C\{\alpha^C + \beta_1^C \Delta Y_t + \beta_2^C t L^6 R_t + \rho^C \Delta Y_{t-1} + \gamma^C Y_{t-1}\} \\ &\quad + D_t^U\{\alpha^U + \beta_1^U \Delta Y_t + \beta_2^U t L^6 R_t + \rho^U \Delta Y_{t-1} + \gamma^U Y_{t-1}\} + \epsilon_t, \end{aligned}$$

where $D_t^C = 1$ during Jan 2008 - July 2010 and zero, otherwise, and $D_t^U = 1$ during Aug 2010 - Dec 2012 and zero, otherwise. The subscripts C and U denote constrained and unconstrained supply regimes, respectively.

The coefficients of the regression with the regime change dummies are as follows (standard errors are given in parenthesis below the coefficients):

$$\begin{aligned} L^6(F_t^{t+12} - F_t) \quad = \quad & D_t^C\{\underset{(0.002)}{-0.012} + \underset{(0.007)}{0.030}\Delta Y_t + \underset{(0.392)}{1.389}tL^6R_t - \underset{(0.007)}{0.015}\Delta Y_{t-1} + \underset{(0.008)}{0.064}Y_{t-1}\} \\ & +D_t^U\{\underset{(0.004)}{0.003} + \underset{(0.016)}{0.026}\Delta Y_t - \underset{(0.402)}{1.014}tL^6R_t - \underset{(0.014)}{0.023}\Delta Y_{t-1} + \underset{(0.017)}{0.045}Y_{t-1}\} \\ & +\epsilon_t, \end{aligned}$$

The goodness–of–fit statistics of the regression are: Multiple $R^2 = 0.79$; Adjusted $R^2 = 0.75$; F-statistic = 18.44 on 10 and 48 DF; p-value = 0.0000. The length of the time series is 58 months.

As predicted by theory, the regression coefficient on speculative floating oil is positive and statistically significant at the one percent level during the constrained regime and remains positive, but is not statistically significant during the unconstrained regime. The coefficient on the cost of shipping cost is positive and statistically significant at the one percent level during the constrained regime, but turns negative and nearly statistically insignificant during the unconstrained regime.

To check for nonstationarity in individual time series we conduct an efficient ADF, DF-GLS unit root tests (see Table III). According to the tests, Y_t is mean-reverting along a trend, I(0) at 1% significance level, in the Restricted Supply regime. After the structural break, Y_t starts following a stochastic I(1) trend in the Unrestricted Supply regime. Slope, $L^6(F_t^{t+12} - F_t)$ is a I(1) process consistently. Freight cost, L^6R_t is trend stationary in the full sample, but I(1) in both subsamples.

D. Cointegration

According to theory, arbitrage activities of commodity cash-and-carry traders in the physical and futures markets impact the risk premium because both commodity carry traders and commodity producers demand a short position in the futures markets against the limited capacity of capital-constrained commodity asset managers. This introduces additional limits to hedging for commodity producing firms and manifests itself in the risk premium. Thus, the time series for the risk premium, storage costs, and floating inventory are likely to exhibit a conitegrating relationship.

We test for the presence a cointegrating relationship in the three times times series both with and without a structural break. Johansen (1991) test results for the full times series without accounting for a possible structural break are in Table IV. As shown in the table, we fail to reject the null hypotheses that the number of cointegration vectors is less than one, but reject the hypothesis the at the number of cointegration vector is less than two against the alternative that there is one cointegrating vector. However, after adjusting for the small sample size, the statistical significance of test results weakens.

We repeat the test using a modified test by Johansen et al. (2000) that accounts for known structural breaks in the cointegrating relationship. Results of the Johansen et al. (2000) test are presented in Table V. As shown in the table, the results strongly statistically confirm the presence of a single cointegration vector linking the three times series.

Fitting cointegrating vectors estimated under the two regimes gives rise to a the two fitted series for speculative floating oil. The fitted speculative floating oil series under the two regimes as plotted in Figure 14.

<Insert Figure 14>

There is a clear difference both in the level and the variance of the two fitted series, but the signs and the sizes of coefficients are as expected — positive signs on the slope of the term structure of futures prices and negative signs on the cost of shipping.

V. Conclusion

We introduce commodity cash-and-carry traders, who arbitrage between the physical and futures markets, into the limits to arbitrage extension to speculative storage framework by Acharya, Lochstoer and Ramadorai (2013). Carry traders are assumed to have access to off-shore costly storage technology — shipping vessels that can be used to store physical commodity in one period and deliver it in the subsequent period. We show that in equilibrium, arbitrage activities of commodity carry traders are associated with additional limits to hedging for commodity producing firms and affect equilibrium commodity prices.

