Arbitrage breakdown in WTI crude oil futures: An analysis of the events on April 20, 2020

Christopher B. Burns[1], Stephen Kane

Commodity Futures Trading Commission, 1155 21st Street NW, Washington, D.C. 20581

Abstract

This study provides novel analysis of the events in the WTI crude oil futures market on April 20, 2020. We detail how the arbitrage linkages between the NYMEX CL contract and the e-mini NYMEX QM contract broke down and report new information about the unusual market conditions on that date. After establishing that most price discovery happens in the more liquid CL contract, we show how these two contracts decoupled in the May 2020 spot period. Next, using supervisory CFTC data, we document that the typical arbitragers did not participate in the WTI crude oil markets on April 20. This change in the composition of arbitragers had important implications for the unusual settlement prices in the CL contract. Third, we use generalizable non-parametric methods to rank the values observed in terms of price deviations, realized volatility and spreads to similar crude oils. We find the May 2020 spot month to have the largest values of these measures across all spot periods from 2011 to 2020. Finally, we show that natural gas futures markets did not experience a similar price decoupling, suggesting the lack of storage capacity at Cushing played an important role in the WTI crude negative price event.

Keywords: Arbitrage, Price decoupling, West Texas Intermediate, Non-parametric methods

---

[1]Corresponding author, *email address: cburns@cftc.gov*. The research presented in this paper was co-authored by Christopher Burns and Stephen Kane, CFTC employees who wrote this paper in their official capacities. The analyses and conclusions expressed in this paper are those of the authors and do not reflect the views of other members of the Office of the CFTC Chief Economist, other Commission staff, or the Commission itself.

1. Introduction

On April 20, 2020, the penultimate day of the NYMEX West Texas Intermediate (WTI) May contract (CME ticker symbol CL), a confluence of factors, including a shock to global demand caused by the COVID-19 pandemic, an oversupply of oil, and limited storage capacity at the delivery point of Cushing, OK, led the contract to settle at -$37.63 per barrel. This was the first time the NYMEX CL contract had settled or traded below $0 since its inception in 1983 (CFTC 2021). This was an unprecendented event that was widely covered in the media.[2]

As the price of the CL May contract went negative, the price of the NYMEX E-mini WTI contract (CME ticker symbol QM), which financially settles to the larger CL contract, decoupled and stopped transacting. This was a highly unusual event, as the law of one price predicts that opportunities for arbitrage would keep prices for these two contracts closely linked. An examination of all spot periods from 2011-2020 reveals the average difference in the per-barrel price between these two contracts is near zero, with small, short-run deviations observed. The one exception is seen on April 20, 2020(figure 1).

[2]For example, see: `https://www.nytimes.com/2020/04/21/upshot/negative-oil-price.html`.

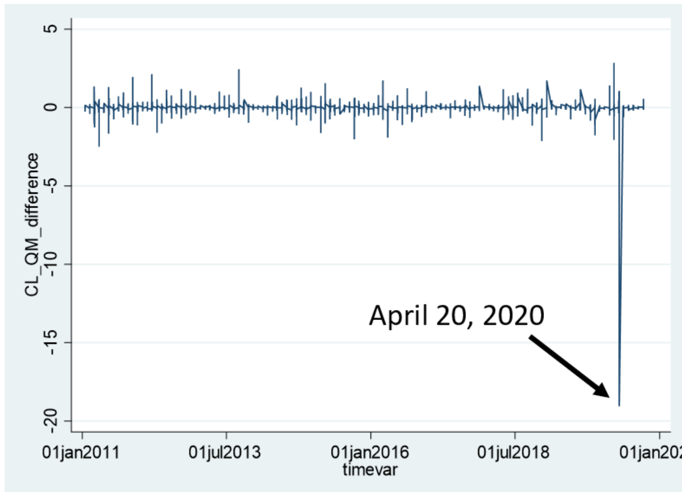

Figure 1: Difference in volume-weighted average prices (per barrel) of CL and QM contracts, aggregated at 10-minute level during the spot month, 2011-2020

Our study makes several empirical contributions to the literature on arbitrage and price discovery in crude oil markets. First, we show that the price decoupling between the CL and QM contracts was likely caused by the lack of typical arbitrageurs in the WTI crude oil market.[3] Without the arbitrageurs that typically provide liquidity and capture price inefficiencies between these markets, the conditions for large price movements were made possible. Second, we use novel non-parametric methods to rank the observed measures of price deviations, realized volatility, and price spreads with similar crude oils on April 20, 2020 (May 2020 spot period), with all spot periods from 2011 to 2020. We find the May 2020 spot month to have the largest values observed across all these different assessments.[4] Finally, we discuss potential policies that could

[3]We note that participation in U.S. futures markets with limit orders is voluntary, with no designated market makers. In contrast, many equities exchanges, such as the New York Stock Exchange (NYSE), have Designated Market Markers (DMMs) that can augment limit orders to ensure fair and orderly markets.

[4]This analysis concerns only the observations by the authors and does not represent a legal conclusion with respect to the applicability of any provision of the Commodity Exchange Act

help prevent a similar event from happening again in WTI crude oil futures markets.

Much of our empirical analysis uses non-parametric methods due to the large outliers (i.e. "fat tails") in our data. All of our non-parametric methods are based on widely used statistical assessments of financial markets used in the finance literature. To understand where price discovery occurs we use a co-intergrated quantile regression model that allows the dependence structure between the prices in the CL and QM contracts to vary across the quantiles of these distributions. This method allows us to model the price relationship between these two contracts in a more flexible framework, and to show how they decoupled on April 20, 2020. Our non-parametric techniques are generalizable to any assessment that permits a rank ordering (monotonic relationship). We use non-parametric assessments of price deviations, realized volatility, and price spreads to accommodate the large outliers in our data.

The rest of the article is organized as follows: The next section provides more background on the CL and QM contracts, followed by a review of the relevant literature. The data and non-parametric methods sections discuss the price data and explore why it does not fit parametric assumptions. The empirical section summarizes our measures of price discovery, and rankings for the price spreads, deviations, realized volatility, and distribution of arbitrageurs across all spot periods from 2011-2020. The final section summarizes our results and includes a policy discussion.

2. Background

The price discovery role of the CL contract is relied on by commercial and non-commercial entities for purposes of hedging and risk-management in the crude oil market. Daily volume traded is approximately 1.2 million contracts

---

or any relevant regulations. The analysis is based upon the information available to the authors at the time the paper was written, and any different, changed, or omitted facts or circumstances would require additional analysis and might result in different conclusions.

per day (CFTC 2021). As a global benchmark, WTI serves as the key reference rate for physical and financial oil transactions around the world in the cash (spot) and futures market. One contract of CL is for 1,000 barrels of physically deliverable crude oil to Cushing, OK.

The QM contract was created by NYMEX in 2002 as a response to market participant interest (Tse & Xiang, 2005). One contract of QM is financially settled for 500 barrels of WTI crude oil. On a typically day the QM contract has only 1% of the volume of the CL contract. The QM price is tethered to the final settlement price of CL on the day before the expiration of the physically-settled CL contract. Both contracts are free to float before expiration of QM. Market segmentation between participants in the financially-settled QM contract and the physically-settled CL contract is likely because delivery is not required to stay in the financial contract. The QM contract is mostly used by financial entities to gain price exposure to WTI for investment purposes, though some entities (commercials) use it for hedging and other purposes.

The next section of the paper describes the relevant literature on crude oil markets and non-parametric methods in finance. We show that these methods give the researcher the flexibility to develop test statistics that do not depend on distributional assumptions. Our analysis then turns to the price discovery linkages between the CL and QM contracts.

3. Related literature

Theory predicts that futures contracts with similar underlying assets should have a long-run equilibrium price relationship. If prices for the same asset are different across markets, then a profit can be made by buying in one market and selling in the other. This arbitrage mechanism should keep prices coupled (law of one price) across different markets. Kawaller et al. (1987) show that S&P 500 futures and S&P 500 index prices are linked through arbitrage and that most price discovery occurs in the E-mini futures market. The growth of e-mini contracts in futures and equity markets has expanded possible markets

Tse & Xiang (2005) find that e-mini contracts contribute to price discovery in WTI crude oil and natural gas futures markets. The authors find that the smaller contracts improve market quality by reducing the bid-ask spread, and contribute to price discovery in the more liquid physically-settled contract.

The relationship between storage, price spreads, and convergence in futures markets is an area of active research. In the crude oil markets, Buyuksahin et al. (2013) examine the link between inventory conditions and the futures term structure, as well as the WTI-Brent spread. They find storage constraints at Cushing play an important role in spread prices. A recent study by Ederington et al. (2021) finds that storage capacity limits in U.S. crude oil markets can impede cash-and-carry arbitrage. Work by Irwin (2020) and Garcia et al. (2015) demonstrate that non-convergence in grain markets during the mid-2000s occurred when the market price of the physical grain storage exceeds the storage rate on delivery instruments. We can see a parallel in the WTI crude oil market on April 20, 2020, where storage capacity was extremely limited, and futures prices settled below \$0, while cash market transactions occurred at positive prices.[5] The Commodity Futures Trading Commission's Staff report also remarks that working storage at the Cushing facility was near capacity by March 2020 and that some industry participants were already preparing for the prospect of negative prices in late March and early April.[6]

There is a substantial body of research on forecasting crude oil prices. Miao et al. (2017) improve on existing methods of forecasting crude oil prices by including explanatory variables about commodity and financial markets, supply and demand, speculative activity, and geopolitical factors into a LASSO regression model framework. Other factors important to predict oil prices include

[5] Media reports confirm that there was strong interest in leasing oil tanks at Cushing prior to the May 2020 spot period, but that all tanks had been leased by mid-March: `https://pgjonline.com/news/2020/04-april/remaining-oil-storage-in-cushing-ok-is-already-booked-traders`.

