THE OXFORD INSTITUTE FOR ENERGY STUDIES

January 2021

A Lesson from Physics on Oil Prices:

Revisiting the Negative WTI Price Episode

OIES Energy Comment

Ilia Bouchouev, OIES Research Associate and Adjunct Professor at New York University, Courant Institute of Mathematical Sciences

Introduction

The historic event of negative WTI prices on April 20, 2020 was not simply a one-off aberration in futures. The episode was emblematic of the modern world of oil trading, a natural culmination of the steady trend in oil financialization. It brilliantly demonstrated how intertwined physical and financial markets are, and how they feed off each other.

While the media and regulators exclusively blame the lack of storage capacit for sending WTI prices negative, many market practitioners remain highly skeptical. Even though the storage tightness was indeed the catalyst leading to the event, the market knows that the cost of storage as implied by the front futures spread, cannot change by nearly $60 per barrel and back within few hours. We look at the event from a different perspective and show that oil commentators can learn a great deal from physics, the science that has been dealing with such extreme events for centuries.

In physics, when something goes to infinity in a single instant as the result of a shock to the system and subsequently diffuses with time, it is described by a so-called impulse function. In trading, conceptually the same phenomenon is colloquially known as the squeeze[1].

The Theory of Storage, and the Model of the Squeeze

The methodology proposed in this paper is inspired by challenges to convincingly apply the classical economic theory, the so-called canonical storage problem, to the oil market. The storage theory originated in agricultural markets with a pioneering work of Gustafson and was subsequently extended in numerous articles and applied across various markets[2]. The initial work was motivated by governments' interests in using storage as an alternative price stabilization mechanism for various agricultural commodities. By and large, the canonical storage theory was trying to solve Robinson Crusoe's problem: to optimally allocate the finite supply of food between today's consumption and storing it for tomorrow given the uncertainty about replenishing the supplies tomorrow.

In the language of economics, the problem was recast as the problem of finding an equilibrium where both the price and quantity of the product are determined jointly as the result of supply and demand. The commodity price is theoretically backed out from inverting the so-called net demand function. This problem becomes much harder when the commodity can be stored. Since storage shifts supplies across multiple time periods, the same equilibrium relationship between price and quantity must hold at any time, which makes the problem dynamic. In theory, the price today depends on the supply and demand today, and the price tomorrow depends on the supply and demand tomorrow. At the same time, the price today is linked to the price tomorrow via the cost of storage; and the supply today is connected to the supply tomorrow by the decision of how much to store and shift from today to tomorrow. The problem becomes circular and it can only be solved using sophisticated numerical techniques with several monographs and many articles devoted to this subject[3].

The solution to the canonical storage problem hinges on two important assumptions. The first assumption is the ability to execute the storage carry trade, where one buys the unit of a physical commodity at the spot price and sells the futures at the higher price to lock in a guaranteed profit, provided that the spread covers the cost of storage. The trade works only if inventories do not fall to zero, in which case storage becomes impossible and the relationship between spot and futures prices breaks down. Without storage, the price must be determined by forces of supply and demand, and it

---

[1] The word squeeze is chosen entirely for pedagogical preferences and does not imply any impropriety of trading by any market participant.

[2] R. L. Gustafson (1958), Carryover Level for Grains, U.S. Department of Agriculture, Technical Bulletin No. 1178. The model was applied to oil markets by E. Dvir, and K. Rogoff (2014), Demand effects and speculation in oil markets: Theory and evidence, *Journal of International Money and Finance,* vol. 42, pp. 113-128.

[3] See J. Williams and B. Wright (1991), Storage and Commodity Markets, Cambridge University Press, and C. Pirrong (2012), Commodity Price Dynamics: A Structural Approach, Cambridge University Press, and references therein.

can rise without any bound if the demand exceeds the supply, and neither can quickly adjust. In market terms, such scenario is known as a stock-out, or an extreme backwardation.

