Closer to One Great Pool? Evidence from Structural Breaks in Oil Price Differentials

Michael Plante^a and Grant Strickler^b

ABSTRACT

We show that the oil market has become closer to "one great pool," in the sense that price differentials between crude oils of different qualities have generally become smaller over time. We document, in particular, that many of these price differentials experienced a major structural break in or around 2008, after which there was a marked reduction in their means and volatilities. Differentials between residual fuel oil, a low-quality fuel, and higher-valued products, such as gasoline and diesel, experienced similar breaks during the same time period. A growing ability of the global refinery sector to process lower-quality crude oil and the U.S. shale boom, which has unexpectedly boosted the supply of high-quality crude oil, are two factors consistent with these changes. Differentials between crude oils of similar quality in general did not experience breaks in or around 2008, although we do find evidence of breaks at other times.

Keywords: Crude oil price differentials, Oil, Structural breaks, Refining

https://doi.org/10.5547/01956574.42.2.mpla

1. INTRODUCTION

Morris Adelman (1984) famously wrote "The world oil market, like the world ocean, is one great pool." If this were literally true, it would imply that all crude streams would be perfectly substitutable for one another in the refining process. We would then expect to see generally small price differentials between different crude streams over the long-run, reflecting primarily transportation costs. In reality, crude oil can have a wide range of physical properties that play an important role in how one crude stream is priced relative to another, and one often observes large price differentials between crude streams of different qualities.

Price differentials between different grades of crude are important for many oil market participants. For refiners, they can affect profitability and influence investment decisions about specific equipment, such as cokers, that could improve the profitability of processing lower grades of crude.¹ Oil producers and fiscal authorities are concerned about these differentials because of the impacts they can have on revenues earned from producing or taxing certain types of oil.² They can also affect

- 1. This topic has received attention from trade press and market analysts since at least the early 2000s. More recently, the shale boom and IMO 2020 have renewed interest in these issues. See, for example, Evans and Mowler (2002), Petroleum IntelligenceWeekly (2005b), Piotrowksi (2009), Tuttle (2019) and recent analysis on IMO 2020 by the U.S. Energy Information Administration and the International Energy Agency.
 - 2. For example, Khrennikova and Mazneva (2018) and Dipaola (2019).
- a Corresponding author. Federal Reserve Bank of Dallas. Send correspondence to 2200 N Pearl Street, Dallas, TX 75201. E-mail: michael.plante@dal.frb.org.
- b Federal Reserve Bank of Dallas.

a government's choice of the benchmark used to set official selling prices.³ Finally, for academics, analysts and others interested in understanding the upstream and downstream oil markets, these differentials provide important signals about how supply and demand conditions change over time for one type of crude relative to others.

In a certain sense, these price differentials reflect limits on the global refining sector's ability to treat various crude streams as substitutes for one another when it comes to transforming them into the valuable petroleum products that consumers desire. In this paper, our question of interest is whether the average values of price differentials between different quality crude oils have declined over time. That is, can we find evidence that crude oils of different types may have become more substitutable for one another and that the oil market has become closer to "one great pool"?

To answer this question, we construct price differentials between numerous crude oils of various types and then test whether these differentials have experienced shifts in their means using a structural breakpoint test. While it is well known in the industry and literature that changing market conditions can cause short-run variations in these price differentials, little has been said about whether they have been affected by structural breaks that have more permanently changed their average levels.

To provide some motivation for our interest in structural breaks, we plot in Figure 1 examples of price differentials between higher and lower grade crudes for four parts of the world: Midland, TX; the U.S. Gulf Coast; Europe; and Asia. These are log-differentials using monthly data from 1997 to 2018 that consider the price of a high-quality crude relative to a lower-quality one. Visually, there is strong evidence of at least one break in the means of these differentials, occurring sometime around 2007 or 2008. The vertical lines denote the breakpoints as determined by a more formal statistical test. Visual inspection of other differentials, not shown here, strongly suggests the existence of structural breaks in many of their means, as well.

Our first contribution is to document more rigorously and systematically the extent to which differentials between crude oils of different types have experienced structural breaks in their means. Using the sequential breakpoint test of Bai (1997), we find that almost all of the differentials we look at have experienced at least one break in their mean. In particular, we find that most price differentials between crude oils of different types—25 out of 27 cases, to be exact—experienced a significant break around 2008. We then investigate how the means have shifted over time and find that most of those differentials have narrowed. A reduction in volatility has accompanied those changes. After the break, the means and volatilities are often half of their pre-break levels. Related to these findings, we show that differentials between higher-valued petroleum products, such as gasoline and diesel, and residual fuel oil, a low-quality fuel produced in greater abundance in low-grade crude oil, have also narrowed significantly and become less volatile following breaks that also occurred around 2008.

We then investigate what has changed in the oil market that would be consistent with these differentials having become smaller. We document that the global refinery sector has added a significant amount of upgrading capacity, thereby increasing its ability to process low-grade crude oils into high-value petroleum products. At the same time, we show that the U.S. shale oil boom has boosted the relative supply of high-quality, light crude oil, which does not require highly sophisticated refiners to fully process. Both changes are consistent with the findings of smaller price differentials. We also show that changes in environmental regulations on sulfur content in petroleum products and

^{3.} See, for example, Kemp (2009) on the 2009 Saudi Aramco decision to switch its benchmark from West Texas Intermediate Crude to the Argus Sour Crude Index.

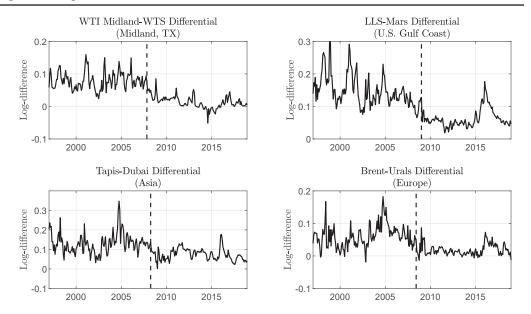


Figure 1: Oil price differentials in four areas

Notes: Figure plots log (percent) differentials between a higher and lower grade crude oil from 1997 to 2018 in four areas of the world, calculated as $\ln P_{high} - \ln P_{low}$. West Texas Intermediate (WTI), Louisiana Light Sweet (LLS), Tapis and Brent are light, sweet crudes. West Texas Sour (WTS) is light, sour, while Dubai, Mars and Urals are medium, sour crudes.

changes in the relative demand for residual fuel oil are not consistent with smaller price differentials between high and lower-quality crude oil.

Regarding the timing of the cluster of breaks, we consider the potential roles of the Great Recession, supply changes, and speculation. Data show the Great Recession significantly reduced demand for petroleum products, especially lighter petroleum products such as gasoline and diesel. At the same time, due to long lead times associated with refinery investment decisions, significant upgrading capacity additions continued unabated during the downturn. We use simple supply and demand curves to highlight how these two occurrences would help generate smaller differentials. We find no notable changes in the supply of different grades of crude oil around that time. The limited available data also does not point to any particular role for speculation with regard to the changes in the price differentials.

Finally, we also investigated whether oil price differentials between crudes of the same type, for example, two light, sweet crude oils, experienced a similar set of breaks, particularly around 2008. If that were true, it would suggest a broader change in the oil market not necessarily connected to crude quality. Overall, we do not find any evidence for this. We do, however, find that differentials between similar-type crude oils have experienced their own set of breaks. Many appear connected to changing market conditions in the United States, occurring either in the mid-2000s or after 2010, and affecting numerous differentials related to crude oils in the U.S. Gulf Coast, particularly light, sweet crude oil. A modest contribution on our part is to show that these breaks are more prevalent than previously documented in the literature.

We note that our work is connected with previous research papers, such as Weiner (1991), Sauer (1994), Ripple and Wilamoski (1995), Gulen (1997), Gulen (1999), and Bachmeier and Griffin (2006) that have considered to what extent Adelman's statement holds true. Those works have mainly looked at the degree to which oil prices move together across space and time, often using

cointegration models. Or, to elaborate on Adelman's metaphor, these works ask if there is a disturbance in one part of the pool that generates waves (price movements), do the waves spread out and affect other parts of the pool?