[6] Report available at: `https://www.cftc.gov/PressRoom/PressReleases/8315-20`.

price jumps (Buncic & Gisler, 2017), economic news (Elder et al., 2013), and crude oil inventory announcements (Miao et al., 2018). Recent studies have examined more sophisticated methods for forecasting crude oil realized volatility and prices, such as standard mixed data sampling (MIDAS) (Chen et al., 2022), and combining variation mode decomposition with random sparse integrated Bayesian learning (Li et al., 2021).

Crude oil markets with similar quality crude oil are often co-integrated, meaning they have a long-run relationship (Galay, 2019). In addition, crude oil markets can be co-integrated across different oil producing regions (Weiner, 1991),and co-integrated with their respective spot markets (Schwarz & Szakmary 1994; Silvapulle & Moosa 1999). Some studies argue that the world oil market is "one great pool" (Adelman, 1984) and prices co-move together, while others have found evidence of regionalization in crude oil markets (Gulen, 1997).

Vector error correction models (VECM)(Johansen, 1988, 1991)) are commonly used to estimate price dynamics in co-integrated markets, such as between WTI and Brent(Liu et al., 2015). However, linear models, such as a VECM, have been shown to have low power when nonlinearities are present in the data (Hiemstra & Jones, 1994). In such circumstances, non-parametric quantile regression methods can provide a more flexible modeling approach (Koenker & Hallock, 2001). A recent study by Yang et al. (2021) uses causality-in-quantiles methods to estimate the effects of oil shocks on commodities across the quantiles of the distribution. Price discovery patterns are widely estimated using methods such as Hasbrouck (1995)'s Information Shares and Gonzalo & Granger (1995)'s Permanent-Transitory Common Factor Weights.

A few studies have examined how the financialization of commodity markets (Cheng & Xiong, 2014) could impact the functioning of the WTI oil market and other commodity markets. Fernandez-Perez et al. (2020) use a Granger-causality framework and find no evidence that the largest WTI crude oil Electronically Traded Fund (ETF), the United States Oil Fund (USO), had a price impact on the WTI market on April 20th. In fact, most ETFs had rolled their positions into the active contract (June) well before the spot month.

A recent study has explored the direct impacts of negative WTI crude prices on energy producers. Gilje et al. (2020) show how low prices in WTI on April 20th caused oil producing firms to stop production (i.e. shut-in wells) in areas of the United States far from Cushing, but were affected via WTI purchase contracts. Many firms have crude oil purchase contracts that are indexed to WTI futures and the low prices dramatically affected the profitability of wells, particularly in North Dakotas where production could be shut off more easily due to the use of fracking technology[7].

In contrast to previous studies, our paper examines how the price events in the CL contract on April 20, 2020 interacted with cash-settled contracts on near substitutes, such as the QM and BZ (CME Brent crude oil) contracts. Further, we provide new evidence on the distribution of arbitrageurs and show how it was significantly different from any spot period in the CL contract across the time period 2011-2020.

4. Data

Our data set comes from the Transaction Capture Reporting (TCR) system of the U.S. Commodity Futures Trading Commission. We use TCR data to capture all trades in the CL and QM contracts in the spot month, or the last three days before expiration, between January 3, 2011 and December 31, 2020. January 2011 is the first month where reliable WTI crude data from TCR is available. The TCR database includes fields about the counterparties, which side of the trade is represented, and provides identification about the executing firm and the trading account. In addition, we use TCR data on NYMEX

[7] We note that not all spot markets reflected the abnormal pricing in the futures market on April 20. Platts did not report negative spot prices for WTI crude. We believe they used a combination of spreads to Brent and WTI Midland prices to create the WTI benchmark when futures prices went negative. This override on their price methodology prevented a reporting of negative spot prices. For more details on Platts pricing methodology, go to:https://www.spglobal.com/platts/plattscontent/_assets/_files/en/our-methodology/methodology-specifications/platts-assessments-methodology-guide.pdf

financially-settled Brent Crude (CME ticker BZ) for the same time period as WTI CL, in order to compare the price differentials between the two crude oils.

We also examine data from TCR on Trade-at-Settlement (TAS) volume in the WTI crude oil market (CL). TAS trades are an order type that allows a participant to execute a trade at a differential or defined number of tick increments, above or below that day's settlement price, at any time during the trading session. We only use data on TAS volume for the purposes of analysis in this paper, not TAS trades. In this way, we consider only price-forming transactions. Our final TCR data includes Natural Gas trade prices in the spot month for both the financially settled (NYMEX ticker HH) and physically settled (ICE ticker H) contracts from January 1, 2015 through December 31, 2020. January 2015 was the earliest that reliable TCR data on natural gas was available.

4.1. Summary Statistics

In order to make the data usable for event-time analysis, we create a volume-weighted average price (VWAP) from the TCR data on the CL and QM contracts. Our data are measured at the 10-minute level in every spot month from 2011-2020, giving a total of 120 spot months. However, we do not observe a CL and QM trade in every 10-minute window for each spot month. Our final dataset contains 11,923 observations.

To test whether our price series have a unit root, we perform a Augmented Dickey-Fuller Generalized Least Squares (ADF-GLS)(Elliott, 1999) test. Optimal lags for the test are chosen using the Akaike Information Criteria (AIC). The ADF-GLS test finds that both the CL and QM price series are not stationary in levels. After taking first-differences of both price series, we reject the null hypothesis that CL and QM prices have a unit root process at the 1% level of significance. Summary statistics for CL and QM prices and their first-differences are shown in table 1.

Table 1: Summary statistics for volume-weighted average prices (VWAP) in CL and QM contracts and first differences (Δ), 2011-2020.

	Obs	Mean	SD	Skewness	Kurtosis	ADF-GLS
CL	11,923	71.527	109.844	-0.162	1.786	-0.843
QM	11,923	71.529	109.919	-0.159	1.776	-0.848
Δ CL	11,922	-0.0003	0.030	-10.967	2569.260	-76.160***
Δ QM	11,922	-0.002	0.146	-93.481	9439.384	-77.176***

Note: *** ,** and * indicate the ADF-GLS test unit root test is rejected at the 1% level.

We note that even after taking a first-differences transformation both the CL and QM prices show significant deviations from normality. For example, the Δ CL price data have significant probabilities in the tails of the distribution, likely due to the large price movements on April 20. These "fat tails" can be seen in the QQ-norm plot in figure 2.

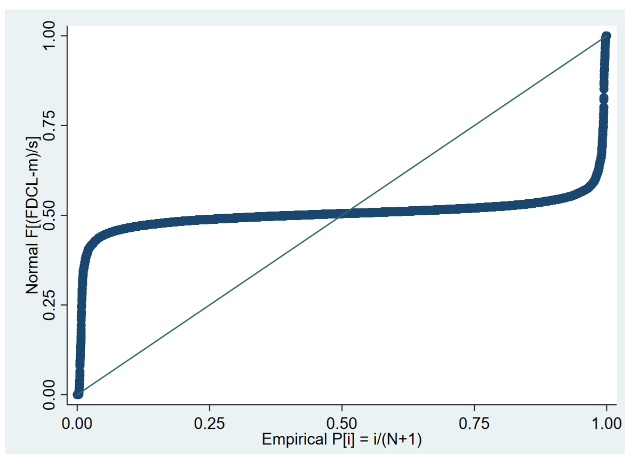

Figure 2: Normal quantile plot of first-differences in CL prices

A similar finding is seen with the Δ QM price data. Additional univariate tests for normality reject that either data series does not have skewness or kurtosis at the 1% significance level. We also see that joint plots of first differences in CL and QM prices are not spherical (appendix figure 4) and do not come from a clear parametric family. These findings motivate our use of non-parametric methods in the empirical methods section.

5. Non-parametric methods

Our non-parametric estimators of price deviations, price spreads, realized volatility, and price discovery are based on well-known parametric estimators used to measure arbitrage conditions. We use non-parametric methods because our price data contain "fat tails" but these methods are generally applicable to any financial data that violate the classical distributional assumptions used for statistical inference (Eom et al. 2019, 2021). In such circumstances, the use of non-parametric or distribution-free methods are preferred for statistical analysis because they are more robust to large outliers (Maritz 1995; Kane 2004).

Statistical inference relies on the empirical cumulative probability distribution of the rank-ordered statistics. Working with the empirical distribution, we give each observation the probability of 1/N, where N is the total number of observations. If two observations are tied, then we give the probability 2/N, etc. Each observation of the test statistic is selected from a uniform draw with replacement. This is known as bootstrap sampling(Efron, 1982) and can be used with any assessment that permits ranking thought a monotonic relationship. P-values for our test statistics are computed by the relative frequency of observations bigger than or equal to the observed value.

6. Empirical Results

The first part of our analysis tests whether CL and QM prices are co-integrated. After establishing co-integration, we examine which contract contributes more to price discovery using a co-integrated quantile regression model. Next, we examine the extent of price decoupling in the May 2020 spot month by using non parametric order statistics that rank the price deviations, realized volatility, and price spread with Brent crude oil, for each spot period from 2011 to 2020. We then examine the distribution of arbitrageurs and test whether it was significantly different on April 20, 2020. Finally, we estimate a difference-in-differences model that compares the spread between the cash-settled and physically-settled contracts in the natural gas markets with WTI crude oil markets in the May 2020 spot period. All empirical analysis shown was estimated using STATA or SAS.