Transferring this theory from agricultural markets to oil storage is easier said than done. The first assumption about the carry trade generally holds well, but the carry arbitrage trade must instead be applied to the spread between two futures with different delivery dates, because physical barrels always trade for forward rather than instantaneous delivery. Furthermore, in agricultural markets, storage is cheap and nearly unlimited, so storage costs are typically assumed to be constant. For oil, on the other hand, such assumptions are clearly too restrictive. The spread of oil futures can only be determined by the cost of carry if not only the inventory of oil is available, but also if there is sufficient capacity to store the inventory. The latter scenario is the mirror image of the former, which could result in the futures spread weakening without limit when the storage capacity is full, leading to the scenario of super contango.

The primary challenge, however, lies with the second assumption of the traditional storage theory about invertibility of the net demand function for price. Short-term price elasticities of oil demand and oil supply are extremely low. While there is ongoing and interesting debate on this subject, especially given the shorter cycle of shale production[4], the consensus is that price elasticities remain low as consumers simply cannot change their consumption habits quickly enough in response to prices, and producers cannot adjust their production overnight either. In addition, as global incomes grow, the energy costs represent a smaller portion of the overall expenditure making consumers less willing to adjust their behavior in response to prices. For many regional and landlocked markets, elasticities are even smaller, in fact, close to zero, because fundamental response to prices is often further constrained by the rigidity of the operational infrastructure with both the demand and the supply becoming practically inelastic on the short time scale relevant for traders.

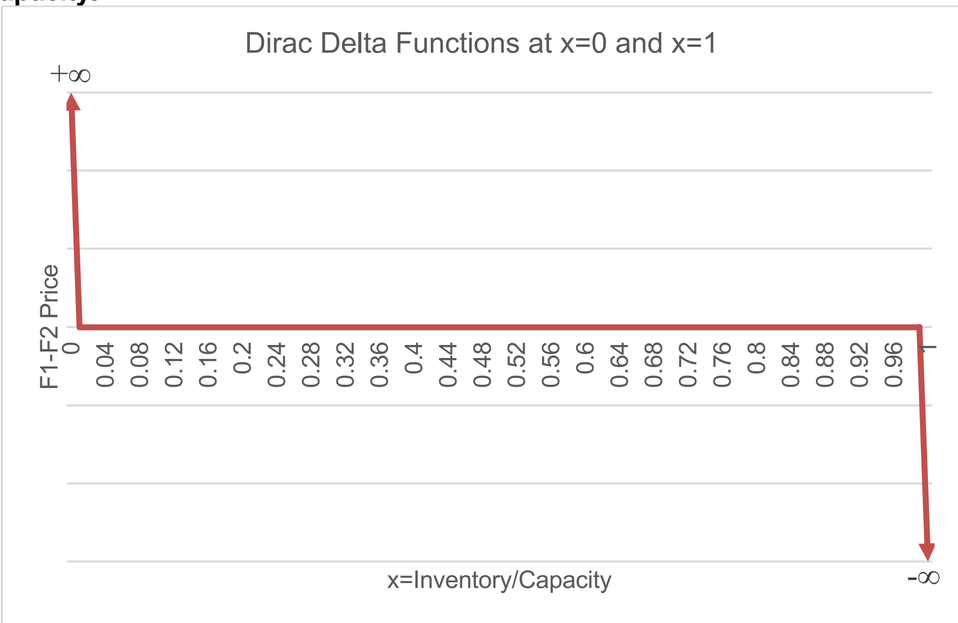
When elasticities drop to zero, the inverse demand function becomes vertical and the corresponding price can no longer be backed out from demand. The solution to the canonical storage problem, which typically describes price dependency on the total supply, effectively degenerates as the price goes to infinity in the case of zero inventories but remains unchanged with respect to all other states of inventories unless the maximum storage capacity is reached. The case of maximum storage capacity, or the so-called tank tops scenario, is usually not considered by the traditional theory, but one can attempt to incorporate it by exogenously imposing non-linear storage costs. In this case, the price could theoretically go to a negative infinity when the storage is full, like a mirror image of the vertical inverse demand function when inventory is zero. The solution to the problem will then resemble the function shown in Figure 1.