In our work, we consider the idea that long-run price differences also tell us something about how close the oil market is to being "one great pool" but from the quality perspective. Or, elaborating on the metaphor again, to what extent can the global refining sector literally dip into any part of the pool to get the crude oil it needs? Because of the different perspective, we focus on structural breaks in the long-run average size of the price differentials between high and low-quality crude, rather than modeling dynamics using cointegration models.⁴

Prior work in the literature has also discussed the occurrence and importance of structural breaks affecting price differentials related to key benchmarks for light, sweet crude, such as West Texas Intermediate (WTI) and Brent. See, for example, Buyuksahin et al. (2013), Borenstein and Kellogg (2014), Scheitrum et al. (2018), and Agerton and Upton (2019). These primarily focused on the implications of the shale oil boom and the ensuing logistical bottlenecks. Buyuksahin et al. (2013) also discussed Canadian production growth and issues related to storage. Our work differs from the prior literature in two main regards: (1) our main focus is on price differentials between crude oils of differing qualities; (2) we offer additional insight into the role of the downstream (refining) sector as an important market factor.

The rest of the paper is organized as follows. Section 2 offers a brief introduction to crude quality and oil price differentials. In Section 3, we discuss our data and econometric methodology. Section 4 presents evidence regarding the presence of structural breaks and documents how they have affected the differentials. Section 5 discusses what changes in the oil market are consistent with our econometric findings. We then conclude.

2. CRUDE OIL PROPERTIES AND PRICE DIFFERENTIALS

While the previous literature has found that oil prices tend to move together over time, i.e. they are cointegrated, crudes usually do not sell for the same price because of differences in their physical characteristics. Two properties of particular importance are a crude oil's American Petroleum Institute gravity, hereafter API gravity, and sulfur content.⁵ The industry has found it convenient to lump crude oils into several major groups based on these properties. It is common to label oils as light, medium or heavy depending upon their API gravity and sweet or sour depending upon whether they have low or high sulfur content.

There is a hierarchy of quality in terms of density, with light at the top and heavy at the bottom, and in terms of sulfur content, with sweet crudes preferred to sour. In terms of prices, light, sweet crudes usually command a premium relative to other grades, while heavy, sour crude oils usually sell at a discount to other grades. In this section, we discuss why these physical characteristics generate such price differentials and how refineries can attempt to arbitrage those differentials. We also briefly discuss how transportation costs can influence oil price differentials, as they play a role in some of our later econometric findings.

^{4.} Another line of work has shown the usefulness of threshold regression models when modeling the dynamic behavior of crude oil price differentials, for example Hammoudeh et al. (2008), Fattouh (2010) and Ghoshray and Trifonova (2014). Our work focuses on structural breaks in the means of price differentials, as opposed to modeling the dynamics of those differentials.

^{5.} API gravity for most crudes is a number between 10 and 70. The lower the value, the denser the oil.

2.1 The refining process and API gravity

The first step of refining crude oil involves using an atmospheric distillation unit, also referred to as a crude distillation unit (CDU), to distill the crude into various "cuts" or fractions. All refiners, from the simplest to most complex, undertake this step. In general terms, it is helpful to imagine that every barrel of crude oil can be distilled into three fractions: light distillates (naphtha/gasoline), middle distillates (diesel/gas oil) and a residual, often referred to as atmospheric residue, which is literally the bottom of the barrel. These categories are determined by their boiling points and density, with light distillates possessing the lowest densities and boiling points, and the atmospheric residue possessing the highest density and boiling point.

The API gravity of a crude is related to the proportions of the different cuts within a specific type of crude oil. Light crudes, i.e. those with a high API gravity, tend to have greater proportions of gasoline and diesel than residual products, while medium and heavy crude oils usually contain greater amounts of residual products. The exact proportions for a specific crude oil are sometimes publicly available in the form of a chemical analysis known as a crude oil assay; we now use some of these analyses to discuss the relationship between API gravity and the residual content. The inherent yields of atmospheric residue for West Texas Intermediate (WTI) and Brent, two benchmarks for light, sweet crude, are 33.3 and 34.2 percent, respectively. Mars, a benchmark medium, sour crude in the U.S. Gulf Coast, contains about 47 percent residual while Maya, a heavy, sour crude produced by Mexico, has 61.2 percent residual.⁶ The circles in Figure 2 show the relationship between API gravity and the amount of atmospheric residue present for 54 crude oils. The amount of light and middle distillates for each crude is simply 100 minus the amount of atmospheric residue.

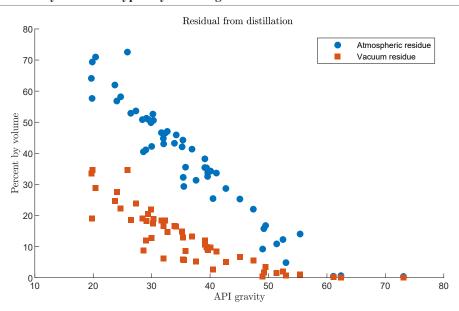


Figure 2: Heavy crude oils typically contain greater volumes of residual

Notes: Figure plots the amount by volume of atmospheric residue and vacuum residue as a function of API gravity for 54 crude oils. The data comes Exxon's library of crude oil assays.

6. These are based on assays from the Oil&Gas Journal (08/15/1994), Exxon, BP and the Oil&Gas Journal (05/15/2000), respectively. Atmospheric residual here has a boiling point over 650 degrees Farenheit and includes both vacuum gas oil and residual fuel oil. For Mars, the boiling point listed is 696 degrees.

A major difference between simple and more complex refineries is the latter's ability to transform the bottom of the barrel into other petroleum products. A simple refinery essentially has no ability to do so. More complex refineries have additional equipment to convert the residual into higher valued petroleum products. Collectively, this capital is referred to as secondary processing units, upgrading capacity or conversion capacity.

Moderately complex refineries have a vacuum distillation unit (VDU), which further distills the atmospheric residue from the CDU into vacuum gas oil (VGO) and vacuum residue, which is essentially residual fuel oil. They also have crackers, equipment that processes the VGO into lighter products. The residual fuel oil is produced in greater concentrations in lower-quality crudes. It makes up 9 percent of WTI, 9.7 percent of Brent, 25 percent of Mars and 36.9 percent of Maya. The squares in Figure 2 show the relationship between API gravity and vacuum residue for 54 crude oils.

The most complex refineries, in addition to a VDU and crackers, also have a coker. This expensive piece of capital equipment allows the refiner to break down residual fuel oil and transform it into lighter products and petroleum coke. A coker significantly reduces the amount of residual fuel oil produced during refining. As the residual content is highest in heavy crude oil, cokers are most often used by refiners that specialize in processing heavy crude. However, it can also be used to upgrade medium crude oils, such as Mars.

Residual fuel oil generally sells at a much lower price than gasoline or diesel. This inherently makes medium and heavy crude less valuable than light crude. Complex refineries take advantage of this price differential by processing lower grade crude oils, using crackers and cokers to increase the production of higher-valued products, while at the same time reducing the production of residual fuel oil.

2.2 The refining process and sulfur content

Sulfur is a pollutant and also prevents the use of sophisticated emissions control technologies in vehicles. As a result, many countries' environmental regulations require gasoline and diesel to meet strict specifications for sulfur content. Regulations also exist for residual fuel oil use in the shipping industry. Due to these policies, sulfur is usually removed during the refining process. This requires investment in desulfurisation units, also known as hydrotreaters. The compliance costs associated with these regulations create a premium for sweet crude oil, as it generally requires less processing than sour crude oil.