6.1. Price Discovery in WTI crude oil

The quantile regression method allows us to model the relationship between the CL and QM prices in a more flexible manner. It allows for a dependence structure that is not linear and does not depend on the errors being normally distributed (i.e. data with "fat tails"). Baur (2013) notes that the quantile-based approach allows for asymmetric dependence structures, where there may be more dependence in the right or left tail of the distribution. Koenker & Hallock (2001) observe that quantile regression methods provide an alternative approach for robust inference because they allow the researcher to explore a range of conditional quantiles, allowing for forms of conditional heterogeneity. Quantile regression uses linear programming methods (e.g. simplex method) to solve the minimization problem instead of least squares in the case of conditional mean.

A closer analysis of our price data series reveals there is a structural break in late 2014 and early 2015. There is a significant drop in WTI crude oil prices during this period caused by the U.S. shale oil boom and increased the supply from the Bakken fields in North Dakota. To keep the econometric model parsimonious, we opt to start our data series after January 2015. The remaining five years of crude oil data gives us 6,040 observations for our analysis.

We test for the maximum order of integration by using the (Johansen, 1991) test. We find that CL and QM prices are indeed co-integrated with one common factor at the 5% level of significance. Next, we test for the number of lags using the Akaike Information Criteria (AIC). The AIC selects 13 lags as optimal for a vector-autoregressive model. This means our co-integrated quantile regression model will have 12 lags.

Following Bianchi et al. (2020) and Troster et al. (2018) we extend a standard quantile auto-regressive model by taking first differences and including an error correction term. We also interact a dummy variable with the error-correction term, where the dummy variable is equal to one in May 2020 spot period and 0 otherwise. This interaction variable will allow us to test whether the long-run relationship between the CL and QM contracts decoupled in the May 2020 spot period. We then specify a bivariate error-correction model for each conditional quantile as follows:

$$\begin{aligned} Q_{\Delta CL_t}(\tau|\boldsymbol{X}) &= \alpha_{QM,\tau} + \beta_{CL,1,\tau} ECT_{t-1} + \beta_{CL,2,\tau} D_{May20,t} ECT_{t-1} \\ &\quad + \sum_{i=1}^{12} \gamma_{1,i,\tau} \Delta CL_{t-i} + \sum_{i=1}^{12} \gamma_{2,i,\tau} \Delta QM_{t-i} + \epsilon_{CL,t}, \\ Q_{\Delta QM_t}(\tau|\boldsymbol{X}) &= \alpha_{QM,\tau} + \beta_{QM,1,\tau} ECT_{t-1} + \beta_{QM,2,\tau} D_{May20,t} ECT_{t-1} \\ &\quad + \sum_{i=1}^{12} \gamma_{3,i,\tau} \Delta CL_{t-i} + \sum_{i=1}^{12} \gamma_{4,i,\tau} \Delta QM_{t-i} + \epsilon_{QM,t}, \end{aligned} \tag{1}$$

where $Q_{\Delta .}(\tau|\boldsymbol{X})$ denotes the conditional quantile of the first-differenced CL or QM price series, $0 < \tau < 1$ denotes the quantile, and ECT_{t-1} is the lagged error correction term ($ECT_t = CL_t - \theta_0 - QM_t$). The error correction term proxies for the speed of adjustment caused by a price disequilibrium with the other contract. The $D_{May20,t}$ term represents a dummy variable, equal to one in the May 2020 spot period, and zero otherwise. The coefficients on the error correction term, $\beta_{CL,1}$ and $\beta_{QM,1}$, capture the effect of the equilibrium adjustments in all periods excluding the May 2020 spot period. The coefficients, $\beta_{CL,2}$ and $\beta_{QM,2}$, capture any additional equilibrium adjustment effect during the May 2020 spot period.

In order to compute a proxy for price discovery, we use a version of the Permanent-Transitory Common Factor Weights (CFW) that relies on the absolute relative magnitude of the error correction terms to determine each contract's contribution to price discovery (Cabrera et al., 2009). The CFW for the CL and QM contracts evaluated at quantile τ are:

$$CFW_{CL,\tau} = \frac{|\alpha_{QM,\tau}|}{|\alpha_{QM,\tau}| - |\alpha_{CL,\tau}|} \quad , \quad CFW_{QM,\tau} = \frac{|\alpha_{CL,\tau}|}{|\alpha_{CL,\tau}| - |\alpha_{QM,\tau}|} \tag{2}$$

Estimated parameters for the error correction terms and common factor weights are shown at each conditional quantile (i.e. q10, q20, ..., q90) in table 2. We find that the error correction term to be significant at the 5% level in both CL and QM contracts across all spot periods in 2015-2020, excluding May 2020. This is seen across quantiles 10 through 80 (i.e. q10 - q80). This suggests that both contracts contribute information to the equilibrium price across this time period. The much smaller coefficients for β_{CL} reflect CL prices incorporate information earlier and thus, require smaller equilibrium adjustments to QM prices. This small adjustment factor is seen across all quantiles of the ΔCL_t distribution. In contrast, the estimates of β_{QM} are relatively larger, reflecting that it makes larger adjustments to changes in CL prices.

Table 2: Estimates of error correction term coefficients and Common Factor Weights (CFW) by quantile for all spot periods in 2015-2020, excluding May 2020 (top), and the May 2020 spot period (bottom)

	***q*10**	***q*20**	***q*30**	***q*40**	***q*50**	***q*60**	***q*70**	***q*80**	***q*90**
$\beta_{CL,1}$	-0.14**	-0.11***	-0.11***	-0.12***	-0.12***	0.11**	-0.12*	-0.16*	-0.19
CFW_{CL}	82%	85%	84%	81%	80%	80%	77%	70%	62%
$\beta_{QM,1}$	0.62***	0.64***	0.56***	0.50***	0.49***	0.43***	0.42***	0.37**	0.31
CFW_{QM}	18%	15%	16%	19%	20%	20%	23%	30%	38%
$\beta_{CL,1} + \beta_{CL,2}$	0.12	0.13	0.09	0.10	0.09	0.11	0.13	0.15	0.12
CFW_{CL}	88%	86%	90%	88%	78%	70%	66%	58%	54%
$\beta_{QM,1} + \beta_{QM,2}$	0.88***	0.81**	0.80**	0.72*	0.32	0.26	0.25	0.21	0.14
CFW_{QM}	12%	14%	10%	12%	22%	30%	34%	42%	46%

Note: *** ,**, and * indicate the parameter estimate is significant at the 1, 5, and 10 percent level, respectively. Quantile regression standard errors are based on 100 bootstrap resamplings.

We find that most price discovery happens in the more liquid CL contract over the 2015 - 2020 period, excluding May 2020. Evaluated at the median (q50), the CL contract contributes 80% to price discovery in WTI crude oil, while the QM contract contributes 20%. This is consistent with the larger, more liquid, CL contract taking in information more rapidly than the smaller QM contract. We also find that price discovery contribution of the CL contract varies between 62-85% across the ΔCL_t distribution, with a lower share of price discovery at the higher quantiles.

Next, we examine the coefficients on the error correction term in the May 2020 spot period and test whether they are statistically significant from zero.[8] The bottom portion of table 2 displays results of a test of the sum of $\beta_{CL,1}$ and $\beta_{CL,2}$, representing the effect of the error correction term in the May 2020 spot period for CL contract. The test results reveal they are not statistically different from zero at any quantile of the distribution. This provides evidence that the CL contract stopped incorporating information from the QM contract during the May 2020 spot period. This is consistent with the CL contract decoupling from the smaller financially settled QM contract. We also find no evidence of a significant error correction effect in the QM contract at quantiles 50 through 90. However, we do see evidence that QM prices were still responding to changes in CL prices at quantiles 10 through 40, where more negative price changes occurred. This is consistent with the observation that QM prices followed CL prices as they plunged towards $0 per barrel on April 20, 2020.

When we examine price discovery in the May 2020 spot period, we see still observe more price discovery occurring in the CL contract. However, compared with the 2015 - 2020 period, a larger share of price discovery happens in the CL contract at the lower quantiles (i.e. q10 - q40), where CL price changes are more negative. We note that because many of the error correction terms are not significant, the price discovery estimates contain a lot of variability. The lack of arbitrage happening in the May 2020 spot period between these two contracts likely contributes to the noise in measures of price discovery.

Overall, the co-integrated quantile regression model provides clear evidence that most price discovery happens in the CL contract and that these two contracts decoupled in the May 2020 spot month. The more flexible model also shows how price discovery and equilibrium adjustments between these two contracts vary over their respective price distributions. Our next set of analyses

[8] We have 197 observations in the May 2020 spot period.

show the extent of the price decoupling using non parametric methods.

6.2. Price deviations between CL and QM contracts

We measure the price deviations between the CL and QM contracts across all spot periods from 2011-2020 and show the extent of price decoupling in the May 2020 spot month. Our price deviation statistic is similar to a paired Wilcoxon signed-rank test or a t-test for a difference in means for matched trades in CL and QM. One difference is that a t-test assumes normality while our method does not. Our statistic is as follows:

$$Price\ deviation = \frac{\sum_{i=1}^{n} Q_i * |QMP_i - (\frac{CLP_i^- + CLP_i^+}{2})|}{TotalVolume} \tag{3}$$

where n is the number of trades in the spot period in QM, QMP_i is the price of trade i in QM, Q_i is the quantity of trade i in QM, CLP_i^- is the most recent price of a trade in CL before trade i, CLP_i^+ is the most recent price of a trade in CL after trade i, and Total Volume $=\sum_{i=1}^{n} Q_i$.