It shows that the futures spread goes to infinity in the case of zero inventories and falls to a negative infinity in the case of zero storage capacity. Otherwise, when both inventories and storage capacity are available the spread is constant equal to the normal cost of storage. We can think about this function as representing two events of default by the trader who attempts to profit from the carry arbitrage but fails to make or take the delivery. In one case, the trader sells the spread when the market is backward but fails to deliver physical barrels at expiry as barrels are not available. In the second case, the trader buys the spread at steep contango but fails to take delivery of the barrels as there is no space to store them.

The function of this type is the cornerstone of science, frequently used to model the system behavior resulting from the sharp instantaneous shock, the unit impulse applied at a single point in time and space. It is known as Dirac delta function, which is equal to infinity at the point of application and zero everywhere else with some additional normalization property. Mathematically, it is not even a function in the traditional sense, and it is typically understood as a limiting sequence of normal probability densities whose variance progressively shrink to zero.

[4] The latest summary of various estimation techniques is provided by L. Kilian (2020), Understanding Estimation of Oil Demand and Oil Supply Elasticities, *Federal Reserve Bank of Dallas.*

Figure 1: The solution to the storage problem for inelastic supply and demand resembles two Dirac delta functions for futures prices which correspond to zero inventories and zero available storage capacity.



We can now borrow from physics and restate the problem of storage and valuing futures and inventories in simpler and more practical terms. Instead of making any assumptions about unobservable supply and demand functions, we skip this step entirely, and model the futures spread directly as a function of inventories. We assume that inventories follow mean-reverting process converging towards some normal inventory level as supply and demand are forced to rebalance by the presence of hard boundaries at zero inventories and zero storage capacity. The futures spread behaves as a financial derivative of the state variable, which is the inventory. The standard analytical machinery of derivatives pricing can then apply with prices determined by the famous Black-Scholes equation with the pair of Dirac delta functions from Figure 1 acting as the boundary condition and representing the risk of default by the futures trader.

The solution, which propagates the risk of potential default at expiry of futures back to the present time, is given in the Appendix[5]. In this framework, the value of the cost-adjusted futures spread is determined by the difference between two probability density functions, one for the upside squeeze at zero inventories, and the other one for the downside squeeze at zero storage capacity. If inventories are in the normal range, then the futures spread is expected to converge to the cost of storage. However, if inventories are within reach of the two danger zones, then the spread is driven by the likelihood of approaching the extremes and risks of financial traders being squeezed as they lack capabilities to make or take physical delivery. Note that Dirac delta function in physics represents a largely unattainable state, and the same is true in trading. The actual event of default will never happen as futures contracts will be liquidated by the clearing broker, but the price forced by such liquidation may not have any bound.

The Case Study of WTI and Cushing Inventories

In contrast to economic models of supply and demand equilibrium, the primary advantage of our simple framework is in its practicality. Instead of exogenously imposing the structure of supply and demand,

[5] The details are given in I. Bouchouev (May 2020), A Stylized Model of The Oil Squeeze, New York University, working paper.

we work directly with observable data on inventories. The recent episode of negative oil prices provides us with a good opportunity for an important case study.

We calibrate the pricing formula for futures spreads to weekly data published by the Energy Information Administration (EIA). The data is scaled by the working capacity which is also available from EIA since 2011. Such normalization is essential for the behavior of prices around the extreme upper boundary. In other words, we use storage utilization capacity as our core fundamental metric.

Figure 2: WTI spread model calibration to Cushing inventories (2011-2020)