2.3 Transportation costs

While the main focus of this paper is on issues related to crude quality, transportation costs also influence oil price differentials. In some areas, the demand for crude oil from refineries will exceed the local supply of crude. In that case, crude oil is imported from elsewhere and there are associated costs. More specifically, imagine two areas, a and b, where a needs to import crude from b. For this example, we imagine the crude oil is of the exact same type. Then the steady state price of crude oil at location a will equal $p_a = c + p_b$, where p_b is the price of crude at b and c is the per-unit transportation cost. Given a permanent and large enough supply increase in a, however, it is possible that demand from local refineries could become satiated and that a could become an exporter of crude oil to b. In that case, we would have $p_a = p_b - c$. This would show up as a structural break

^{7.} See Samuelson (1952) for a useful and more detailed discussion on these issues.

because the mean of the differential, p_a-p_b , would shift from c to -c. The previous literature has documented breaks of these kinds in light, sweet crude differentials (see, for example, Buyuksahin et al., 2013; and Agerton and Upton, 2019).

3. DATA AND METHODOLOGY

3.1 Prices

We work with a set of 14 crude oil prices. Table 1 lists the crude oils along with their API gravity and sulfur content. The crudes are divided by location, which refers to the geographic area where the crude is priced. For the U.S. crudes, these groupings are straight forward. Waterborne crude oils outside the U.S. are broken into two groups: a Europe/Atlantic Basin group and a Middle East/Asia group. We assign Dubai and Oman into the same group as Tapis because Dubai has long been an important benchmark for a large amount of oil sold into the Asian market (Energy Intelligence, 2009).

The table also categorizes crude oils into light, medium or heavy and sweet or sour. There are no formal definitions for these categories but we define a light crude oil as any oil with an API above 33, while heavy crudes have an API below 25. A sweet crude is defined as any with a sulfur content below 0.50 percent. We note here that these categories are intended to help make the analysis more manageable by grouping crude oils of roughly the same characteristics. In reality, there is a continuum of quality. With that being said, our series includes light sweet crudes, such as Brent and Louisiana Light Sweet (LLS); medium, sour crudes, such as Dubai and Mars; and one heavy, sour crude, Maya. We have tried to include a broad set of crude oils that, while not necessarily on par with a benchmark crude, are relatively well known to ensure that the price data are of reasonable quality.

Table 1: Oil Price Series

Name	API gravity	Sulfur	API category	Sulfur category
Cushing, OK			,	
WTI Cushing (WTIC)	39.0	0.34	Light	Sweet
Midland, TX				
WTI Midland (WTIM)	39.0	0.34	Light	Sweet
West Texas Sour (WTS)	34.0	1.90	Light	Sour
U.S. Gulf Coast (USGC)				
Heavy Louisiana Sweet (HLS)	33.7	0.39	Light	Sweet
Louisiana Light Sweet (LLS)	35.7	0.44	Light	Sweet
Mars	28.9	2.05	Medium	Sour
Maya	21.1	3.38	Heavy	Sour
Europe/Atlantic Basin			·	
Brent	38.1	0.41	Light	Sweet
Saudi Heavy to Europe (SHE)	27.0	2.80	Medium	Sour
Urals	31.5	1.44	Medium	Sour
Middle East/Asia				
Dubai	31.0	1.70	Medium	Sour
Oman	33.0	1.10	Medium	Sour
Saudi Heavy to Asia (SHA)	27.0	2.80	Medium	Sour
Tapis	44.6	0.03	Light	Sweet

All price series come from Bloomberg except for Urals, which comes from the Haver database. We consider a common sample that runs from January 1997 to December 2018. We start in 1997 as that is the first year we have data available for Mars. Our data are daily except for Urals,

which comes as monthly averages. The data appendix provides the exact series name for each crude stream. Data on API gravity and sulfur content come from Bloomberg for all of the crude streams except Brent and Urals, which comes from Platts (2018).

One point worth mentioning is the lack of a price series for Canadian heavy crude oil. Given our topic of interest, it would seem natural to include such a price. We do not, however, because Bloomberg data for the current benchmark price, Western Canadian Select (WCS), only starts in 2008. Given this, we have decided to exclude WCS prices from the analysis.

Finally, we also present some additional results using monthly averages, which allows us to include two additional light, sweet crude prices: Algerian Saharan and Bonny Light. Details on the data and results can be found in the appendix.

3.2 Differentials

We consider log-differentials of the price series, as in Gulen (1997), Gulen (1999), Hammoudeh et al. (2008) and Fattouh (2010). If we denote the level of two arbitrary oil prices as P_i and P_i , the log-differential between them at time t is given by

$$p_{ij,t} = \ln P_{i,t} - \ln P_{j,t}. \tag{1}$$

The use of log-differentials has the advantage of converting units to percent differences. An additional benefit is that the log-differential is equivalent to the log of a relative price. As such we do not need to worry about the effects of inflation on the differential over time.

We generally construct the differentials so that $P_{i,t}$ is the higher-quality crude. For the daily data, we construct pair-wise differentials on all days where there is an observation for both prices, and exclude any day where we are missing one or both prices. The number of observations, therefore, varies slightly from differential to differential but, in general, we have roughly 5500 data points per differential. For the analysis using monthly data, some of the price series are only available as monthly averages. To ensure comparability across series, we take a monthly average of the daily price data when they are available. Differentials are then calculated based on the monthly averages.

Even with the limited number of price series we work with, there are a large number of differentials that can be constructed. We have found it convenient to break the differentials into two groups. The first grouping contains differentials between various crude streams within the same area, as defined in Table 1. We hereafter refer to this group as the within-area differentials. The second group consists of differentials between crude oils that are priced in different areas. We hereafter refer to these as the across-area differentials. In the results section, we also consider a further breakdown where one sub-group includes differentials where the two crudes are the same type, and another sub-group where the two crudes are of a different type.

In addition to being convenient, the breakdown into within-area and across-area differentials also has some intuitive appeal given our topic of interest. Over long periods of time, we expect transportation costs to play a smaller role for the within-area differentials, as they are priced in closer proximity to each other. This should provide a cleaner signal in regards to arbitrage across quality.

^{8.} We have access to daily data for Urals but only from 2002 to 2013. Our main conclusions for Urals-related differentials are unchanged whether we use daily or monthly data. Details can be found in the online appendix.

^{9.} For example, tanker rate data from Bloomberg shows that over our sample period the cost of shipping crude oil to the U.S. Gulf Coast from Mexico averaged about 75 cents a barrel. From the Middle East, the cost averaged \$1.93 per barrel, despite a larger tanker size. Likewise, tanker rates from the Baltic Sea to the United Kingdom averaged \$1.03 per barrel while shipping from the Middle East averaged \$1.90 a barrel.

3.2.1 Within-area differentials

The within-area differentials are constructed starting with the crude oil that has the highest API gravity in the area and then working down. For example, in the USGC we construct differentials between LLS and the three other crudes. After LLS we calculate log-differentials between HLS and the two heavier crudes, Mars and Maya, and finally the differential between Mars and Maya. There are 16 differentials in this group.

3.2.2 Across-area differentials

We follow the same procedure as before and construct across-area differentials beginning with the highest quality crude, with the following exception: the differentials between light crudes in the USGC and light crudes outside the U.S. At the start of the sample the Gulf Coast was a net importer of light crude and LLS and HLS sold at a premium to many light crudes outside of the USGC. We put LLS and HLS in the numerator of those differentials. We have also excluded all but two of the across-area differentials involving WTI Midland, WTI Cushing and WTS. These differentials show extreme changes in behavior after 2010 due to the shale boom and pipeline bottlenecks and, as many of these issues have been discussed elsewhere, for brevity's sake we do not include them in our analysis. This procedure leads to a total of 27 across-area differentials.

3.2.3 Summary statistics

To conserve space, tables containing the full set of summary statistics are in the appendix. To summarize, we find that the differentials are typically larger for those pairs of crude streams that are further apart in terms of API gravity and sulfur content, in line with the intuition developed in section 2 and previous works. For example, the mean of the LLS-HLS differential was only 1.5 percent while it was almost 23 percent for the LLS-Maya differential. A few differentials, primarily light, sweet crude differentials, do not follow this pattern as transportation costs play a large role in their average values. The LLS and HLS differentials with Brent are both positive while the two WTI-LLS differentials have negative means, for example. We also find that the greater the differences in API gravity and sulfur content, the more volatile the differential tends to be.