We have 120 spot periods between 2011-2020. The largest outlier is the May 2020 contract, with a price deviation assessment of 1.28. The largest outlier has a p-value of 1/120 ($p < 0.01$). The next largest assessment is 0.31 from the February contract of 2016. This supports the claim that the financially-settled contract decoupled from the physically-settled contract in the spot period for crude oil. A complete table of price deviation rankings by spot month are available in table 3 in the appendix.

6.3. Distribution of arbitrageurs

To analyze the frequency of arbitrage trades in WTI crude during the spot month, we concatenate the executing firms with the trading accounts to form trader-ids. We compute benchmark prices in QM by taking the average of the trade immediately before and after in CL. We estimate the amount of arbitrage done by a trader-id. If the price in QM is above the benchmark then we examine the trader-id that sold QM and consider *trade quantity*(trade price – benchmark price)* as the amount of arbitrage for that trade. Similarly, if the price of QM is below the benchmark, then we add an arbitrage quantity to the trader-id that bought QM as *trade quantity*(benchmark price – trade price)*. We then sum all the arbitrage for each trader-id.

We assume $\sum_{i=1}^{n} \frac{(Observed_i - Expected_i)^2}{Expected_i}$ has a Chi-Squared distribution with degrees of freedom $n-1$, where n is the number of trader-ids. We form the observed by computing the amount of arbitrage during the spot period for QM for each trader id. We form the expected by computing the amount of arbitrage over all 120 spot periods and adjusting for the amount of arbitrage in May 2020 by dividing by the sum of the total amount of arbitrage overall and multiplying by the amount of arbitrage in May 2020. This gives us a well-defined expectation. We use this approach because there may be trader-ids for the May 2020 contract that did not trade in any other spot period. This would cause issues with the test statistic, as the numerator would be divided by zero in those cases.

Our Chi-squared statistic is 1,798 with 1,200 degrees of freedom. The p-value is less than 0.0001. This result shows that traders who were performing the arbitrage between QM and CL were different for the May 2020 contract than in other contract months. Indeed, we find that many of the typical arbitragers did not do any trades that we computed as arbitraging on April 20, 2020. This finding has important implications for the smooth functioning of the crude oil market. As we show later, the unusually large values of realized volatility and price spreads observed on April 20 can likely be tied to the lack of typical arbitrageurs in the market. For a complete listing the chi-square statistics for arbitrageur distribution rankings by spot month, see appendix table 4.

As a robustness check to our test statistic, we compute a p-value using the assumption that any ordering of the contract month is equally likely under the null hypothesis because the different contract month test statistics may be serial correlated due to arbitragers moving into and out of market making activities. This lack of independence makes the underlying assumption of the Chi-squared distribution dubious. We note that the two p-values from the different methods are not identical as they are derived from different assumptions. Here the May 2020 contract is the largest outlier with a p-value of 1/120. This provides collaborating evidence that the arbitragers were different for the May 2020 contract. As a further robustness check, we delete the May 2020 contract and reran the chi-square tests with 119 observations. We still have outliers that are significant at 0.0001 but they are substantially smaller (see appendix table 5).

6.4. Spread deviations between Brent and WTI crude oil futures

The Brent crude oil (BZ) futures contract is a close substitute to WTI crude contract (CL) because it is similar in grade, but is a waterborne contract that is cash settled. We analyze the spread between these two contracts to understand how specific delivery issues at Cushing, OK impacted the CL contracts. The choice of the test statistic between Brent and WTI is analogous to a standard deviation calculation around volume-weighted average (VWA), because we want to average over contracts that are traded. Examples of standard deviations being used to measure price volatility in WTI and Brent crude oil include Milonas & Henker (2001). The VWA Spread Proxy functions like a mean in a standard deviation calculation where the dispersion between the price of the two contracts is assessed. We do not use an absolute deviation because the spread between Brent and WTI should vary over time due to economic conditions and not be fixed. Consequently, each spot period has its own VWA spread.

We calculate the spread deviation assessment (SDA) as:

$$SDA = \frac{\sum_{i=1}^{n} Q_i * |Spread\ Proxy_i - VWA\ Spread|}{Total\ Volume} \tag{4}$$

where

$$Spread\ Proxy_i = P_i - \frac{B_i^- + B_i^+}{2} \tag{5}$$

and

$$VWA\ Spread = \sum_{i=1}^{n} Q_i * \frac{Spread\ Proxy_i}{TotalVolume} \tag{6}$$

where n is the number of trades in the spot period, P_i is the price of trade i in WTI crude, Q_i is the quantity of trade i, B_i^- is the most recent trade in Brent after trade i, and $Total\ Volume = \sum_{i=1}^{n} Q_i$.

When we analyze the 120 spot periods between 2011-2020, we find the largest outlier (5.35) to be the May 2020 futures contract in late April of 2020. The largest outlier has a p-value of 1/120 ($p < 0.01$). The next largest deviation is 1.52 in the March contract of 2011, likely caused by political turmoil in the Middle East at the time. All the other deviations are less than one. This gives support for the spread deviation being induced by the WTI contract with physical delivery at Cushing, as the price Brent contract was relatively stable. While storage capacity at Cushing is substantial, approximately 75.8 million barrels according to EIA, the hub itself is landlocked. This contrasts with Brent crude oil, which is a waterborne contract settled around the North Sea. A complete table of spread deviation rankings by spot month can be found in table 6 in the appendix.

6.5. Realized Volatility in the spot month for CL

We measure the realized volatility in each spot period from 2011-2020 and rank them. We use volume-weighted realized volatility as an assessment of how much the futures prices are moving around from trade to trade. This is similar to testing whether the variance in prices in each spot month are equal using a parametric test, such as an F or Chi-square distribution.

We calculate realized volatility as follows:

$$Volume\text{-}Weighted\ Volatilty = \frac{\sum_{i=a}^{n} Q_i * (P_i - P_{i-1})^2}{Total\ Volume} \tag{7}$$

where n is the number of trades in the spot period in CL, P_i is the price of trade i in CL, Q_i is the quantity of matched trade i in CL, the constant a is 1 for the marking period and 2 for the non-marking period, and $Total\ Volume = \sum_{i=1}^{n} Q_i$.

There are 120 spot periods between 2011-2020. The largest outlier has a p-value of 1/120 ($p < 0.01$). In the two-minute marking period for QM and for trade at settlement transactions in CL on the penultimate day, we observe a volume-weighted volatility of 317.19 in the May 2020 contract. The next highest being 0.682 for the April 2020 contract. There were only 135 contracts matched in CL in the marking period, the lowest in our sample. We conclude that the the lack of arbitrageurs in the May 2020 spot period resulted in significantly higher realized volatility. A complete table of weighted volatility rankings by spot month can be found in table 7 in the appendix.

6.6. Trade at settlement analysis

In this section we analyze some of the unusual behavior in the settlement of the CL contract on the penultimate day for the May 2020 contract. A trade at settlement transactions results from a limit order that have been matched in the TAS differential order book which has a price derived from settlement prices. The differentials in the order book range from \$-0.10 to \$0.10 in outrights and \$-0.20 to \$0.20 in spreads in CL. A trade at settlement transactions results from a limit order that have been matched in the TAS differential order book which has a price derived from settlement prices which occurs at the end of the trading day.

To compute the settlement price, the CME uses the spread between the active contract and the spot contract and the settlement price in active contract in CL. The active contract is the next calendar month from the spot month. The exchange computes the volume-weighted average price during [14:28,14:30] for regular outright and implied outright transactions in the active contract. The exchange computes the spread between the spot month contract using the regular spread and implied spread transactions using the volume-weighted average price during [14:28,14:30] with the spot month when available. We observed it in each spot month of our sample for CL.

There was some unusual behavior in the settlement of the CL contract on the penultimate day for the May 2020 contract. There was the largest number of futures contracts settling to the settlement price on the penultimate day since any contract month from 2011-2020. There were over 103,734 total outrights (both buys and sells) trade-at-settlement in CL on April 20. There were 50,418 spread trade-at-settlement (both buys and sells), and 5,141 QM contracts (QM is a half notional contract) that went to expiration and cash settled. This is the largest amount of futures transactions settling on the penultimate day of the contract, ranking 120 out of 120 with the next largest being November 15, 2012 (December 2012 contract) with 92,186 outright TAS, 7,948 spread TAS, and 9,299 QM contracts that went to expiration. There are also derivative contracts such as balance of the month futures contracts, swaps and forwards that referenced the closing price of CL, too.

The CL contract also saw the lowest trading volume both in the spread between the spot and the active (285 contracts), and the spot (135 contracts). The active contract did not see a record low volume (3246 contracts). Further, the spread contract between the spot and active contract was not that volatile in the May 2020 contract month. However, the spot and the active contract month were the most volatile contract months.[9] This means that liquidity rarely crossed between the spot and active contract even though the realized volatility was the highest ever observed in both contracts.[10] Only 135 contracts were matched between the active and spot month by spreads. This means that the price of limit orders between spot and active contract moved down in near lock step, otherwise more limited orders would have matched between the active and spot contract.