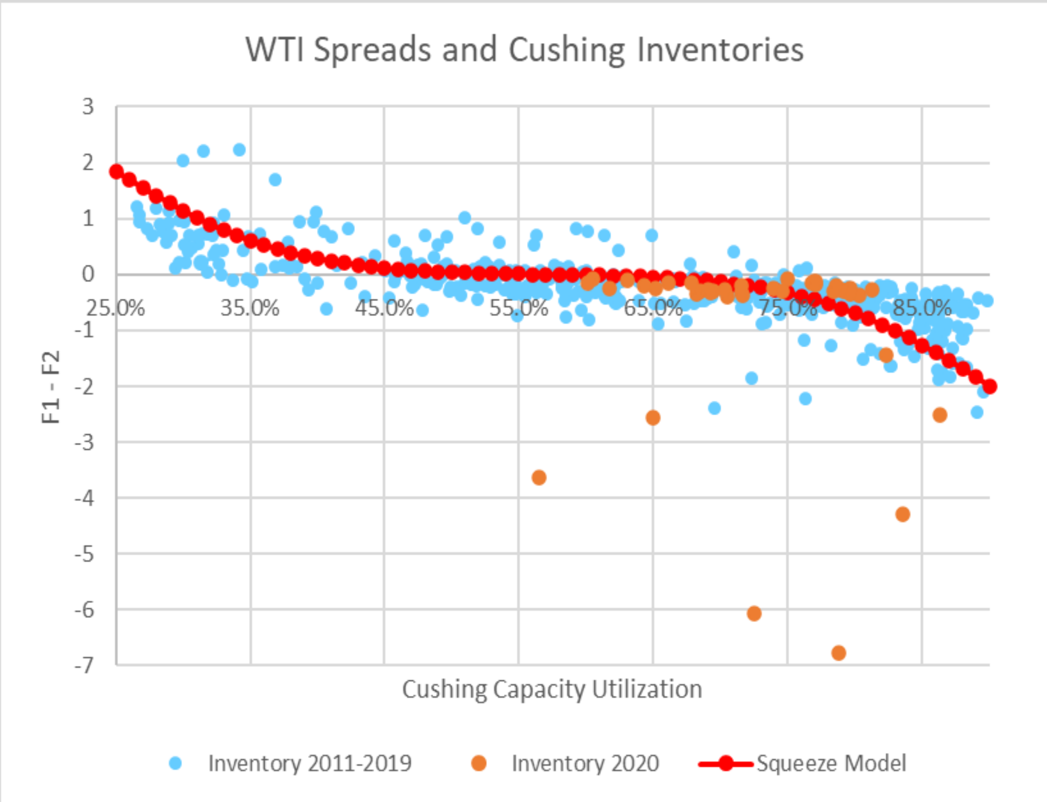
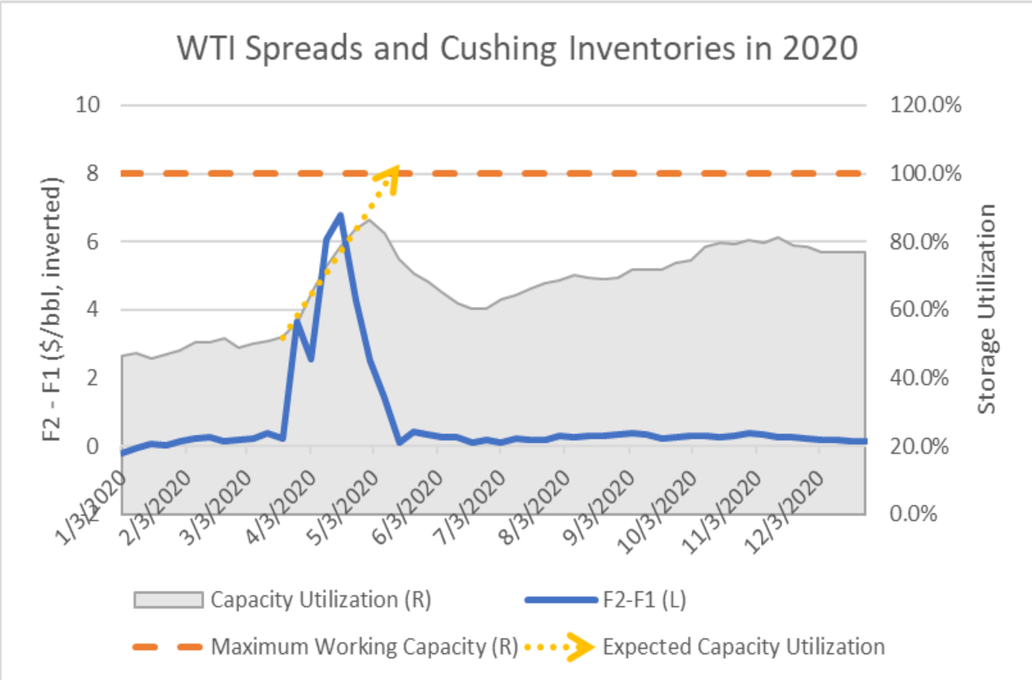