We test empirical predictions of the model using a novel data set with granular information on every tanker that delivered seaborne crude oil into the United States during 2008-2012. As predicted by the model, we find that the slope of Brent futures prices is strongly positively related to speculative floating inventory, which we derive from the total seaborne crude imported into the U.S. after taking into account the costs of using vessels for use as speculative storage.

In addition, consistent with extensions to the canonical speculative storage framework by Dvir and Rogoff (2009, 2014), speculative floating inventory imported into the U.S. increased during January 2008–August 2010 when the global supply response was constrained and then decreased to zero during the second half of 2010 and into 2011 when the U.S. was increasing its domestic production of crude oil even though the term structure of (globally determined) Brent futures prices remained in contango.

References

Acharya, Viral V., Lochstoer, Lars A., and Tarun Ramadorai, 2013, Limits to arbitrage and hedging: Evidence from commodity markets, *Journal of Financial Economics* 109 (2), 441-465.

Ahn, S., and G. Reinsel, 1990, Estimation for partially nonstationary autoregressive models, *Journal of the American Statistical Association* 85, 813-823.

Atkins, Eric, 2016, Why oil speculators are turning to ships as floating storage, *The Globe and Mail*. Available online at https://www.theglobeandmail.com/report-on-business/industry-news/energy-and-resources/why-oil-speculators-are-turning-to-ships-as-floating-storage/article28846311/

Berk, Jonathan, and Richard Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269-1295.

Deaton, Angus, and Guy Laroque, 1992, Behaviour of commodity prices, *Review of Economic Studies* 59 (1), 1–23.

Deaton, Angus, and Guy Laroque, 1996, Competitive storage and commodity price dynamics. *Journal of Political Economy* 104 (5), 896–923.

Deaton, Angus, and Guy Laroque, 2003, A model of commodity prices after Sir Arthur Lewis. *Journal of Development Economics* 71 (2), 289-310.

Dvir, Eyal, and Kenneth Rogoff, 2009, Three Epochs of Oil, *NBER Working Paper* No. 14927.

Dvir, Eyal, and Kenneth Rogoff, 2014, Demand effects and speculation in oil markets: Theory and evidence, *Journal of International Money and Finance* 42, 113–128.

Frankel, Jeffrey, 2014, Effects of speculation and interest rates in a "carry trade" model of commodity prices, *Journal of International Money and Finance* 42, 88-112.

Hamilton, James, 2009, Understanding crude oil prices, *Energy Journal* 30 (2), 179–206.

Johansen, Soren, 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica* 59 (6), 1551–1580.

Johansen, Soren, Mosconi, Rocco, and Bent Nielsen, 2000, Cointegration analysis in the presence of structural breaks in the deterministic trend, *Econometrics Journal* 3, 216–249.

Kilian, Lutz and Thomas Lee, 2014,Quantifying the speculative component in the real price of oil: The role of global oil inventories, *Journal of International Money and Finance* 42, 71-87.

Kilian, Lutz, and Daniel Murphy, 2014, The role of inventories and speculative trading in the global market for crude oil, *Journal of Applied Econometrics* 29 (3), 454-478.

Knittel, Christopher, and Robert Pindyck, 2013, The simple economics of commodity price speculation, MIT mimeo.

Routledge, Bryan, Seppi, Duane, and Chester Spatt, 2000, Equilibrium forward curves for commodities, *Journal of Finance* 55 (3), 1297-1338.

Shleifer, Andrei, and Robert Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35-55.

Table I: Speculative Floating Oil: January 2008 - December 2012

	Y_t	$ln(Y_t)$	ΔY_t	$\Delta ln(Y_t)$
Mean	0.14	-2.18	0.00	0.00
STD	0.09	0.73	0.07	0.65
Skewness	0.83	-0.75	0.51	0.43
Kurtosis	3.23	3.64	4.00	3.66
ADF probability	0.35	0.40	0.01	0.01
ρ_1	0.68	0.60	-0.27	-0.35
	(0)	(0)	(0.03)	(0.01)
ρ_3	0.46	0.51	-0.08	0.09
	(0)	(0)	(0.11)	(0.02)
ρ_6	0.30	0.35	-0.06	0.11
	(0)	(0)	(0.34)	(0.02)
ρ_9	0.26	0.35	0.03	0.17
	(0)	(0)	(0.6)	(0.02)
ρ_{12}	0.11	0.17	-0.06	-0.04
	(0)	(0)	(0.76)	(0.06)

This table presents summary statistics for the monthly time series of Speculative Floating Oil, Y_t (in mln barrels per day). ADF probability refers to the p-value of the ADF test for the null of unit root with two lags used for error term correction. ρ_τ is autocorrelation ρ at lag τ. Values in brackets below autocorrelation coefficients refer to p-values of the Portmanteau Q-test for no serial correlation at 1,3,6,9,12 lags, Ljung & Box (1978).