Finally, the largest amount of linked trading between TAS and the outrights was observed, summed by trader-ids. The May 2020 has a linked trade open interest that is 25% more than next largest outlier and 2.7 times larger than the median. Linked-trade open interest is calculated as follows:

$$\begin{aligned} Linked\,Trades = MIN[\sum_{k=0}^{n} Q(Outright_{sell}), \sum_{k=0}^{m} TAS_{buy}] \\ + MIN[\sum_{k=0}^{l} Q(Outright_{buy}), \sum_{k=0}^{p} Q(TAS_{sell})] \end{aligned} \tag{8}$$

[9]The first six months of 2011 lack descriptors for both TAS and specific calendar spreads. Consequently, the relative frequencies are out of 114 months with spread and TAS transaction test statistics. The volatility of the spread between the spot and active month ranked 89 out of 114 which is not statistically significant at the 5% level.

[10]The CME has limit order functionality that allows an outright in the spot month and a spread between the active and spot month to match against an outright in the active month (or visa versa).

This demonstrates that traders executed TAS transactions where the price was to be determined by settlement at the end of the day and traded in the opposite direction in the outright contract. CFTC Commissioner Berkovitz has criticized not bounding the amount of linked trades. He asserts that since the price of the TAS trade isn't determined, it gives parties with large TAS positions an incentive to mark the settlement period to benefit their TAS transactions.[11]

6.7. Comparison of WTI and Natural Gas Markets in May 2020 spot month

Our last empirical analysis examines whether the storage problems observed in WTI crude oil were isolated or whether another energy market, such as natural gas, experienced a similar capacity problem. We estimate a difference-in-differences model to test whether the decoupling between the cash and physically settled contracts in WTI crude oil was observed in the natural gas futures prices as well. Our null hypothesis is that this event was isolated to WTI crude oil markets connected to the pipeline at Cushing.

To test the suitability of a difference-in-differences model, we first examine the trends in differences between the cash settled and physically settled prices for both natural gas and WTI crude oil. Our parallel trend analysis is shown in appendix figure 5. The two sets of lines have a similar trend around zero until our negative price event in April 2020. While there is more noise in the WTI crude oil price series, the overall trend is consistent with the natural gas series.

We note that the contracts for these two commodities expire at different times in the month, so we compare the prices in their respective spot months. As such our analysis is in event-time, rather than chronological time.[12] The time frame for this analysis runs from December 2015 to December 2020. We have five years of data with 12 spot months per year, giving us 60 spot months. There are 72 hours in each spot month, giving us a total possible number of

[11] `https://www.cftc.gov/PressRoom/SpeechesTestimony/berkovitzstatement031521`

[12] NYMEX CL trading terminates 3 business days prior to the 25th calendar day of the month prior to the contract month. NYMEX NG trading terminates on the 3rd last business day of the month prior to the contract month.

observations of 4,320. However, because hourly trades are not always observed in both contracts, our dataset has 2,890 observations.

We specify a difference-in-differences model with spot month fixed effects as follows:

$$y_{it} = \sum_{m=1}^{M-1} \alpha_m D_{m,it} + \beta_0 D_{may20,it} + \beta_1 D_{WTI,it} + \beta_2 D_{WTI,it} D_{may20,it} + \epsilon_{it} \tag{9}$$

where y_{it} is the difference in the hourly VWAP between the cash settled and physically settled prices in natural gas or WTI crude oil contracts, where the commodity type (i.e. natural gas or WTI crude oil) is denoted by i, the hour is denoted by t, and spot month denoted by m. $D_{m,it}$ is a dummy variable for the mth spot month, leaving out May 2020. $D_{may20,it}$ is a dummy variable, equal to one in the May 2020 spot month, and zero otherwise. Finally, the interaction of the dummy variables $D_{WTI,it}$ and $D_{may20,it}$ is equal to one when the commodity is WTI crude oil and the spot month is May 2020, and zero otherwise. The coefficient on this interaction term, β_2 is the treatment effect we wish to measure.

We scrutinize the parameter β_2, which assesses the the average difference-in-differences between the physically settled and cash settled contracts of Natural Gas and WTI crude oil in the May 2020 spot period. The estimate for β_2 is -0.218 and has a p-value of less than 0.001. This finding indicates that there was a significant difference in the average price between the cash and physically-settled contracts in WTI crude oil, compared with Natural gas, during the May 2020 spot month. As a robustness check, we conduct a series of falsification tests using the four spot months before and after May 2020. These tests do not find a significant difference in prices between the pair of contracts in WTI crude oil and Natural Gas during any other spot periods.

The difference-in-differences model results confirm storage problems with WTI crude likely played a role in negative price event. We note that natural gas has seasonal storage variability with lower capacity in fall prior to the winter months with higher capacity in the Spring because natural gas is drawn down during the cold winter months and consumed. Consequently, there was plentiful natural gas storage available to a long position holder taking delivery at Henry Hub with the May 2020 contact in late April.

7. Summary and Discussion

The negative price settlement in the WTI crude oil futures contract on April 20, 2020 was a historic and unprecedented event. Our study leverages CFTC supervisory data to show that the typical arbitrageurs were not present in the market on that date, leading to a substantial loss of liquidity. Our study uses generalizable non-parametric methods to assess and rank the price movements, volatility, and spreads with similar crude oils observed on April 20 with the past 10 years of spot periods. We demonstrate that the NYMEX CL and QM contracts, usually linked by arbitrage, decoupled on that date. Consistent with previous studies, we find that storage constraints at Cushing affected the WTI-Brent spread, finding the spread was significantly larger than usual on April 20. We confirm the lack of storage at Cushing as a contributing factor in the negative price event by comparing the cash-physically settled price spread in WTI crude oil futures with natural gas futures.

We identify two policy areas that might prevent a similar market event in WTI crude futures, specifically, 1) increasing market liquidity with designated market makers (DMMs), and, 2) changing the contract specifications to include variable storage rates or additional delivery locations. Our suggestions seek to improve liquidity and facilitate better cash and futures price convergence at expiration. Better price convergence accommodates the hedging of price risk by commercial entities in the physical marketing channel.[13]

Academic research has shown that DMMs can improve market liquidity and

[13] We note that futures exchanges institute their policies and procedures absent objection from the CFTC when they certify that they have satisfied core principles stipulated under the Commodity Exchange Act. As such, much of the discretion lies with the futures exchange listing a contract for trading.

decrease transaction costs (Tse & Zabotina 2004; Clark-Joseph et al. 2017). Many equities and options exchanges already have DMMs that stand ready to buy and sell stocks listed on the exchange, including the London Stock Exchange, Euronext, and the New York Stock Exchange.[14] The exchange could also create incentives for market makers to provide liquidity. This could be accomplished through rebates or discounts for participants who have matched passive limit orders during the marking period in the active contract or with the spread between the active and spot contract. We note that making markets during the spot period for a physical commodity may require contingency planning for making or taking delivery.

Storage constraints affect the price of deliverable futures contracts. Adding a variable or market-based storage rate to the WTI contract might help with cash and futures price convergence issues. Studies by (Irwin, 2020) and (Garcia et al., 2015) found that fixed storage rates used in grain contracts in the mid-2000s contributed to convergence issues between cash and futures prices, particularly in wheat markets. These convergence issues were addressed when the Chicago Mercantile Exchange (CME) introduced variable storage rates on the wheat contracts. Variable storage rates are a market-based determinant of the maximum allowable storage rates for outstanding wheat shipping certificate. It triggers higher maximum allowable storage rates that allow wide spreads when spreads are near full carry, while also allowing lower maximum storage rates when spreads are narrow or inverted. We note that delivery of WTI crude is slightly different than grain contracts because it happens one week to ten days hence, not immediately in the spot period, due to the need to schedule pipelines, etc.

Another potential enhancement to confront storage issues with WTI is to add more delivery options. The WTI-Houston contract was introduced by CME in 2018. It has same chemical specification for crude oil as WTI-Cushing but delivery is at the port of Houston. Houston has the potential to serve as an

[14] We note that participation in futures markets is voluntary for traders.

alternative delivery location (with commensurate differentials) that could allow longs additional ability to take delivery of WTI if they could elect this location under the contract after delivery in Cushing because crude oil may be loaded into tankers or placed into nearby storage facilities.

References

Adelman, M. A. (1984). International oil agreements. *The Energy Journal*, *5*. doi:10.5547/ISSN0195-6574-EJ-Vol5-No3-1.

Baur, D. G. (2013). The structure and degree of dependence: A quantile regression approach. *Journal of Banking & Finance*, *37*, 786–798. doi:10.1016/j.jbankfin.2012.10.015.

Bianchi, R. J., Fan, J. H., & Todorova, N. (2020). Financialization and de-financialization of commodity futures: A quantile regression approach. *International Review of Financial Analysis*, *68*, 101451. doi:10.1016/j.irfa.2019.101451.

Buncic, D., & Gisler, K. I. (2017). The role of jumps and leverage in forecasting volatility in international equity markets. *Journal of international Money and Finance*, *79*, 1–19. doi:10.1016/j.jimonfin.2017.09.001.

Buyuksahin, B., Lee, T. K., Moser, J. T., & Robe, M. A. (2013). Physical markets, paper markets and the wti-brent spread. *The Energy Journal*, *34*. doi:10.5547/01956574.34.3.

Cabrera, J., Wang, T., & Yang, J. (2009). Do futures lead price discovery in electronic foreign exchange markets? *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *29*, 137–156. doi:10.1002/fut.20352.

Chen, Z., Ye, Y., & Li, X. (2022). Forecasting china's crude oil futures volatility: New evidence from the midas-rv model and covid-19 pandemic. *Resources Policy*, *75*, 102453. doi:10.1016/j.resourpol.2021.102453.

Cheng, I.-H., & Xiong, W. (2014). Financialization of commodity markets. *Annu. Rev. Financ. Econ.*, *6*, 419–441. doi:10.1146/annurev-financial-110613-034432.