Figure 2 shows the relationship between weekly inventories reported and futures time spread. Note that the day of April 20, 2020 when oil prices went negative and the futures spread settled at negative $58.06 per barrel is not shown here, as we only use end of the week prices to match the timing of published inventories. Given the magnitude of the move on April 20, 2020, any attempt to calibrate the model to this outlier will make other data points invisible, so we will discuss this one-off event separately below. The graph generally captures non-linear effects induced by the presence of two extreme boundaries reasonably well with model parameters specified in the Appendix. Calibrating any analytical formulas to noisy fundamental data, such as inventory, is a daunting exercise, but it clearly improves the accuracy of standard linear regression-based methods, which are powerless in modelling tail events. Not surprisingly, other large outliers occurred during the events preceding the price collapse on April 20, 2020. To understand the cause, it is helpful to zoom in on the trajectory of weekly inventories during that time, which is shown in Figure 3.

The first immediate observation is that contango started to steepen when Cushing inventories were still less than 50% of the working capacity, which is below even their historically normal level. At that time, the risk of tank tops in Cushing appeared to have been a very remote possibility. The graph highlights an important, but often overlooked feature of the oil markets, that futures trade not based on the currently observed inventories, but rather based on forward looking expectations where inventories will be at the time of delivery.

Figure 3: The futures spread is trading based on forward looking expected inventories, not contemporaneously measured inventories



The proper way to look at the relationship between inventories and futures would require shifting the process for inventories in our model about one month forward. Obviously, in this case the starting point for the inventory process is not known today. In this case, our structural assumption about longer-term inventories mean-reversion could be supplemented with some short-term inventory persistence. Since the supply and demand are slow to adjust, the imbalance that exists today will likely be similar to the imbalance tomorrow. The simplest way to estimate such imbalance forward is to extrapolate the slope of the inventory path forward for a few weeks. This is exactly what the market did in April 2020. If the slope of inventory builds during March and April were extrapolated, then Figure 3 shows that the inventory would have indeed reached the tank tops by the end of May when physical barrels must be delivered.

While the situation of tank tops creates a hard boundary that cannot be exceeded, oil inventories can rarely be analyzed in isolation for a single location, such as Cushing, Oklahoma. Given enough time, inventories could be easily transported to alternative destinations, where plenty of storage capacity was still available across various U.S. storage locations. This is where the mean-reverting nature of the inventory process kicks in, and the slope of the inventory trajectory starts changing as invisible hands of supply and demand force gradual rebalancing. The expectations of tank tops to be reached in May 2020 never materialized.

Our simple model is only a stylized representation of the market reality which should not be used as an econometrically rigorous framework. Nevertheless, it can only attribute around $5 per barrel to the limitation of storage. This estimate simply reflects the cost of storage in alternative locations and consistent with where futures spreads traded before and after the event. The remaining $55 per barrel clearly came from non-fundamental factors as the result of the classical financial squeeze. The magnitude of the price move was probably as close to infinity, prescribed by Dirac delta function, as one can find in any financial markets.

CFTC Report and The Myth of Storage

We have previously summarized our interpretation of the events that led to the collapse of WTI futures on April 20, 2020[6]. To recap, many over-the-counter financial products were settling on this day, some of which turned out to have material flaws in their design. Among the most vulnerable products was the over-the-counter derivative contract, known as Yuányóu Bǎo, heavily marketed by the Bank of China to its domestic retail clients looking to buy cheap oil. The product did not, however, emphasize the fact that any holder of a long oil futures is implicitly involved in the storage-like carry trade without having access to the actual storage facility. The return on buying futures is always made up of the change in the price and an indirect cost from rolling the position, which is effectively the payment for outsourcing the function of storage[7]. As such, at expiry the longs are always vulnerable to the squeeze, or the risk described above by the Dirac function, which goes to a negative infinity when the buyer cannot take the delivery.