Table II: Tests for Constancy of Autoregression Parameters and Error Variance

	$\Delta Y_t = \alpha + \rho \Delta Y_{t-1} + e_t$			
	AR(1) parameter, ρ	Intercept, α	Joint, ρ and α	Variance, σ^2
supW	5.8	1.31	8.49	10.61**
expW	1.08	0.11	2.07	3.72***

This table presents test results for a change in regime of the ΔY_t process assuming it follows an autoregressive process of order 1 and test for the instability in variance of the residuals. Variance of the residuals is calculated as a sum of squared residuals divided by a sample size, n, $\frac{1}{n}\sum e_t^2$. We used the Quandt Likelihood Ratio (QLR) (Quandt, 1960) a.k.a. maximum Wald statistic (*supW*) and the logarithm of Andrews and Ploberger exponential Wald statistic (*expW*). The tests check for structural breaks under the assumption of unknown break date. The *supW* test statistic is the largest value of all the sequence of Wald F-statistic calculated for each candidate breakdate and the *expW* test statistic is the exponential transformation of the F-statistic. We use restricted time interval for candidate breakdates, $\Pi = [.15n, .85n]$, where n denotes the length of the sample size as suggested in Andrews (1993). Significance: '***' 0.001, '**' 0.01, '*' 0.05.

Table III: DF-GLS Unit Root Tests

	Full Sample, Jan 2008 - Dec 2012		Restricted Supply, Jan 2008 - Aug 2010		Unrestricted Supply, Sept 2010 - Dec 2012	
	Constant, trend	Constant	Constant, trend	Constant	Constant, trend	Constant
$L^6(F_t^{t+12} - F_t)$	-1.57***	-1.37***	-1.47**	-1.37**	-1.42***	-1.25***
$L^6 R_t$	-2.52*	-2.17*	-2.18	-1.68	-2.3***	-2.03***
Y_t	-2.43***	-2.08***	-2.63**	-2.03**	-1.81***	-1.45***
ΔY_t		-5.29***		-4.82***		-5.14***

This table presents test results for nonstationarity in individual time series using efficient ADF, DF-GLS unit root tests. Y_t denotes monthly speculative floating oil. $L^6(F_t^{t+12} - F_t)$ denotes six month lag of the slope of the terms structure of Brent futures prices. $L^6 R_t$ denotes lagged freight costs. Significance: '***' 0.001, '**' 0.01, '*' 0.05.

Table IV: Johansen (1991) Tests for the Existence of Cointegrating Vectors: January 2008 - December 2012

	Restricted Supply, Jan 2008 - Jul 2010			Unrestricted Supply, Aug 2010 - Dec 2012		
Cointegrating Rank, r	0	1	2	0	1	2
Trace Statistics	33.78***	16.51***	6.36	54.77***	16.42***	4.71
Trace Statistics, adjusted	27.24**	13.31	5.13	44.17***	13.24**	3.8
10 % Critical Value	39.06	22.76	10.49	39.06	22.76	10.49
5 % Critical Value	42.44	25.32	12.25	42.44	25.32	12.25
1 % Critical Value	48.45	30.45	16.26	48.45	30.45	16.26
Number of observations	31			29		
Lags	2			2		

This table presents test results for the existence of cointegration vectors using Johansen (1991) trace test. Adjusted trace statistics adjust for small sample size as suggested in Ahn and Reinsel (1990). P-values are calculated according to the approximation method proposed by Doornik (1998). Significance: '***' 0.001, '**' 0.01, '*' 0.05.

Table V: Johansen et al. (2000) Modified Trace Test for Cointegration in the Presence of a Single Known Structural Break: January 2008 - December 2012

	Cointegrating Rank		
	0	1	2
Trace Statistics	72.41***	34.00***	13.18*
Trace Statistics, adjusted	63.17***	30.60***	11.86
10 % Critical Value	55.46	34.46	16.77
5 % Critical Value	59.09	37.42	18.90
1 % Critical Value	66.32	43.41	23.35

This table presents test results for the existence of cointegration vectors using Johansen et al. (2000) modified trace test in the presence of a single known structural break. Adjusted trace statistics adjust for small sample size as suggested in Ahn and Reinsel (1990). P-values of adjusted statistics are calculated according to the approximation method proposed by Doornik (1998). Asymptotic critical values are from Giles and Godwin (1991) for one breakpoint and its relative location in the sample. Significance: '***' 0.001, '**' 0.01, '*' 0.05.