Clark-Joseph, A. D., Ye, M., & Zi, C. (2017). Designated market makers still matter: Evidence from two natural experiments. *Journal of Financial Economics*, *126*, 652–667. doi:10.1016/j.jfineco.2017.09.001.

Ederington, L. H., Fernando, C. S., Lee, T. K., Linn, S. C., & Zhang, H. (2021). The relation between petroleum product prices and crude oil prices. *Energy Economics*, *94*, 105079. doi:10.1017/S0022109020000204.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM.

Elder, J., Miao, H., & Ramchander, S. (2013). Jumps in oil prices: the role of economic news. *The Energy Journal*, *34*. doi:10.5547/01956574.34.3.10.

Elliott, G. (1999). Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. *International Economic Review*, *40*, 767–784. doi:10.1111/1468-2354.00039.

Eom, C., Kaizoji, T., Livan, G., & Scalas, E. (2021). Limitations of portfolio diversification through fat tails of the return distributions: Some empirical evidence. *The North American Journal of Economics and Finance*, *56*, 101358. doi:10.1016/j.najef.2020.101358.

Eom, C., Kaizoji, T., & Scalas, E. (2019). Fat tails in financial return distributions revisited: Evidence from the korean stock market. *Physica A: Statistical Mechanics and its Applications*, *526*, 121055. doi:10.1016/j.physa.2019.121055.

Fernandez-Perez, A., Fuertes, A.-M., & Miffre, J. (2020). Understanding the negative pricing of the nymex wti crude oil may 2020 futures contract. *Available at SSRN 3748321*, . doi:10.2139/ssrn.3748321.

Galay, G. (2019). Are crude oil markets cointegrated? testing the co-movement of weekly crude oil spot prices. *Journal of Commodity Markets*, *16*, 100088. doi:10.1016/j.jcomm.2019.02.003.

Garcia, P., Irwin, S. H., & Smith, A. (2015). Futures market failure? *American Journal of Agricultural Economics*, *97*, 40–64. doi:10.1093/ajae/aau067.

Gilje, E. P., Ready, R., Roussanov, N., & Taillard, J. P. (2020). The day that wti died: Asset prices and firm production decisions. *Working Paper*, .

Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, *13*, 27–35. doi:10.2307/1392518.

Gulen, S. G. (1997). Regionalization in the world crude oil market. *The Energy Journal*, *18*. doi:10.5547/ISSN0195-6574-EJ-Vol18-No2-6.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance*, *50*, 1175–1199. doi:10.1111/j.1540-6261.1995.tb04054.x.

Hiemstra, C., & Jones, J. D. (1994). Testing for linear and nonlinear granger causality in the stock price-volume relation. *The Journal of Finance*, *49*, 1639–1664. doi:10.1111/j.1540-6261.1994.tb04776.x.

Irwin, S. H. (2020). Trilogy for troubleshooting convergence: Manipulation, structural imbalance, and storage rates. *Journal of Commodity Markets*, *17*, 100083. doi:10.1016/j.jcomm.2018.11.002.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254. doi:10.1016/0165-1889(88)90041-3.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, (pp. 1551–1580). doi:10.2307/2938278.

Kane, S. (2004). Scientific methods in finance. *International Review of Financial Analysis*, *13*, 105–118. doi:10.1016/j.irfa.2004.01.003.

Kawaller, I. G., Koch, P. D., & Koch, T. W. (1987). The temporal price relationship between s&p 500 futures and the s&p 500 index. *The Journal of Finance*, *42*, 1309–1329. doi:10.1111/j.1540-6261.1987.tb04368.x.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, *15*, 143–156. doi:10.1257/jep.15.4.143.

Li, T., Qian, Z., Deng, W., Zhang, D., Lu, H., & Wang, S. (2021). Forecasting crude oil prices based on variational mode decomposition and random sparse bayesian learning. *Applied Soft Computing*, *113*, 108032. doi:10.1016/j.asoc.2021.108032.

Liu, W.-M., Schultz, E., & Swieringa, J. (2015). Price dynamics in global crude oil markets. *Journal of Futures Markets*, *35*, 148–162. doi:10.1002/fut.21658.

Maritz, J. S. (1995). *Distribution-Free Statistical Methods* volume 17. CRC Press.

Miao, H., Ramchander, S., Wang, T., & Yang, D. (2017). Influential factors in crude oil price forecasting. *Energy Economics*, *68*, 77–88. doi:10.1016/j.eneco.2017.09.010.

Miao, H., Ramchander, S., Wang, T., & Yang, J. (2018). The impact of crude oil inventory announcements on prices: Evidence from derivatives markets. *Journal of Futures Markets*, *38*, 38–65. doi:10.1002/fut.21850.

Milonas, N. T., & Henker, T. (2001). Price spread and convenience yield behaviour in the international oil market. *Applied Financial Economics*, *11*, 23–36. doi:10.1080/09603100150210237.

Schwarz, T. V., & Szakmary, A. C. (1994). Price discovery in petroleum markets: Arbitrage, cointegration, and the time interval of analysis. *The Journal of Futures Markets (1986-1998)*, *14*, 147. doi:10.1002/fut.3990140204.

Silvapulle, P., & Moosa, I. A. (1999). The relationship between spot and futures prices: evidence from the crude oil market. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *19*, 175–193.

Troster, V., Shahbaz, M., & Uddin, G. S. (2018). Renewable energy, oil prices, and economic activity: A granger-causality in quantiles analysis. *Energy Economics*, *70*, 440–452. doi:10.1016/j.eneco.2018.01.029.

Tse, Y., & Xiang, J. (2005). Market quality and price discovery: Introduction of the e-mini energy futures. *Global Finance Journal*, *16*, 164–179. doi:10.1016/j.gfj.2005.04.001.

Tse, Y., & Zabotina, T. (2004). Do designated market makers improve liquidity in open-outcry futures markets? *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *24*, 479–502. doi:10.1002/fut.10117.

Weiner, R. J. (1991). Is the world oil market "one great pool"? *The Energy Journal*, *12*. doi:10.5547/ISSN0195-6574-EJ-Vol12-No3-7.

Yang, D.-X., Wu, B.-B., & Tong, J.-Y. (2021). Dynamics and causality of oil price shocks on commodities: Quantile-on-quantile and causality-in-quantiles methods. *Resources Policy*, *74*, 102246. doi:10.1016/j.resourpol.2021.102246.

Appendix

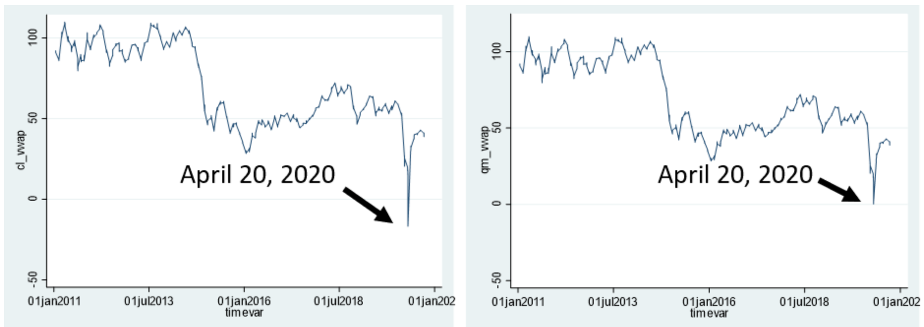

Figure 3: Time series of volume-weighted average prices in CL (left graph) and QM (right graph) contracts, measured at the 10-minute level during the spot month, 2011-2020

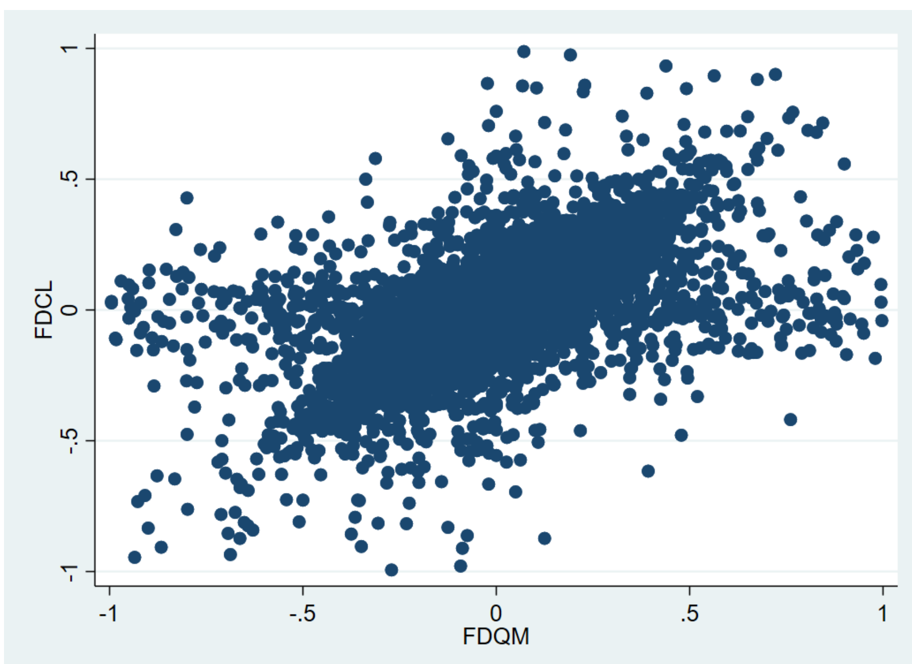

Figure 4: Joint plot of first-differences in prices for CL and QM contracts, 2011-2020 spot months