3.3 Methodology

There are numerous econometric methods available to test for structural breaks in a time series. We use the sequential breakpoint test of Bai (1997), which allows one to determine both the number of breaks present and their timing. Here, we provide a brief sketch of the procedure. For details on the theory we refer the reader to Bai (1997). Critical values come from Bai and Perron (2003), which also provides a discussion on more practical matters related to various structural break tests. We note here that we make use of the repartition technique introduced in Bai (1997), which makes the asymptotic distributions of the sequential test equivalent to those of the simultaneous breakpoint tests of Bai and Perron (1998).

For each differential, we consider a model of pure structural change where we estimate regression equations of the following form,

- 10. See, for example, Bacon and Tordo (2005) and Giulietti et al. (2015).
- 11. Perron (2006) provides a more general overview of structural breaks.

$$p_{ij,t} = c_{ij} + u_{ij,t}, (2)$$

and test for breaks in the intercept term, c_{ij} . This specification has the advantage of allowing for fairly general properties of the residual, including serial correlation.¹²

Time is denoted by t and the sample runs from 1 to T. There are m possible breaks and M=m+1 regimes. The test requires us to choose a maximum number of breaks to be considered. Visual inspection of the data usually pointed to no more than three breaks but we allow for a maximum of five, i.e. $0 \le m \le 5$. The breakpoint test also requires us to choose a trimming parameter, ε , which controls the minimum number of observations allowed for each regime. More specifically, if h is the minimum observations allowed, $h=\varepsilon T$. The trimming parameter can be set as low as 0.05 but we set ε to 0.15. As discussed in Bai and Perron (2003), the higher value helps mitigate against potential size distortions that can occur when the data are serially correlated. For our time series, the minimum regime size is a little over 3 years.

The first step of the procedure is to estimate the regression equation for a price differential using the full sample of data. The test searches for breaks over all allowable sub-samples and the null of no breaks versus one break is then considered for the candidate break that maximizes the test statistic. ¹³ We use the robust version of the test statistic found in Bai and Perron (1998) where the estimate of the variance-covariance matrix is robust to heteroscedasticity and autocorrelation. The matrix is estimated using the Quadratic Spectral kernel and the automatic bandwidth method of Andrews (1991). ¹⁴ If the null can be rejected at the 1 percent level, we accept the candidate break. The sample is then split in two at the estimated breakpoint and the procedure is repeated individually for the two sub-samples. This process continues until the null hypothesis cannot be rejected for any of the subsamples or until we find 5 breaks. When a candidate break is accepted, the initial estimates for breakpoints and break fractions are denoted as k_s^0 and $\tau_s^0 = k_s^0 / T$ for s = 1,..., m.

We make use of a refinement of the sequential procedure, called repartitioning, introduced in Bai (1997). This process re-estimates the dates for the breakpoints, modifying the sub-samples to take into account the initial breakpoints identified by the sequential procedure. In the case of two breaks, the repartition process re-estimates the breaks using the subsamples $\begin{bmatrix} 1, k_2^0 \end{bmatrix}$ and $\begin{bmatrix} k_1^0, T \end{bmatrix}$. The final estimates for the break fractions and breakpoints, after the repartition process, are denoted as τ_s^* and k_s^* , respectively. Under the repartition technique, the asymptotic distributions for the sequential test are the same as those of the simultaneous breakpoint tests discussed in Bai and Perron (1998).

There is a well known issue with these types of structural breakpoint tests where the test can fail to reject the null of no breaks versus 1 but finds evidence for rejecting the null of 1 versus 2 breaks. This occurs particularly when the series experiences a second break where the mean shifts back to a level close to its initial value. Visual inspections show that several of our series experience potential breaks of these types. As a result, in the few cases where the null of no breaks is not rejected we also consider the UDmax test described in Bai and Perron (1998). This test reports the maximum test statistic up to *m* breaks, in this case 5. If the UDmax test provides statistical evidence

^{12.} We also considered regression equations that explicitly modeled auto-correlation by including lags of the dependent variable. In that case the test statistics are only valid when there is no serial correlation in the residuals. In many cases, particularly with monthly data, we found it difficult to a priori properly determine the lag length, which is not surprising given the nature of the breaks we are investigating. Given our interest in testing for breaks in the mean and our concern about potential misspecification, we decided to work with the more parsimonious model in equation (2).

^{13.} Technically speaking, the test and its asymptotic properties are defined in terms of the break fraction rather than the breakpoints. We follow Bai (1997) and base our discussion around the breakpoints.

^{14.} In preliminary analysis we also considered the Bartlett kernel. We found that this generally led to smaller standard errors for the estimates and, as a result, somewhat more breaks being accepted by the test.

for more than 1 break, we then report the results for all of the cases up to and including the last break that is statistically significant at a 1 percent level. This occurs for only two cases.

Since we use daily data, we note the test will assign a specific day to any structural break detected. Our goal with using this test, however, is not to literally ascribe a particular day to a break. Instead, we use the test to provide a more rigorous, statistically-backed answer to whether a long-lasting shift has occurred in the mean of a differential and, if so, at what approximate time. By construction, the test will pick the observation where the evidence is strongest. As a robustness check, we confirm that our results are not sensitive to using the monthly data.

Finally, the breakpoint test relies on an assumption that the oil price differentials are stationary. However, as discussed in Perron (1989) and many other papers, stationarity tests can be biased by the structural breaks we are interested in. To account for this, we tested the differentials using a breakpoint unit root test. The test is an Augmented Dickey-Fuller test where the lag length is chosen using the SIC. All of the differentials are found to be stationary after allowing for a break in their mean. Additional details are in the appendix.

4. RESULTS

4.1 Identifying the structural breaks

Our first goal is to document the presence of structural breaks in the crude oil differentials. To begin, we focus specifically on pairs of crude oils of different types, first for the within-area differentials and then for the across-area differentials. Results for same-type crudes, such as the light, sweet differentials, are introduced thereafter.

We begin with the WTIM-WTS differential, which is a differential between a light, sweet crude and a light, sour crude. The test identifies two breaks that are significant at a 1 percent level. The dates and test statistics are listed at the top of the upper panel in Table 2. The first break is in December 2007 and the second in February 2013. These dates refer to the month that contains the last day of a given regime. The F-statistic for the first break is 156.51 vs. a critical value of 12.29. The second break has a test statistic of 14.14 vs. a critical value of 13.89. We list the breaks in the order the test finds them, which is related to the size of the test statistic that each break generates for the null of 0 vs. 1 break.

The middle portion of the upper panel in Table 2 shows the identified breakpoints for the USGC. Our main finding is that there is strong evidence for a break in the mean of all the series sometime between mid-2007 to mid-2008. This is similar to the timing of the first break in the WTIM-WTS differential. We also find evidence for the existence of a second break at the end of 2001 in HLS-Mars differential. A similar break is detected for the LLS-Mars differential but is not listed, as it is only significant at a 5 percent level.

Finally, we run the breakpoint tests using the differentials in the Europe and Asia groups. As with the USGC differentials, we find evidence of a break affecting all of the differentials in or around 2008.

We next consider the across-area differentials for different crude types, with the results presented in the bottom panel of Table 2. The test finds that all of the differentials, with just two exceptions, experienced a break around 2008. The test identifies a few other breaks for differentials involving light, sweet crude in the USGC and two breaks involving Mars differentials after 2010.