Figure 7: A Bill of Lading

VICS Standard BOL: WWW.VICS.ORG For Complete VICS BOL Guideline Information

Date: 02/01/1999

BILL OF LADING

Page 1

SHIP FROM

Name: ABC Company
Address: 1000 ABC Drive
City/State/Zip: Any City, AB, 10000
SID#:
FOB: ☐

Bill of Lading Number: _06141411234567890_

SHIP TO

Name: XYZ Company
Location #: _0669_
Address: 9000 XYZ Drive
City/State/Zip: Some City, ZY 90000
CID#:
FOB: ☐

CARRIER NAME: _LTL Transportation_
Trailer number:
Seal number(s):

SCAC: ABCD
Pro number: 12345678901234567890

THIRD PARTY FREIGHT CHARGES BILL TO:

Name:
Address:
City/State/Zip:

Freight Charge Terms: *(freight charges are prepaid unless marked otherwise)*

Prepaid ______ Collect _X_ 3rd Party ______

SPECIAL INSTRUCTIONS:

☐ (check box)

☐ (check box) Master Bill of Lading: with attached underlying Bills of Lading

CUSTOMER ORDER INFORMATION

CUSTOMER ORDER NUMBER		WEIGHT	PALLET	ADDITIONAL SHIPPER INFO
45012345698	350 ctns	1750 lbs	Y	
6805673	50 ctns	250 lbs	Y	
GRAND TOTAL	400 ctns	2000 lbs		

CARRIER INFORMATION

HANDLING UNIT		PACKAGE				COMMODITY DESCRIPTION	LTL ONLY	
QTY	TYPE	QTY	TYPE	WEIGHT	H.M. (X)		NMFC #	CLASS
5	plts	100	ctns	500 lbs		Sport Accessories	154865 00	70
		250	ctns	1250 lbs		Video, Tape Recording	168955 03	92.5
		50	ctns	250 lbs		Recordings, Sound, Disc, Tape	168945 01	100
5		400		2000 lbs		GRAND TOTAL		

Where the rate is dependent on value, shippers are required to state specifically in writing the agreed or declared value of the property as follows:
"The agreed or declared value of the property is specifically stated by the shipper to be not exceeding
________________ per ________________"

COD Amount: $ ________________
Fee Terms: Collect: ☐ Prepaid: ☐
Customer check acceptable: ☐

NOTE Liability Limitation for loss or damage in this shipment may be applicable. See 49 U.S.C. § 14706(c)(1)(A) and (B).

The carrier shall not make delivery of this shipment without payment of freight and all other lawful charges.

________________ Shipper Signature

SHIPPER SIGNATURE / DATE

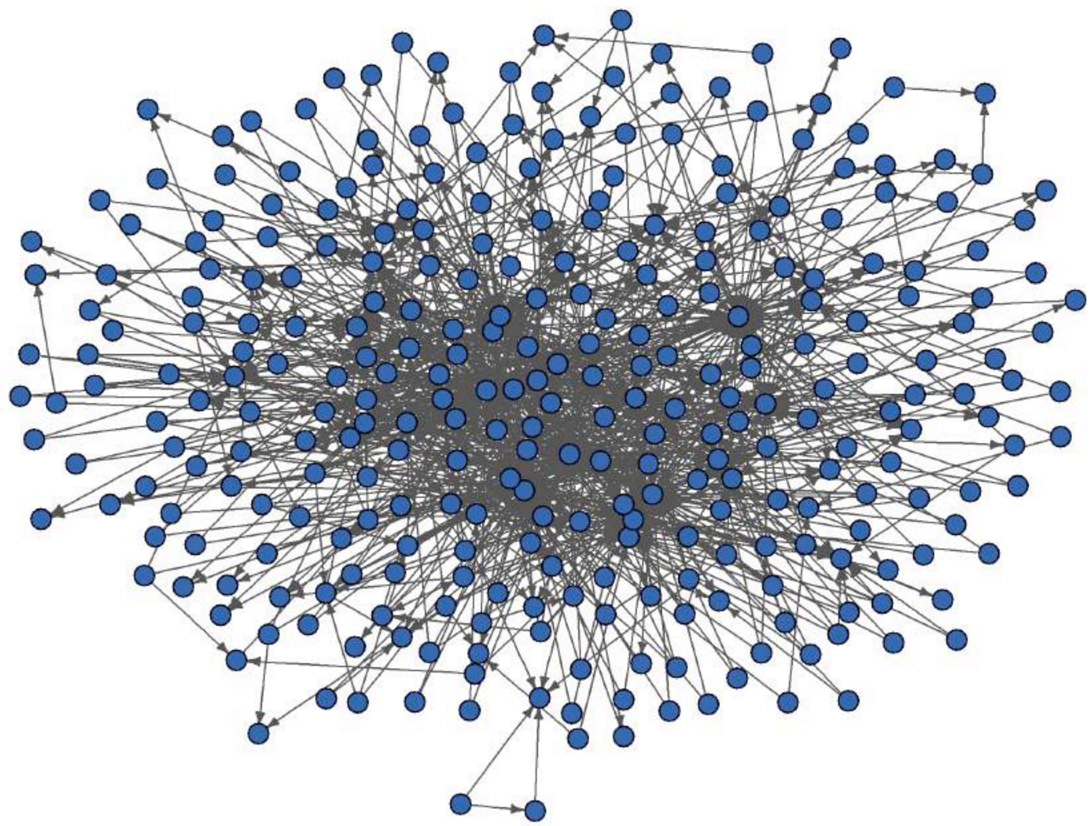
Trailer Loaded:
☒ By Shipper
☐ By Driver