The story of the winners in this trade has also been recently documented[8]. Exact details behind both stories are not yet known with big hopes placed on the widely anticipated study of the event by the Commodities Futures Trading Commission (CTFC). Unfortunately, the recently published Interim Report by CFTC largely disappointed[9]. Instead of acknowledging that the event was caused by financial flows, it focused on reciting various statistics which added little to the explanation of the actual event. Ironically, the regulators in China already recognized the role that the Bank of China's derivative product played in the event and even penalized the bank.[10]

Traders can only applaud one dissenting Statement by the CFTC Commissioner acknowledging that "the issuance of an incomplete preliminary Report is a disservice to the public"[11]. As highlighted in the Statement, "the Report does not undertake any analysis of the actual storage situation in Cushing" and references only "anecdotal reports that crude oil storage capacity was in short supply at Cushing, Oklahoma". In addition, according to the Executive Summary of the Report, "by March 2020, the working storage available at the Cushing facility was near capacity", which is inaccurate, as can be clearly seen in Figure 3. The inventories had not reached high level until a month later. Moreover, Figure 3 shows that as recently as November 2020, the level of Cushing inventories was higher than they were throughout most of March 2020, while oil was trading at positive rather than negative $40 per barrel.

The Report does not explicitly acknowledge that the episode of negative prices was the result of the financial squeeze. While the storage capacity constraints can explain up to $5 per barrel, according to our model, it accounts for only ten percent of the actual price move on April 20, 2020. With daily futures volumes exceeding daily oil consumption by a factor of at least forty, the tightness of storage was only an early catalyst leading to the event. Every market participant knows that storage cannot change by almost $60 back-and-forth within a few hours, and one must look elsewhere.

The Report also did not study the netting of positions held by the largest traders across related instruments and trading venues. While the Report provided a comment on over-the-counter reportable swaps being reviewed, it did not include them in the analysis. Interestingly, the Report discovered that the largest long positions on April 20, 2020 were held not by hedge funds or commercial participants

---

[6] I. Bouchouev (2020a), The Rise of Retail and the Fall of WTI, Oxford Institute of Energy Studies.
https://www.oxfordenergy.org/wpcms/wp-content/uploads/2020/05/The-rise-of-retail-and-the-fall-of-WTI-.pdf

[7] I. Bouchouev (2020b), From risk bearing to propheteering, Quantitative Finance, vol. 20, no. 6, pp. 887-894.

[8] L. Vaughan, K. Chellel, and B. Bain (2020), The Essex Boys: How Nine Traders Hit a Gusher with Negative Oil, *Bloomberg Businenessweek*, December 10.

[9] Commodity Futures Trading Commission. Interim Staff Report: Trading in NYMEX WTI Crude Oil Futures Leading up to, on, and around April 20, 2020., November 23, 2020.

[10] Regulator fines Bank of China over loss-making product linked to crude oil, *Reuters*, December 5, 2020.
https://www.reuters.com/article/china-bank-of-china-oil-regulator/regulator-fines-bank-of-china-over-loss-making-product-linked-to-crude-oil-idINKBN28F0EA

[11] Statement of Commissioner Dan M. Berkovitz Regarding the CFTC Staff Report on the Trading of NYMEX WTI Crude Oil Futures Contracts On and Around April 20, 2020,
https://www.cftc.gov/PressRoom/SpeechesTestimony/berkovitzstatement112320a

trading futures, but by over-the-counter swap dealers. Many of these dealers rely on various hedge exemptions against over-the-counter products, which indirectly confirms that the largest price risks on that day were held by over-the-counter speculators. In addition, large in-the-money put options on calendar spreads were also settling on the same day. Since all put options were deeply in-the-money, they behaved identically to synthetic futures, forming another significant market segment which has been ignored.