Contract date	L1norm	Rank	Contract date	L1norm	Rank	Contract date	L1norm	Rank	Contract date	L1norm	Rank
202101	0.011386861	1	201612	0.035565	31	201405	0.059372	61	201204	0.093838	91
202011	0.0124	2	201401	0.03557	32	201608	0.059476	62	202006	0.09434	92
202012	0.014051095	3	201806	0.036585	33	201709	0.059507	63	201809	0.096742	93
201905	0.018856502	4	201407	0.037305	34	201812	0.060399	64	201909	0.098344	94
202009	0.019918033	5	202003	0.037615	35	201708	0.060646	65	201511	0.10528	95
201303	0.02035061	6	201302	0.038104	36	201601	0.062014	66	201501	0.109456	96
201304	0.023	7	201712	0.039301	37	201907	0.062365	67	201911	0.111	97
202008	0.024104478	8	201208	0.040117	38	201903	0.062368	68	201505	0.11455	98
201711	0.02474359	9	201908	0.040376	39	201102	0.063054	69	201502	0.114648	99
201801	0.025033557	10	201803	0.041259	40	201411	0.063555	70	202004	0.122788	100
201910	0.025573248	11	201205	0.042086	41	201610	0.065073	71	201811	0.123801	101
201312	0.026904025	12	202002	0.042303	42	201512	0.0665	72	201912	0.125559	102
201707	0.027	13	201412	0.04325	43	201108	0.068053	73	201509	0.126057	103
201710	0.028017621	14	201802	0.043695	44	201107	0.068228	74	201503	0.131403	104
201202	0.029883041	15	201206	0.044681	45	201311	0.068804	75	201805	0.133134	105
201705	0.030286344	16	201207	0.046297	46	202007	0.071168	76	201603	0.135947	106
201309	0.030522088	17	201106	0.047851	47	201506	0.071892	77	201605	0.138004	107
201703	0.030526316	18	201110	0.048044	48	201607	0.071934	78	201409	0.140439	108
201402	0.030657895	19	201508	0.048177	49	201404	0.072151	79	201810	0.14192	109
201704	0.030970149	20	201212	0.048967	50	201103	0.076561	80	201609	0.152262	110
201611	0.031065574	21	201109	0.049126	51	201901	0.078592	81	201510	0.153573	111
201307	0.031727941	22	201201	0.050666	52	201804	0.078641	82	201112	0.15409	112
201209	0.03253125	23	201807	0.051582	53	201902	0.078885	83	201808	0.161614	113
201904	0.033140097	24	201701	0.05188	54	201410	0.078945	84	201604	0.170786	114
201211	0.033670382	25	201306	0.052279	55	201308	0.079122	85	201403	0.183289	115
201906	0.033891129	26	201301	0.052592	56	201706	0.079974	86	201504	0.197605	116
201305	0.034237968	27	201507	0.052668	57	201702	0.083821	87	201310	0.221826	117
201111	0.034477612	28	201203	0.052795	58	201406	0.08676	88	201104	0.243184	118
201408	0.034723127	29	202010	0.057407	59	201105	0.088364	89	201602	0.305682	119
202001	0.0348659	30	201606	0.05787	60	201210	0.088421	90	202005	1.279365	120

Table 3: Rankings of price deviations between QM and CL by spot month, 2011-2020

Chisq stat	Expiration	Rank	Chisq stat	Expiration	Rank	Chisq stat	Expiration	Rank	Chisq stat	Expiration	Rank
12.508	01Nov2017	1	172.646	01Aug2016	31	377.630	01Mar2015	61	590.605	01Aug2017	91
28.579	01Mar2018	2	179.470	01Jan2020	32	379.997	01Nov2013	62	591.381	01Aug2019	92
28.840	01Jan2017	3	183.730	01Jun2019	33	380.498	01Feb2018	63	615.785	01Nov2011	93
28.921	01Jul2015	4	189.256	01Jan2021	34	381.421	01Jun2020	64	622.959	01Jun2016	94
40.992	01May2019	5	197.074	01May2018	35	384.154	01Jun2011	65	637.372	01Oct2011	95
43.081	01Nov2018	6	208.415	01Apr2018	36	384.777	01Jan2013	66	641.317	01Apr2013	96
43.754	01Mar2017	7	211.438	01Jan2015	37	386.811	01Jun2017	67	643.520	01Jun2013	97
47.132	01Dec2017	8	225.451	01May2013	38	394.665	01Nov2019	68	680.001	01Sep2017	98
50.330	01Apr2014	9	226.852	01Jul2014	39	395.323	01Sep2014	69	735.782	01Nov2012	99
54.705	01Dec2018	10	228.427	01Jan2016	40	395.740	01Sep2015	70	789.767	01Jan2019	100
54.770	01Apr2019	11	238.781	01Sep2013	41	397.087	01Feb2014	71	807.589	01Dec2012	101
58.531	01Jul2020	12	246.627	01Dec2014	42	407.042	01Jun2012	72	817.386	01Mar2020	102
62.285	01Dec2016	13	257.851	01Sep2012	43	410.633	01May2017	73	853.050	01May2011	103
62.877	01Oct2017	14	267.369	01Nov2016	44	411.158	01Jun2018	74	891.079	01Oct2012	104
66.735	01Jan2018	15	271.006	01May2012	45	416.663	01Oct2019	75	897.381	01Apr2020	105
69.471	01Jul2018	16	279.088	01Jun2014	46	453.136	01Jul2017	76	930.270	01Jul2011	106
71.456	01Aug2015	17	284.251	01Oct2014	47	458.300	01Dec2019	77	971.113	01Sep2011	107
86.049	01Sep2018	18	287.108	01May2016	48	464.249	01Sep2019	78	1004.031	01Jan2014	108
99.495	01Oct2016	19	291.867	01Aug2020	49	469.752	01May2015	79	1051.812	01Mar2013	109
106.985	01Apr2016	20	292.442	01Oct2018	50	473.743	01Oct2013	80	1113.703	01Feb2011	110
115.031	01Aug2018	21	308.913	01Aug2014	51	492.017	01Mar2014	81	1187.400	01Jul2019	111
117.600	01Nov2020	22	310.005	01Nov2014	52	502.374	01Dec2020	82	1187.642	01Mar2019	112
123.234	01May2014	23	314.324	01Aug2013	53	508.351	01Aug2012	83	1220.312	01Sep2016	113
138.381	01Jul2013	24	321.917	01Mar2016	54	518.571	01Jul2012	84	1260.841	01Apr2017	114
139.273	01Dec2013	25	326.036	01Jan2012	55	528.349	01Feb2013	85	1322.684	01Dec2011	115
139.287	01Jul2016	26	345.215	01Feb2016	56	537.369	01Feb2019	86	1435.453	01Mar2011	116
156.544	01Dec2015	27	350.659	01Oct2015	57	562.531	01Apr2015	87	1570.360	01Apr2011	117
159.544	01Sep2020	28	355.658	01Mar2012	58	570.103	01Apr2012	88	1575.776	01Oct2020	118
160.354	01Nov2015	29	369.105	01Jun2015	59	578.202	01Feb2020	89	1618.960	01Aug2011	119
166.662	01Feb2017	30	373.853	01Feb2012	60	589.393	01Feb2015	90	1798.357	01May2020	120

Table 4: Rankings of Chi-square test statistic for frequency distribution of arbitrageurs by spot month, 2011-2020

Chisq stat	Expiration	Rank	Chisq stat	Expiration	Rank	Chisq stat	Expiration	Rank	Chisq stat	Expiration	Rank
9.196	01Nov2017	1	135.392	01Jan2020	31	286.378	01Nov2013	61	463.927	01Aug2019	91
20.340	01Jul2015	2	139.500	01Jun2019	32	287.706	01Feb2018	62	464.146	01Nov2011	92
20.834	01Mar2018	3	145.253	01Jan2021	33	289.987	01Jan2013	63	481.505	01Oct2011	93
23.475	01Jan2017	4	148.836	01Jan2015	34	290.559	01Jun2011	64	483.023	01Jun2016	94
31.274	01Nov2018	5	150.291	01May2018	35	292.138	01Nov2016	65	484.433	01Jun2013	95
31.534	01May2019	6	161.263	01Apr2018	36	293.065	01Jun2017	66	485.702	01Apr2013	96
31.563	01Mar2017	7	170.466	01May2013	37	295.061	01Sep2015	67	523.618	01Sep2017	97
36.323	01Apr2014	8	170.703	01Jan2016	38	296.880	01Sep2014	68	555.927	01Nov2012	98
39.976	01Dec2018	9	170.866	01Jul2014	39	300.675	01Feb2014	69	604.646	01Jan2019	99
44.546	01Apr2019	10	179.426	01Sep2013	40	305.117	01Jun2012	70	609.533	01Dec2012	100
46.028	01Dec2016	11	182.667	01Dec2014	41	307.685	01Nov2019	71	624.831	01Mar2020	101
50.289	01Jan2018	12	191.441	01Sep2020	42	311.471	01Jun2018	72	644.286	01May2011	102
53.368	01Aug2015	13	194.277	01Sep2012	43	317.836	01May2017	73	664.978	01Oct2012	103
54.419	01Jul2018	14	203.246	01May2012	44	328.263	01Oct2019	74	704.027	01Jul2011	104
55.407	01Oct2017	15	209.567	01Jun2014	45	345.533	01Oct2013	75	727.091	01Sep2011	105
55.897	01Dec2017	16	210.048	01May2016	46	349.684	01Sep2019	76	753.798	01Apr2020	106
65.108	01Sep2018	17	210.078	01Oct2014	47	351.922	01Jul2017	77	761.535	01Jan2014	107
74.351	01Oct2016	18	218.873	01Oct2018	48	352.503	01May2015	78	764.652	01Feb2020	108
75.406	01Apr2016	19	228.904	01Aug2020	49	361.550	01Mar2014	79	796.398	01Mar2013	109
79.969	01Jul2020	20	230.802	01Nov2014	50	381.434	01Aug2012	80	844.815	01Feb2011	110
84.980	01Aug2018	21	233.211	01Aug2014	51	390.296	01Jul2012	81	901.243	01Mar2019	111
90.110	01May2014	22	234.213	01Aug2013	52	393.187	01Dec2020	82	904.048	01Jul2019	112
90.347	01Nov2020	23	237.263	01Mar2016	53	399.847	01Feb2013	83	920.950	01Sep2016	113
103.235	01Jul2016	24	244.287	01Jan2012	54	406.105	01Feb2019	84	960.464	01Apr2017	114
103.859	01Jul2013	25	254.347	01Feb2016	55	409.596	01Apr2015	85	984.497	01Dec2011	115
103.931	01Dec2013	26	259.976	01Oct2015	56	409.602	01Dec2019	86	1089.044	01Mar2011	116
116.019	01Dec2015	27	266.815	01Mar2012	57	426.279	01Apr2012	87	1178.158	01Apr2011	117
120.867	01Nov2015	28	274.933	01Jun2015	58	427.878	01Feb2015	88	1196.820	01Oct2020	118
127.865	01Feb2017	29	276.851	01Mar2015	59	447.992	01Aug2017	89	1224.106	01Aug2011	119
128.849	01Aug2016	30	281.668	01Feb2012	60	455.420	01Jun2020	90			