As shown in Table 2, a very large number of breaks occurred between 2007 and early 2009. An immediate question of interest was whether this break affected oil price differentials generally

Table 2: Breakpoint test results for crudes of different qualities

		Part 1: Withi	n-area different	ials		
-					F-statistic	
Differential	Break 1	Break 2	Break 3	0 vs. 1	1 vs. 2	2 vs. 3
Midland, TX						
WTIM-WTS	12/2007	02/2013	_	157.83	14.36	_
U.S. Gulf Coast						
LLS-Mars	02/2008	_	_	62.98	_	
LLS-Maya	05/2007	_	_	50.14	_	
HLS-Mars	05/2008	12/2001	_	58.00	14.39	
HLS-Maya	05/2007	_	_	50.44	_	
Mars and Maya	04/2007	_	_	47.28	_	_
Europe/Atlantic Basin						
Brent-Urals ^(m)	06/2008	_	_	31.96	_	
Brent-SHE	02/2007	_	_	29.69	_	_
Middle East/Asia						
Tapis-Oman	05/2008	_	_	29.78	_	
Tapis-Dubai	05/2008	_	_	39.15	_	_
Tapis-SHA	03/2009	_	_	25.27	_	

Part 2: Across-area differentials

	,				F-statistic	
Differential	Break 1	Break 2	Break 3	0 vs. 1	1 vs. 2	2 vs. 3
Light-medium						
Tapis-Urals ^(m)	05/2008	_	_	30.10	_	_
Tapis-Mars	02/2008	05/2011	_	32.51	20.00	_
Brent-Oman	05/2008	_	_	18.63	_	_
Brent-Dubai	05/2008	_	_	25.74	_	_
Brent-Mars	02/2008	08/2013	_	15.15	52.19	_
LLS-Oman	12/2008	_	_	100.62	_	_
LLS-Urals ^(m)	05/2009	_	_	51.09	_	_
LLS-Dubai	12/2008	05/2005	_	116.83	14.39	_
HLS-Oman	11/2008	_	_	89.49	_	_
HLS-Urals ^(m)	03/2007	04/2012	_	57.55	16.50	_
HLS-Dubai	11/2008	03/2005	_	105.34	17.24	_
Light-heavy						
Tapis-Maya	06/2007	_	_	47.47	_	_
Brent-Maya	07/2007	_	_	33.67	_	_
Medium-heavy						
Oman-Maya	05/2007	_	_	35.64	_	_
Dubai-Maya	03/2002	_	_	18.25	_	_
Urals-Maya ^(m)	02/2002	_	_	14.53	_	_

Notes: Dates are month/year of the last day of the regime. The critical values are 12.29, 13.89, and 14.80 for tests of 0 or 1 break, 1 or 2 breaks, and 2 or 3 breaks, respectively. These reflect a significance level of 1 percent. A ^(m) refers to results based on monthly data.

speaking or if it was limited to differentials between different types of oil. To investigate this, we next tested for breaks in the differentials between crude oils of the same type, i.e. the light, sweet differentials and the medium, sour differentials. The results from those tests are shown in Table 3. The upper panel is for the within-area differentials while the bottom panel is for the across-area differentials.

Our main finding is that while the test identifies a number of breaks, evidence for a large set between 2007 and 2009 is non-existent. We find two breaks impacting the LLS-HLS differential after the start of the shale boom. We also find a set of breaks in the mid-2000s and during the shale

boom that affect across-area, light, sweet differentials, and several breaks for medium, sour differentials that involve Mars crude.

Table 3: Breakpoint test results for crudes of similar type

05/2011

05/2004

01/2005

01/2002

07/2013

08/2013

01/2005

08/2013

08/2013

03/2002

		Part 1: With	in-area differen	tials		
					F-statistic	
Differential	Break 1	Break 2	Break 3	0 vs. 1	1 vs. 2	2 vs. 3
U.S. Gulf Coast LLS-HLS	02/2011	07/2014	_	28.86	27.82	_
Europe/Atlantic Basin Urals-SHE(m)	01/2001	_	_	13.55	_	_
Middle East/Asia						
Oman-Dubai	_	_	_	_	_	_
Oman-SHA	03/2009	_	_	13.61	_	_
Dubai-SHA	_	_	_	_	_	_
		Part 2: Acro	ss-area differen	tials		
					F-statistic	
Differential	Break 1	Break 2	Break 3	0 vs. 1	1 vs. 2	2 vs. 3
Light, sweet						
WTIC-LLS#	04/2010	02/2006	08/2013	11.65	84.81	12.77
WTIM-LLS	01/2011	11/2006	_	16.49	143.70	
LLS-Tapis	01/2005	05/2011	03/2015	75.72	20.83	29.28

120.02

60.35

90.40

9.39

17.50

22.52

38.89

59.25

36.70

14.42

19.14

Notes: Dates are month/year of the last day of the regime. The critical values are 12.29, 13.89, and 14.80 for tests of 0 or 1 break, 1 or 2 breaks, and 2 or 3 breaks, respectively. These reflect a significance level of 1 percent. A ^(m) refers to results based on monthly data.

11/2005

4.2 Grouping the structural breaks

LLS-Brent

HLS-Tapis

HLS-Brent

Medium, sour Oman-Mars#

Urals-Mars(m)

Oman-Urals^(m) Urals-Dubai^(m)

Dubai-Mars

Previous works in the literature have focused on breaks in light, sweet crude differentials, whereas our work focuses on breaks in differentials between light, medium and heavy crude oils. To help provide some context to our findings, we have found it useful to group the breaks. Our analysis leads us to consider four groupings, shown in Table 4. Any year with a break is marked with an X and particular groups have been shaded together.

The breaks involving differentials between different types of crude oil are the ones occurring between 2007 and 2009. The rest of the paper is devoted primarily to discussing results related to these breaks.

Another set is shaded dark gray and occurs from 2010 onwards. These breaks primarily, but not exclusively, affect light, sweet differentials that involve at least one U.S. crude oil. Given the timing and the differentials involved, it is natural to conclude that these breaks are a result of the shale oil production boom in the U.S. and production growth in Canada. Most of these breaks, or ones of similar nature, have been documented previously in the literature. A few, such as the WTIM-WTS break in 2013, do not appear to have been noted before.

Table 4: Grouping the breaks

_	ic 4. Groupi	_																					
	WTIM-WTS											X						Χ					
	LLS-Mars												X										
	LLS-Maya											X											
	HLS-Mars					X							X										
	HLS-Maya											X											
	Mars-Maya											X											
	Brent-Urals												X										
o l	Brent-SHE											X											
Differentials Between Different Grades of Crude	Tapis-Oman												X										
of (Tapis-Dubai												X										
ades	Tapis-SHA													X									
t Gr	Tapis-Urals												X										
èren	Tapis-Mars												X			Χ							
Diff	Brent-Oman												X										
'een	Brent-Dubai												X										
3etw	Brent-Mars												X					Χ					
als I	LLS-Oman												X										
renti	LLS-Urals													X									
iffe	LLS-Dubai									X			X										
	HLS-Oman												X										
	HLS-Urals											X					Χ						
	HLS-Dubai									X			X										
	Tapis-Maya											X											
	Brent-Maya											X											
	Oman-Maya											X											
	Dubai-Maya						X																
	Urals-Maya					X																	
П	LLS-HLS															Χ			X				
	WTIC-LLS										X				X			X					
rude	WTIM-LLS										X					X							
*Light Crude	LLS-Tapis									X						Χ				Χ			
'Ligl	LLS-Brent									X						Χ							
*	HLS-Tapis								X														
	HLS-Brent									X								X					
		97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
												_						_					-

Two breaks affecting WTI-LLS differentials occurred in 2006. These breaks are connected with two pipeline reversals that marked the beginning of increased flows of Canadian crude oil into Cushing and the U.S. Gulf Coast.¹⁵ Another set of breaks affects differentials between light crude in the U.S. Gulf Coast and, primarily, a number of waterborne, light crude oils outside the U.S. These

^{15.} Specifically, the WTI Cushing - LLS break lines up exactly with the reversal of Enbridge's Spearhead Pipeline, which allowed Canadian crude to flow directly to Cushing. Shortly thereafter, Exxon's Corsicana pipeline was also reversed. See Cook (2006) and Scott (2006) for media reports on the reversals.

occur in 2004 and 2005. We speculate these are connected with the boom in demand both for crude oil and oil products that occurred in the Asia Pacific region, and particularly China, around that time.¹⁶

4.3 Quality-related oil price differentials are smaller

We next show in Table 5 how the means of the price differentials between high and low-er-quality crude oils have changed across regimes. The differentials are grouped in a similar manner to Tables 2 and 3. The final column shows how the means have changed from the initial to final regime. As a reminder, a change of -0.1 means a 10 percentage point decline.