Freight Counted:
☒ By Shipper
☐ By Driver/pallets said to contain
☐ By Driver/Pieces

CARRIER SIGNATURE / PICKUP DATE

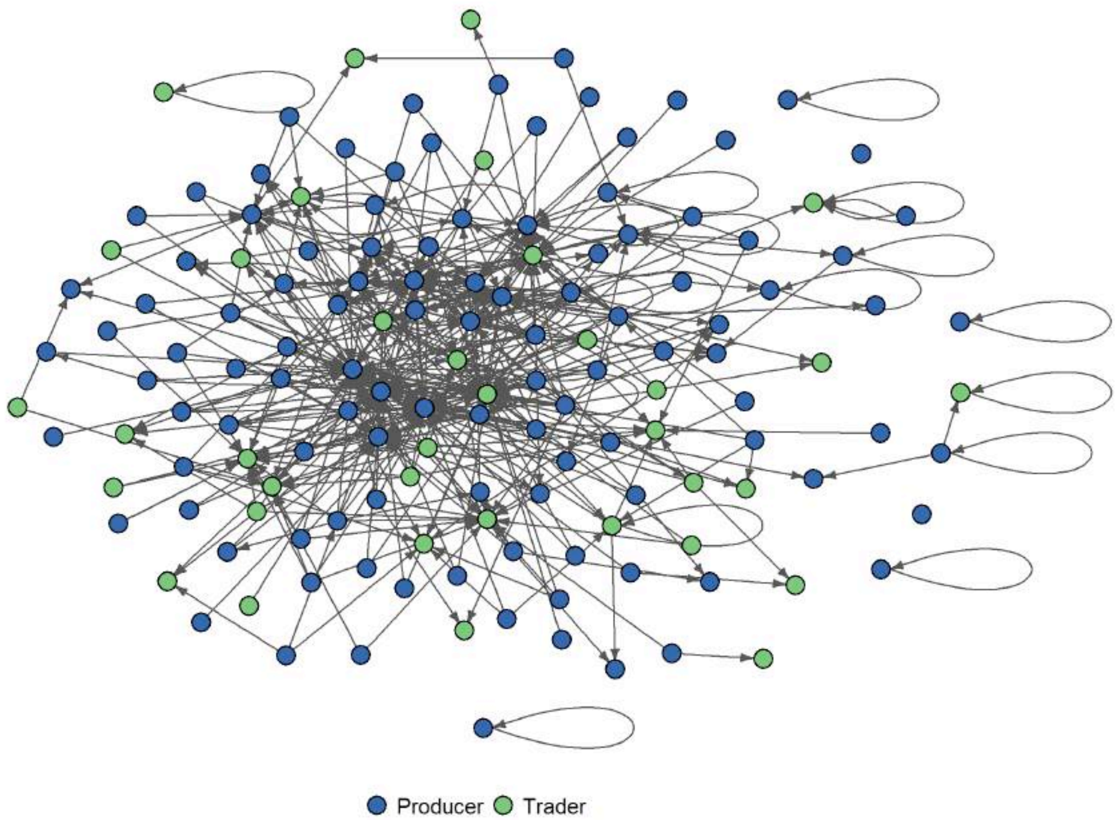
Property described above is received in good order, except as noted.