Nearly all over-the-counter products and options are settled based on the closing price with resulting price risks managed using the instrument, known as trading-at-settlement (TAS)[12]. While the Report acknowledges the important role of this instrument by providing an assortment of graphs and statistical summaries, it remained silent on the real story behind the role of TAS on that day. The most important observation that likely triggered the entire event was the unprecedented break in liquidity in TAS trading, which for the first time reached its daily downside limit. It immediately telegraphed to market participants that some financial traders were unable to liquate their long position using TAS, and their only way out was to sell futures using regular market orders placed near the closing time. The holders of long futures contracts were effectively hand-cuffed by rigid mandates imposed by their over-the-counter derivatives products. Their only motive is to sell futures and match whatever the settlement price at end of the trading day will be. Using any discretion would have inevitably led to complains and legal challenges by their clients. At the same time, for discretionary traders it presented an opportunity to sell futures intra-day and buy them back during the closing window at lower prices pressured down by mandatory liquidation of long positions.

Finally, the Report owes the public some clarity around the effectiveness of intra-day position limits, which prevent any single market participant from holding more than three million barrels of risk during the last three trading days. The existence of such limits appears to contradict much larger losses and profits reported by various media sources. A useful study would have been to net positions held by the largest futures holders by their trader ID across all known derivative products linked to WTI futures. Some statistical summaries of the largest traders' profit and losses, including hedges of over-the-counter products and options, would have been much more relevant for educating the public on what happened that day. This could have been accomplished within CFTC's existing authority without disclosing participants' identities or their trading strategies.

Unfortunately, the Report missed the opportunity to unveil a mystery behind this unprecedented event. The reluctance to openly acknowledge the fact that speculation in derivatives contracts is by far the larger driver of oil prices, is indeed a disservice to the public. Simply pointing to the storage is an easy way out. The hidden realities are much more complex.

[12] TAS instrument was analyzed in C. Pirrong (2019), Derived Pricing: Fragmentation, Efficiency and Manipulation, University of Houston, working paper.

Appendix

We model the spread between two nearby futures F_1 and F_2 as a function of inventories, $x(t)$,

$$S(t;x) = F_1(t,x) - F_2(t,x).$$

The inventory variable is scaled with capacity

$$x(t) = \frac{Inventory(t)}{Capacity(t)}$$

The scaled inventory is viewed as a state variable, which follows the mean-reverting stochastic process

$$dx = k(\bar{x} - x)dt + \sigma dz,$$

where $\bar{x}$ is the long-term normal level of inventories, k is the speed of inventories mean-reversion to the normal level, σ is volatility, and dz represents independent normally distributed increments with zero mean and the variance dt. The process defined for inventories within minimum and maximum operating capacity of the storage tanks

$$0 < x_{min} \leq x \leq x_{max}.$$

At the expiry T of the first futures contract, the spread converges to the negative of the normal cost of storage, U, unless inventory is either zero or at the maximum storage capacity, in which case the futures trader defaults

$$S(T;x) = \begin{cases} -U, & x_{min} < x < x_{max} \\ +\infty, & x = x_{min} \\ -\infty, & x = x_{max} \end{cases}$$

Then the futures spread at any time $t < T$ prior to expiry can be expressed as the spread between two normal probability density functions corresponding, respectively, to the upside squeeze at zero inventories and to the downside squeeze, which occurs at zero available storage capacity, adjusted for storage costs

$$S(t,x) = \frac{1}{\sqrt{2\pi\tau}\sigma}\left\{e^{-\frac{(y-x_{min})^2}{2\sigma^2\tau}} - e^{-\frac{(y-x_{max})^2}{2\sigma^2\tau}}\right\} - U.$$

Here, an auxiliary variable y represents the level of inventories, which dynamically shifts current inventory x towards its long-term mean with the speed k

$$y = \bar{x} + (x - \bar{x})e^{-k(T-t)}.$$

The variance in both probability densities is using the reduced effective time variable

$$\tau = \frac{1}{2k}\left(1 - e^{-2k(T-t)}\right),$$

which is a typical characteristic of mean-reverting processes.

Model parameters could be estimated using various penalty functions to better capture the outliers. The following set of parameters was calibrated for illustration in Figure 2:

$$T - t = \frac{1}{12}, \quad \sigma = 0.4, \ k = 1, x_{min} = 0.15, x_{max} = 1.0, \bar{x} = 0.6.$$