Our main result is that the means of most of these price differentials have shrunk in half, at least. We find many cases where the average values have declined by close to 10 percentage points or more, particularly for differentials connected with the USGC. We find only two notable exceptions to our main finding: the Tapis-Mars and the Brent-Mars differentials. Both experienced a sharp decline in their mean in early 2008, similar to other price differentials between light and medium crude oils, but then experienced a reversal after the beginning of the shale boom.

The table also shows how the means of light, sweet differentials have evolved. We find that the two WTI-LLS differentials have over time gone from being near 0 to negative. Likewise for differentials between USGC light, sweet crude and light crude oils outside the U.S. These findings are consistent with Canadian production growth and the major increase in the supply of light, sweet crude in the U.S., and reconfirm previous results in the literature regarding structural breaks affecting light, sweet crude differentials.

4.4 Quality-related oil price differentials are less volatile

Figure 1 is suggestive of a change in the volatility of many price differentials between high and lower-quality crude, particularly those experiencing a break around 2008. We investigate this a little deeper by comparing the means and standard deviations of these price differentials before and after 2008. While the actual break varies across differentials, we decided to work with a "pre-break" period that runs until the end of 2008 as this simplifies the exposition.

The statistics are shown in Table 6. We include any differential that experienced a permanent drop in its mean since 2008. Overall, we find a marked reduction in both the average level of the differentials, as well as their volatilities. In most cases, the mean in the post-break sample is less than half the size of the pre-break mean. Post-break volatilities are about 1/2 to 3/4 the size of the pre-break volatilities.

4.5 Residual fuel oil differentials are also smaller, less volatile

In Section 2, we discussed the connection between a crude oil's API gravity and its inherent yield of residual products that come from the distillation process. This should create a relationship

16. Data from British Petroleum's Statistical Review of World Energy show a major increase in oil consumption, refining throughput and refinery utilization rates in China around that time. This coincided with a run up in tanker rates. This boom has been highlighted in the academic literature, for example in Hamilton (2009), Kilian (2009) and numerous works thereafter. News reports also corroborate an increased flow of light crude from Africa to China, and the potentially important impacts this would have on other refining centers competing for that oil (Yong, 2004). It is conceivable that growing pressure for that oil to move toward the Asia Pacific region could have boosted the relative price of light oil outside the U.S. Gulf Coast.

Table 5: Regression constant across regimes

Differential	97 98 99 00 01	02 03 04	05 0	6 07	08	09	10	11 12	2 13	14	15	16	17 18	Change
WTIM-WTS	0.	079			0.024 0.003*					-0.076				
LLS-Mars	0.152					0.065					-0.087			
HLS-Mars	0.161	0.1	08						0.05	5				-0.106
LLS-Maya	0.312	2						0.	151					-0.161
HLS-Maya	0.292	2						0.	142					-0.150
Mars-Maya	0.158	3						0.)83					-0.075
Brent-Urals ^(m)	0.	061							0.01	8				-0.043
Brent-SHE	0.198	3						0.0)87					-0.111
Tapis-Oman	0.	116							0.06	9				-0.063
Tapis-Dubai	0.	130							0.07	4				-0.056
Tapis-SHA		0.195							0.	109				-0.087
Tapis-Urals ^(m)	0.	122							0.06	2				-0.060
Tapis-Mars	0.	149			C	0.080)			0.1	10			-0.039
Brent-Oman	0.	055							0.02	4				-0.031
Brent-Dubai	0.	069			0.030					-0.039				
Brent-Mars	0.	088					0.0	34			(0.080		-0.008
LLS-Oman		0.117							0.0	030				-0.086
LLS-Urals ^(m)		0.121							0.0	020				-0.101
LLS-Dubai	0.143			0.101					0.0	035				-0.107
HLS-Oman		0.098							0.0	021				-0.077
HLS-Urals ^(m)	0.108	3			0.0	47				0.0	003*			-0.105
HLS-Dubai	0.124			0.083	0.026					-0.098				
Tapis-Maya	0.308	3						0.	186					-0.121
Brent-Maya	0.	245							0.14	0				-0.105
Oman-Maya	0.192	2						0.	113					-0.078
Dubai-Maya	0.210						0.1	19						-0.091
Urals-Maya ^(m)	0.219			,		().131	Į.						-0.088
WTIC-LLS	0.002*			-0.0	035		-	-0.138			-0.	055		-0.057
WTIM-LLS	-0.008					43				-0.	125			-0.117
LLS-HLS		0.019						_(0.002	*		0.0	13	-0.007
LLS-Tapis	0.015			-0.0	019			_	0.067	,		-0.0	27	-0.042
LLS-Brent	0.075	75 0.0				039 -0.005*				-0.080				
HLS-Tapis	-0.002				-0.045				-0.043					
HLS-Brent	0.056	•				0.01	6					-0.01	9	-0.075

Notes: Change is the difference between the final regime and the first regime for each regression equation. A * means the coefficient is not statistically different from 0 at a 5 percent confidence level. In the table, breaks that occur from July to December in a particular year are assigned to the following year. A (m) refers to results based on monthly data.

between the differentials of residual fuel oil to other, lighter petroleum products, and the differentials between lighter and heavier crude oil. Given this, we investigated whether residual fuel oil differentials have experienced breaks similar to those affecting the quality-driven oil price differentials.

We use daily data to calculate log-differentials, as in Equation 1, between the spot price of high-sulfur residual fuel oil and the following spot prices, all for delivery in the Gulf Coast: heating

Table 6: Summary statistics pre and post-break

		Part 1:	Within-are	a differential	S	
	Pre-	-break	Post	-break	Ratio of mean	Ratio of std. dev.
Differential	Mean	Std. dev.	Mean	Std. dev.	(post/pre)	(post/pre)
Midland, TX WTIM-WTS	0.076	0.031	0.010	0.018	0.13	0.58
U.S. Gulf Coast						
LLS-Mars	0.147	0.052	0.063	0.029	0.45	0.56
LLS-Maya	0.299	0.089	0.141	0.055	0.47	0.62
HLS-Mars	0.128	0.053	0.054	0.023	0.42	0.43
HLS-Maya	0.279	0.086	0.133	0.051	0.48	0.59
Mars-Maya	0.151	0.062	0.079	0.040	0.52	0.65
Europe/Atlantic Basin						
Brent-Urals(m)	0.058	0.036	0.019	0.017	0.33	0.47
Middle East/Asia						
Tapis-Oman	0.114	0.058	0.069	0.033	0.61	0.57
Tapis-Dubai	0.128	0.058	0.074	0.032	0.58	0.55

Part 2: Across-area differentials

	Pre	break	Post	-break	Ratio of mean	Ratio of std. dev.
Differential	Mean	Std. dev.	Mean	Std. dev.	(post/pre)	(post/pre)
Light-medium	1					
Brent-Oman	0.053	0.048	0.025	0.031	0.47	0.65
Brent-Dubai	0.067	0.051	0.030	0.030	0.45	0.59
LLS-Oman	0.116	0.055	0.030	0.043	0.26	0.78
LLS-Urals(m)	0.121	0.049	0.025	0.037	0.21	0.75
LLS-Dubai	0.130	0.056	0.035	0.043	0.27	0.77
HLS-Oman	0.096	0.055	0.022	0.042	0.22	0.78
HLS-Urals ^(m)	0.101	0.046	0.017	0.033	0.17	0.71
HLS-Dubai	0.110	0.057	0.027	0.042	0.25	0.74
Light-heavy						
Tapis-Maya	0.296	0.095	0.181	0.055	0.61	0.59
Brent-Maya	0.235	0.081	0.136	0.055	0.58	0.68
Medium-heavy						
Oman-Maya	0.182	0.078	0.112	0.050	0.62	0.64

Notes: The pre-break sample runs from January 1997 to December 2008. The post-break sample runs from January 2009 to December 2018. A ^(m) means the statistic is based on monthly data.

oil, gasoline, LLS and Mars. Since residual fuel oil is the low-quality product, its price is always in the denominator. A plot of the differentials, found in the appendix, shows a remarkable similarity between them and the differentials plotted in Figure 1. This similarity exists if one uses product prices for New York Harbor and replaces LLS and Mars prices with Brent and Dubai.