Figure 8: A Network Representation of All BOL Shipments During 2011



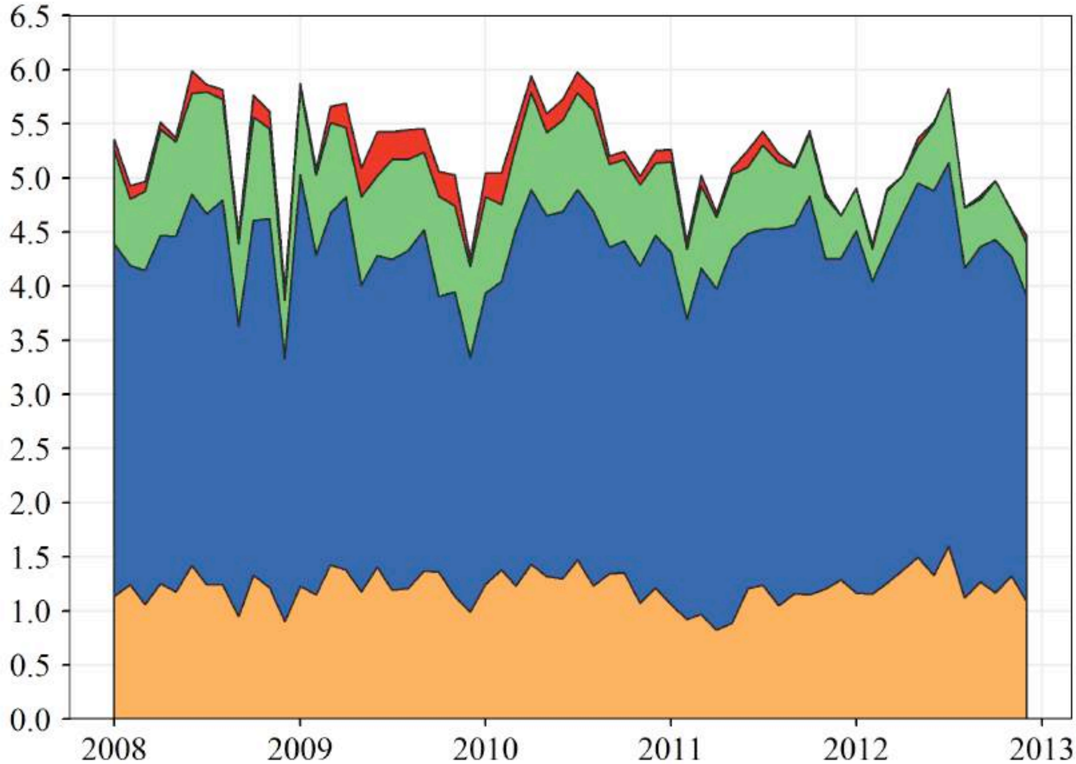
Sources: Calculations of the authors. Vertices represent companies and directed edges represent individual shipments from overseas companies to receivers in the U.S. The graph layout uses a force-based algorithm proposed by Fruchterman and Reingold (purely for aesthetics) "to position the nodes of a graph so that all the edges are of more or less equal length and there are as few crossing edges as possible." (See, "Modern Advances in Intelligent Systems and Tools" by Wei Ding, He Jiang, Moonis Ali, Mingchu Li, 2012).

Figure 9: A Network Representation of All BOL Shipments During 2011 With Ownership Information and Industry Classification of Companies



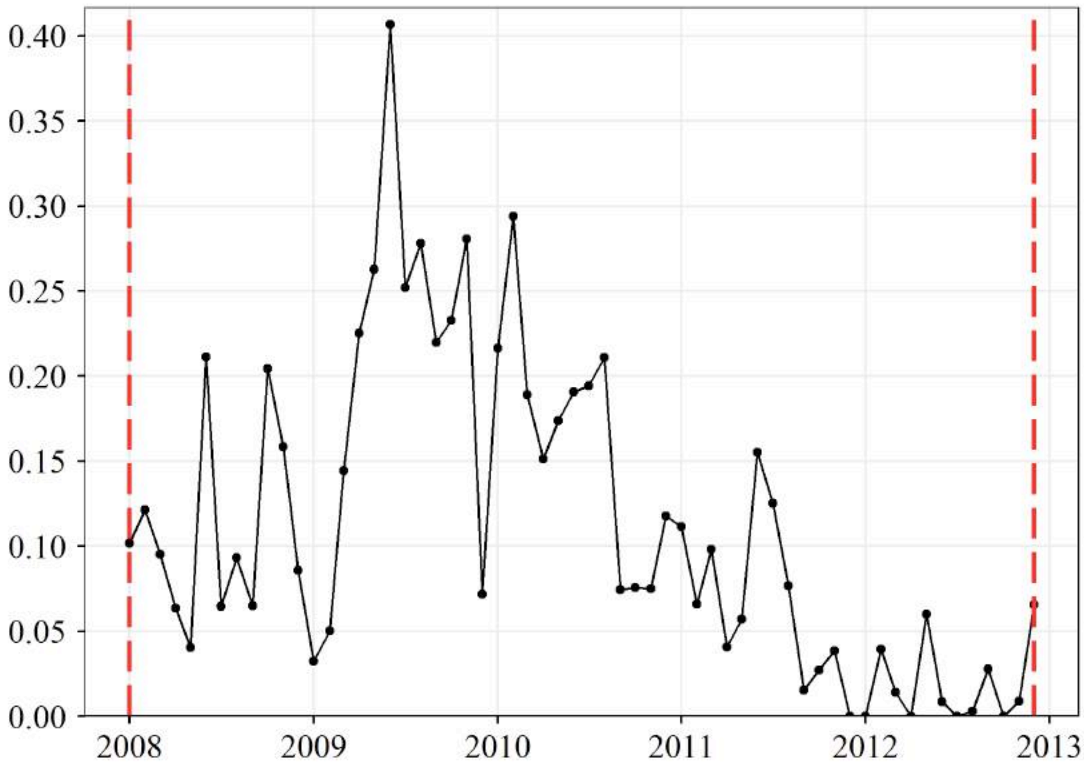
Sources: Calculations of the authors. Vertices represent parent companies and directed edges represent individual shipments from overseas companies to receivers in the U.S. Blue vertices denote producers and green vertices denote traders. Loops represent shipments from one company owned by a parent company to another company owned by the same parent company.