Table 5: Rankings of Chi-square test statistic for frequency distribution of arbitrageurs by spot month without May 2020, 2011-2020

Contract date	VWASD	Rank	Contract date	VWASD	Rank	contract date	VWASD	rank	contract date	VWASD	rank
201802	0.073	1	201307	0.170	31	201604	0.281	61	201108	0.447	91
202008	0.078	2	201803	0.173	32	201209	0.285	62	201503	0.451	92
201708	0.083	3	201601	0.179	33	201811	0.291	63	201403	0.453	93
202012	0.087	4	201701	0.185	34	201402	0.293	64	201602	0.468	94
201706	0.095	5	201608	0.187	35	201211	0.298	65	201312	0.475	95
201711	0.095	6	202007	0.189	36	201412	0.310	66	201210	0.477	96
202001	0.097	7	201704	0.189	37	201805	0.313	67	201502	0.478	97
202101	0.100	8	201709	0.190	38	201910	0.321	68	201410	0.484	98
201801	0.102	9	201905	0.195	39	201901	0.328	69	201106	0.488	99
201707	0.106	10	201908	0.198	40	201907	0.348	70	201308	0.490	100
202009	0.117	11	201609	0.198	41	201406	0.350	71	201309	0.493	101
202002	0.118	12	201906	0.198	42	201310	0.353	72	201207	0.500	102
201904	0.122	13	201809	0.203	43	201505	0.355	73	202006	0.509	103
201612	0.125	14	201607	0.208	44	201401	0.355	74	201110	0.514	104
202010	0.126	15	201506	0.209	45	201304	0.357	75	201102	0.527	105
201911	0.128	16	201611	0.212	46	201405	0.358	76	201409	0.553	106
201705	0.129	17	201411	0.219	47	201202	0.368	77	201504	0.579	107
201606	0.137	18	202003	0.221	48	201810	0.372	78	201206	0.584	108
201703	0.141	19	201201	0.231	49	201512	0.373	79	201407	0.596	109
201804	0.144	20	201305	0.240	50	201105	0.373	80	201107	0.661	110
201702	0.144	21	201306	0.242	51	201303	0.384	81	201408	0.678	111
201610	0.155	22	201712	0.243	52	201111	0.387	82	201404	0.715	112
201710	0.157	23	201603	0.249	53	201104	0.397	83	201311	0.754	113
201909	0.158	24	201208	0.251	54	201501	0.400	84	201203	0.778	114
201902	0.161	25	201605	0.252	55	201302	0.401	85	201109	0.823	115
201812	0.164	26	201212	0.260	56	201807	0.404	86	201205	0.823	116
201508	0.166	27	201301	0.262	57	201806	0.417	87	201112	0.885	117
201903	0.167	28	201511	0.269	58	201808	0.429	88	202004	0.949	118
201912	0.169	29	201509	0.276	59	201204	0.431	89	201103	1.515	119
202011	0.169	30	201507	0.279	60	201510	0.446	90	202005	5.346	120

Table 6: Rankings of volume weighted average spread deviation (VWASD) between CL and Brent prices by spot month, 2011-2020

Contract date	Weighted volatility	Rank	Contract date	Weighted volatility	Rank	Contract date	Weighted volatility	Rank	Contract date	Weighted volatility	Rank
201802	7.25389E-06	1	201111	8.01E-05	31	201601	0.003195	61	201105	0.021341	91
201611	7.50148E-06	2	201411	8.25E-05	32	201309	0.003454	62	201712	0.02167	92
201709	9.32551E-06	3	202009	0.000112	33	201207	0.003512	63	201208	0.021711	93
201702	9.43396E-06	4	201405	0.000123	34	201610	0.003546	64	201706	0.022826	94
201701	1.02715E-05	5	201512	0.000135	35	201301	0.003725	65	201108	0.023353	95
201710	1.04228E-05	6	201107	0.000168	36	201506	0.003901	66	201504	0.025388	96
201705	1.13171E-05	7	201302	0.000183	37	201804	0.004378	67	201510	0.025919	97
201609	1.26924E-05	8	201412	0.000205	38	201608	0.004934	68	201603	0.027261	98
201607	1.6122E-05	9	201201	0.000332	39	201911	0.005195	69	201605	0.033797	99
201703	1.62269E-05	10	201904	0.000494	40	201805	0.005436	70	202003	0.034788	100
201312	1.91431E-05	11	202008	0.000551	41	201708	0.006327	71	202007	0.03499	101
202011	1.95704E-05	12	201402	0.000578	42	201106	0.00652	72	201810	0.035672	102
202012	2.24832E-05	13	201612	0.000823	43	201410	0.006679	73	201807	0.040835	103
201303	2.45127E-05	14	201907	0.000866	44	201306	0.007092	74	201409	0.04246	104
202001	2.75735E-05	15	201308	0.000888	45	201908	0.008113	75	202006	0.046626	105
202002	3.04498E-05	16	201307	0.001182	46	201903	0.008909	76	201104	0.071033	106
201212	3.24459E-05	17	201707	0.00131	47	201806	0.009388	77	201503	0.076135	107
201204	3.53168E-05	18	201202	0.001393	48	201403	0.009787	78	201304	0.079731	108
201110	3.59467E-05	19	201803	0.001405	49	201401	0.010521	79	201211	0.13362	109
201406	3.66864E-05	20	201407	0.001451	50	201511	0.010704	80	201310	0.201667	110
201209	3.84346E-05	21	201809	0.001454	51	201602	0.010957	81	201210	0.207983	111
202010	4.26667E-05	22	201205	0.001584	52	201408	0.011594	82	201502	0.241732	112
202101	4.53488E-05	23	201905	0.001803	53	201902	0.012628	83	201912	0.242358	113
201808	4.56693E-05	24	201206	0.00188	54	201909	0.012825	84	201509	0.252519	114
201910	4.91289E-05	25	201102	0.002251	55	201811	0.013262	85	201501	0.321911	115
201606	6.0559E-05	26	201103	0.002257	56	201404	0.01536	86	201311	0.321919	116
201109	6.32862E-05	27	201711	0.002315	57	201604	0.016411	87	201901	0.322894	117
201801	6.38535E-05	28	201508	0.002338	58	201505	0.017816	88	201112	0.357859	118
201906	7.08273E-05	29	201704	0.002621	59	201203	0.018133	89	202004	0.681548	119
201812	7.65468E-05	30	201305	0.002631	60	201507	0.020957	90	202005	317.1896	120

Table 7: Rankings of weighted volatility in CL contract by spot month, 2011-2020

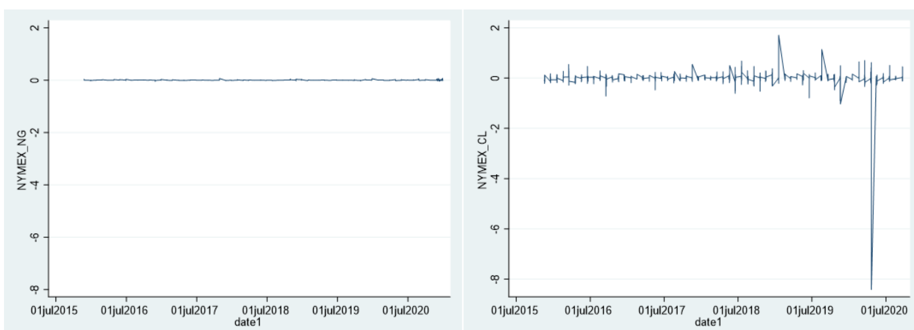

Figure 5: Time series of hourly volume-weighted average price differences between the cash-settled and physically settled contracts in NYMEX NG (Natural Gas) (Left) and NYMEX CL (WTI crude oil) (Right) during the spot month, 2015-2020