More formally, we run breakpoint tests on the Gulf Coast fuel oil differentials and found that all of them experienced a break in their mean around the same time that many of the price differentials between high and lower-quality crude oils did. The gasoline-residual fuel oil differential has a break in September 2007, while the other differentials have a break in January 2009. The decline in the means is on the same order as was documented for the oil price differentials.

4.6 Additional results

We also repeated our analysis using monthly price data for a slightly larger set of crude oils. We find additional evidence of breaks between 2007 and 2009 affecting the differentials be-

tween different grades of crude oil. Specifically, we find breaks for 38 out of 42 possible cases. We also find additional breaks in late 2004 and early 2005 affecting differentials between light, sweet crude oils in the U.S. Gulf Coast and light crude outside the U.S. Finally, we checked if the break dates using monthly data coincide with those using daily data. The test almost always selects the same month or an adjacent month for the breaks in the differentials between crudes of different types. Details on the monthly data and results can be found in the appendix.

As a further robustness check, we also used the cointegration breakpoint test of Gregory and Hansen (1996). This tests for a break in the intercept term of the long-run equilibrium equation, given by $\ln P_{i,t} = c + \beta \ln P_{j,t} + u_t$. Using daily data, the test also finds a large number of breaks in the quality differentials between 2007 and 2009.¹⁷

5. DISCUSSION

The empirical findings show that differentials between light crude oil and other grades of crude have remained low since 2008. We also find evidence of a similar change in the differentials between lighter petroleum products, i.e. gasoline and diesel, and residual fuel oil. In this section, we propose to look at several factors that, based on the discussion in Section 2, should be important to the determination of these price differentials. Our goal, to the extent possible, is to shed light on what changes in the oil market would be consistent with the empirical findings of smaller differentials.

We specifically propose to investigate the following factors: (1) environmental regulations on sulfur content in petroleum products; (2) the demand for residual fuel oil relative to lighter petroleum products; (3) the relative supplies of different types of crude oil; (4) the capacity of the global refining sector to process low-quality crude oil. Our approach is descriptive in the sense that we collect and present as much relevant data as possible for each factor and then use those data to inform our understanding of what has been changing in the oil market over our sample period.

Briefly summarizing our findings, we believe that changes in the relative supplies of different types of crude oil and changes in the refining sector are consistent with the lower differentials seen since 2008, while changes in environmental regulations and the relative demand for residual fuel oil are not. Specifically, the data show that the supply of light crude relative to heavy crude has increased dramatically and somewhat unexpectedly over the past ten years due to the U.S. shale boom, while the global refining sector has become more complex due to increased upgrading capacity, which makes it easier for the sector to process low-quality crude oil. On the other hand, environmental regulations on sulfur have become more stringent over time and cover a growing proportion of consumption of the affected fuels, which should lower the relative demand for low-quality crude oil. Likewise, consumption data show a clear negative trend in the use of residual fuel oil and significant growth in the consumption of other, lighter petroleum products since 1997. Holding all else equal, those trends are not consistent with a pattern of reduced differentials.

Given the timing of the breaks, we also briefly consider events around the time of the Great Recession. Consumption data show the Great Recession played a role by unexpectedly and significantly reducing the global demand for petroleum products in 2008 and 2009, particularly for lighter products. At the same time, additions to upgrading capacity continued uninterrupted due to the long lead times for refiner expansions. We document suggestive evidence that the simultaneous occurrence of a major reduction in demand and continued upgrading capacity additions led to a non-trivial increase in spare coking capacity, something that has persisted to this day. A simple supply-demand diagram is used to provide intuition into how those changes could impact price differentials. We also

look into oil supply changes and the potential role of speculation but do not uncover any evidence that suggests a major role for either of those factors.

Before moving to the details, it is worthwhile to mention important caveats to our analysis. Being descriptive in nature, we are unable to make formal causal inferences or to understand the relative contributions of any particular factor in determining the outcomes we see in the data. It would be preferable to have a formal econometric model, but a lack of data with regard to the global refining sector and crude quality rule this out. Although not as rigorous as a formal model, we still believe our exercise yields important insights. With better data, future research could certainly expand upon our discussion, particularly along the quantitative dimension.

5.1 Important factors

5.1.1 Sulfur regulations increasingly stringent and more widespread

Regulations governing sulfur content of petroleum products are an important factor because sulfur content is generally much higher in heavy crude oil than light crude. As it is costly to remove sulfur during the refining process, these regulations can affect price differentials by reducing the demand for lower-quality crude oil.

Here, we briefly discuss how the prevalence of these regulations has evolved since countries began implementing them in early the 1990s. We also provide estimates for the amount of global consumption covered by the regulations and how that has changed. This is important to consider because from 1997 to 2018 demand growth was driven by developing countries, which may have had fewer regulations in place than developed countries. The appendix provides more details on the data and calculations.

To summarize, we find that these regulations, which affect gasoline, diesel and residual fuel oil, have become more widespread and cover a growing share of consumption of the respective fuels. Those changes are not consistent with the empirical findings of smaller price differentials between light crude oil and other grades of crude.

Gasoline: By 2007, all OECD countries, except Mexico and Turkey, had transitioned to low sulfur gasoline.¹⁸ Based on consumption data from the International Energy Agency, these countries accounted for 65 percent of global gasoline consumption in 2007. By 2018, all OECD countries were using either ultra-low sulfur or low sulfur gasoline. In addition, China was also using ultra-low sulfur gasoline while India switched to low sulfur gasoline. In 2018, those nations accounted for 70 percent of global gasoline consumption. This represents a lower bound, as a handful of other countries, for which we do not have consumption data, have implemented regulations for low or ultra-low sulfur gasoline use.

Diesel: Regulations for diesel are more complicated than those for gasoline as they often distinguish between the fuel's end-use, for example on-highway vs. off-highway use. With that being said, by 2007 China and all OECD countries, except Turkey, had passed regulations mandating low sulfur diesel for on-highway use. ¹⁹ These countries accounted for a little under 65 percent of global diesel consumption in 2007. This likely overestimates the share since it would not apply to all uses of diesel.

By 2018, China and the OECD countries (except Mexico) had transitioned to ultra-low sulfur diesel, and the regulations applied to a wide range of uses. India also transitioned to low sul-

^{18.} Low and ultra-low sulfur gasoline is defined as having less than 150 ppm or 10 ppm, respectively.

^{19.} Low and ultra-low sulfur diesel have less than 500 ppm or 15 ppm, respectively.

fur diesel. In 2018, China, India and the OECD countries accounted for 66 percent of global diesel consumption. Unlike gasoline, a fairly large number of other countries have transitioned to low or ultra-low sulfur diesel. They include Brazil, Russia, more than a dozen countries in Africa and a handful of other countries in Latin America and the Asia-Pacific region. Given this, the 66 percent share is a lower bound.

Residual fuel oil: Major regulations governing sulfur content in residual fuel oil primarily affect its use in the shipping industry. The International Maritime Organization (IMO) sets these rules. Global caps on sulfur content were implemented in 2005 and tightened in 2012. Stricter regulations are in place for certain areas of the US and Europe, which required low sulfur fuel oil in 2010 and ultra-low sulfur fuel in 2015.²⁰ The most significant regulations take effect in early 2020 and require the use of very-low sulfur fuel oil or the installation of scrubbers worldwide.

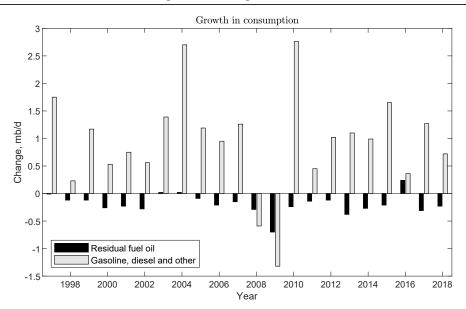


Figure 3: Residual fuel oil consumption declining over time

Notes: Units are annual change in millions of barrels per day. Gasoline, diesel and other includes naphtha, gasoline, jet fuel, middle distillates and biofuels but excludes natural gas liquids, such as ethane.