Figure 10: Shipments Using Transformed Data



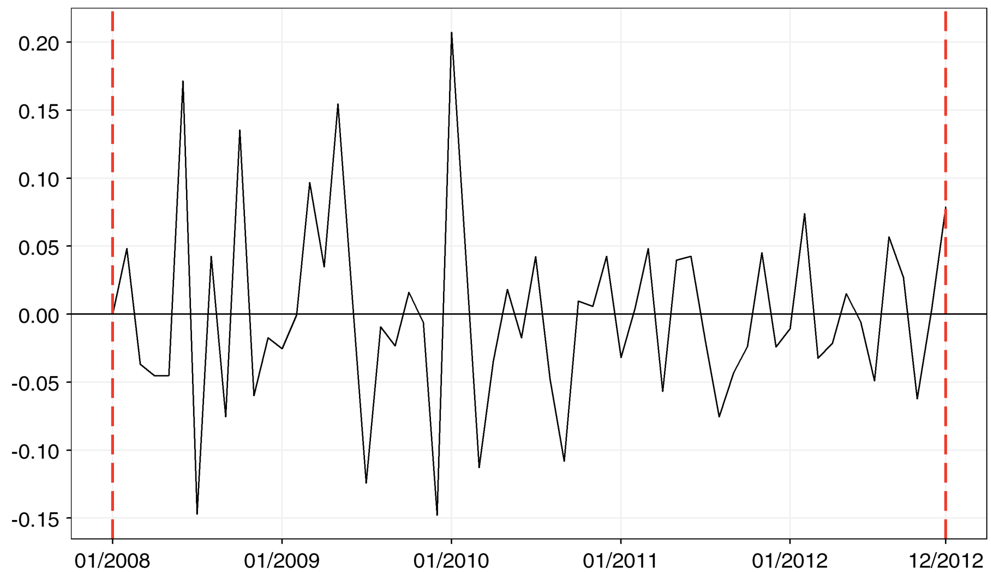
Sources: Calculations of the authors. Monthly shipments in millions of barrels per day. PPEdges are blue, PPTPPPTLoops are orange, TPPTTTEgdes are green, and TTLoops are red.

Figure 11: Speculative Floating Oil: 2008-2012



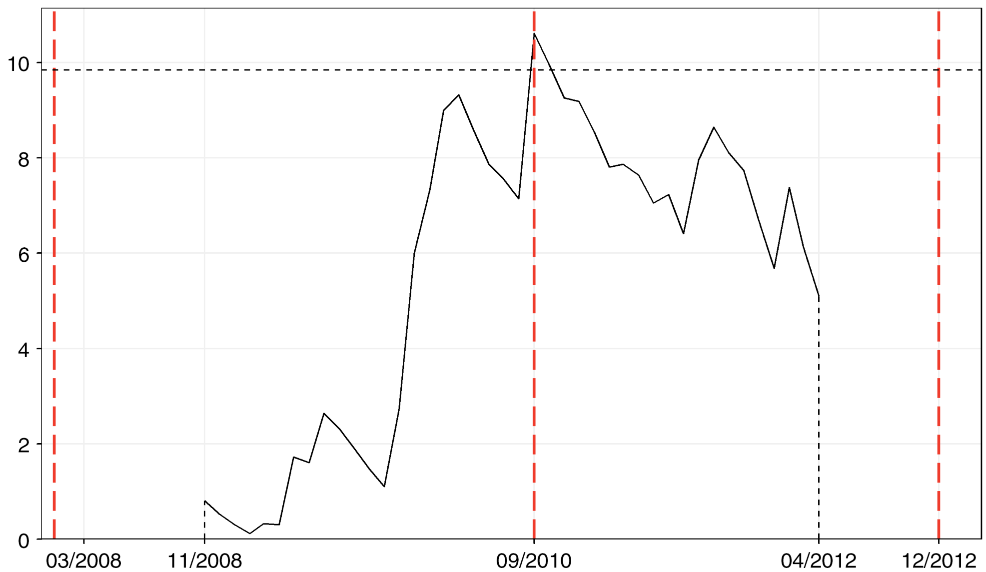
Sources: Calculations of the authors. Monthly shipments in millions of barrels per day.

Figure 12: Speculative Floating Oil in differences, ΔY_t: January 2008 - December 2012



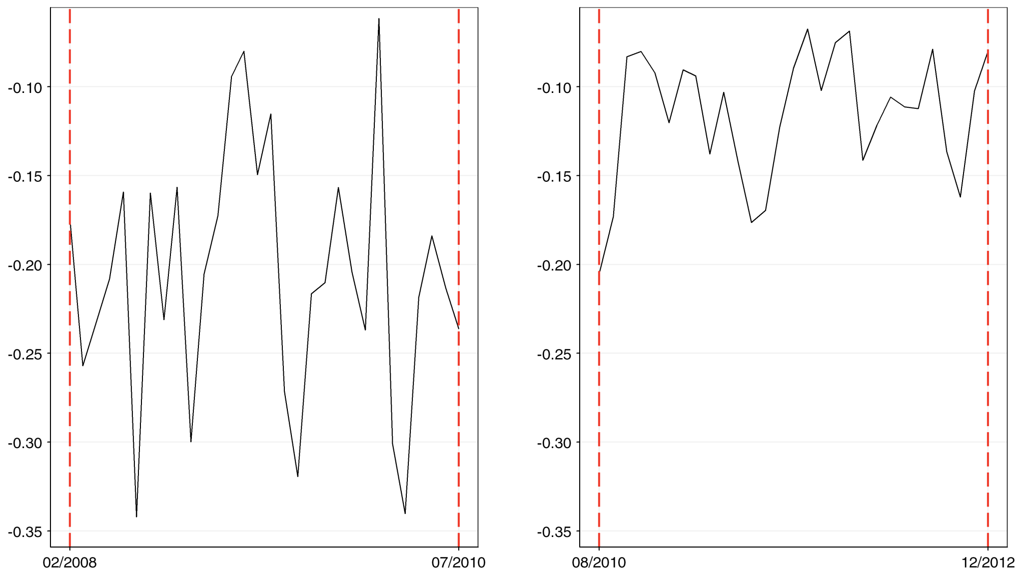
Sources: Calculations of the authors. Changes in monthly shipments, $\Delta Y_t = Y_t - Y_{t-1}$, in millions of barrels per day.

Figure 13: Wald Sequence for Variance: November 2008 - April 2012



Sources: Calculations of the authors. The plot presents time series of Wald statistics, $W(\pi)$, as a function of a single break date, T_1. The specification is $\Delta Y_t = \alpha + \rho \Delta Y_{t-1} + e_t$. The break dates, T_1, are on the x-axis and $W(\pi)$ is on the y-axis. Dotted horizontal black line shows asymptotic critical value at 5%. Restricted time interval of candidate break dates, $\Pi = [.15n, .85n]$, is used as suggested in Andrews (1993). The $Sup_{\pi \in \Pi} W(\pi)$ value of the test is 10.61, where $\pi = \frac{T_1}{n}$, $k = 1$, 5% asymptotic critical value = 9.84, Asymptotic p-value=0.031.

Figure 14: Cointegrating vector $\{\gamma Y_{t-1},\ \beta_1 L^6(F_t^{t+12} - F_t),\ \beta_2 L^6 R_t\}$ in Restricted and Unrestricted Supply Regimes



Sources: Calculation by the authors. Cointegrating vector in restricted supply regime is $-1.12Y_{t-1} + 13.53L^6(F_t^{t+12} - F_t) - 15.93L^6 R_t$; in unrestricted supply regime it is $-0.84Y_{t-1} + 4.06L^6(F_t^{t+12} - F_t) - 5.13L^6 R_t$.