5.1.2 Demand growth tilted towards lighter products

The relative demand for different petroleum products by end users should also be important. If the world desires large amounts of gasoline or diesel relative to residual fuel oil, for example, that would have implications for the relative value of the products and, consequently, the relative price of different types of crude oil.

Holding all else equal, a pattern of increased use of residual fuel oil relative to lighter products would be conducive to generating the smaller differentials seen since 2008. Consumption data, however, shows that the opposite occurred over our period of analysis. Figure 3 plots annual changes in the consumption of residual fuel oil and all other petroleum products (excluding natural gas liquids) from 1997 to 2018. This data comes from various Annual Statistical Supplements from

20. Low, very-low and ultra-low sulfur fuel oil have less than 100 ppm, 50 ppm or 10 ppm, respectively.

the International Energy Agency. Use of residual fuel oil has declined almost every year since 1997, while use of other petroleum products increased at a relatively rapid pace. The only exception was during the Great Recession, when demand for lighter products declined to a greater extent than residual fuel oil.

5.1.3 Global supply of crude unexpectedly lighter

The relative supply of various grades of crude will affect oil price differentials and can also impact product price differentials, since it affects the relative supply of different petroleum products, as shown in the left panel of Figure 2. In Table 7, we show how the supply of light and heavy crude has changed since the year 2000, using data from Eni and the U.S. Energy Information Administration. The data from Eni is only available in five year slices from 2000 to 2015, and annually thereafter.

								Share	of total
	2000	2005	2010	2015	2016	2017	2018	2010	2018
World crude output (total)	68.2	74.1	74.1	80.6	80.5	80.5	82.3	100%	100%
Light (total)	21.6	21.2	22.1	24.5	23.9	24.9	27.3	29.8%	33.2%
U.S. shale	0.4	0.4	0.8	4.8	4.4	5.0	6.5	1.1%	7.9%
Heavy (total)	8.4	10.0	10.3	11.6	11.1	11.0	11.0	13.9%	13.4%

Table 7: Light crude production up since 2005

Notes: Units are millions of barrels per day. In the Eni data, light crude oil has an API gravity 35 and higher while heavy crude is below 26 API. U.S. shale is the monthly U.S. light tight oil production series from the U.S. Energy Information Administration (EIA). Sources: Eni World Oil Review 2019; U.S. EIA; Authors' calculations.

We believe two facts in the table are particularly relevant. First, the production of light crude has increased quite dramatically since the late 2000s, due in large part to the U.S. shale boom. Global production of light crude has increased by about 5.2 mb/d since 2010. U.S. shale production, which is light oil, increased by 5.7 mb/d over the period. Second, the production of heavy crude also increased, but the amounts are modest relative to light crude: 2018 levels exceeded 2010 levels by 0.7 mb/d and 2005 levels by 1.0 mb/d. This is perhaps surprising as there have been some important heavy crude production increases in Canada and Iraq recently. However, these have been partially offset by heavy crude production declines in Mexico and Venezuela, both important producers of heavy crude, as well as several other countries.²¹

The increase in the relative supply of light crude is important for two reasons. First, as shown in Section 2, light oil naturally produces less residual fuel oil than medium and heavy crude oils. All else equal, an increase in the relative supply of light crude should contribute to smaller differentials between lighter, higher-valued petroleum products and residual fuel oil. That in turn should work to reduce the spread between light crude and other types of crude. Second, the extent of the production boom that has occurred in light crude would not have been fully appreciated by refiners until sometime after 2010 because of the unexpected nature of the shale boom.²² This could

^{21.} The appendix provides a more detailed breakout.

^{22.} Projections from the U.S. Energy Information Administration (EIA) and the International Energy Agency (IEA) are in line with this statement. The U.S. EIA first provided a projection for U.S. shale production in its 2012 Annual Energy Outlook report. The reference case had U.S. shale production growing to about 1.1 mb/d by 2018. Actual production levels in 2018 were 6.5 mb/d. The 2013, 2014 and 2015 reports increased those projections to about 2.8 mb/d, 4.8 mb/d and 5.6 mb/d. One reaches a similar conclusion based on IEA projections. Its first projection for shale production, in the 2011 Medium-Term Oil and Gas Markets report, had production reaching only 1.4 mb/d in 2016.

be important because there are significant lead times associated with constructing a complex refinery or installing major pieces of upgrading capacity, and once operating, the capital is a long-lived sunk cost.

One interesting question is whether heavy crude production is likely to decline significantly in the near future due to the high breakeven price for new oil sands projects in Canada and the low oil price environment that has existed since late 2014. Analysis and projections in National Energy Board (2018a) and National Energy Board (2018b) show this to be unlikely because already operating bitumen mines have low operating expenses and decline rates. As a result, oil sands production is fairly insensitive to price declines.

5.1.4 Global refining sector more complex

Upgrading capacity gives refiners the flexibility to adjust outputs from a particular type of crude, rather than changing the quality of their crude input. Specifically, it provides the option of increasing the amount of gasoline and diesel that can be produced from a given barrel of medium and heavy crude oil while reducing the supply of residual fuel oil. Given this, it is important to assess how the global stock of upgrading capacity has evolved over time and to view these changes in light of the trends discussed earlier. While publicly available data on the global refining sector are limited, it clearly shows an increasing ability of the global refining sector to convert low-grade crude oils into high-value petroleum products. Importantly, this has occurred against a backdrop of growing light crude production over the past 10 years, which does not require highly sophisticated refineries to process.

To discuss trends in global refining capacity, we present data on primary and upgrading capacity from Eni's World Oil Review and from the International Energy Agency. Table 8 shows the data from Eni on global refining capacity, as well as two measures of refining sector complexity.²³ The first column shows data on primary capacity, which is crude distillation capacity and condensate splitters. The second column shows data on conversion capacity, which measures the availability of cracking and coking capacity.²⁴ The third column shows the first measure of complexity, which is simply the ratio of conversion capacity to primary capacity. The final column is the Nelson Complexity Index (NCI). This is a commonly used measure of refinery complexity where higher values reflect greater complexity, either at a particular refinery or for a particular area.²⁵ Unlike the conversion capacity data, the NCI reflects both upgrading capacity and desulfurisation capacity.

Since 2000, primary capacity has grown at about 1 percent per year on average, or 0.9 million barrels per day (mb/d). Conversion capacity has grown at a more rapid pace, about 4 percent a year on average, or 1.3 mb/d. This has led to an increase in the conversion capacity ratio and contributed to higher NCI values. The ratio of conversion capacity to primary capacity rose from 38 percent in 2000 to 54 percent in 2017 while the NCI rose from 7.9 in 2000 to 9.3 in 2017. Some of the largest increases in complexity occurred in Asia, where the conversion capacity ratio rose from 36 percent in 2000 to 66 percent in 2017, and the NCI increased from 7.0 to 9.7.

- 23. A more complete time series can be put together using older versions of Eni publications. However, the data have been revised several times, most recently in 2015. As a result, while the numbers in Table 8 are comparable to each other, the longer time series one can put together using older reports are not comparable, strictly speaking. The appendix contains the full series.
- 24. Conversion capacity is fluid catalytic cracking equivalent. Details on the calculation can be found in Eni's World Oil Review 2018.
 - 25. Johnston (1996) provides a good introduction to the index and how it is calculated.

	0. 0.0000 - 0				
Year	Primary capacity (mb/d)	Conversion capacity (mb/d)	Conversion capacity ratio (percent)	Nelson Complexity	
2000	83.2	31.6	38	7.9	
2005	87.3	37.5	43	8.2	
2010	92.4	43.4	47	8.7	
2015	96.5	50.2	52	9.1	
2016	98.1	52.0	53	9.3	
2017	98.7	53.3	54	9.3	

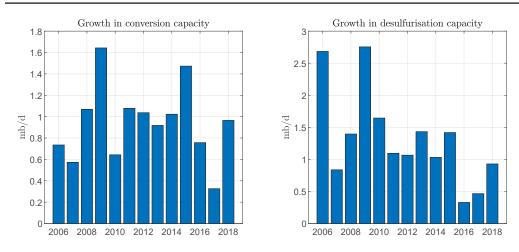
Table 8: Global refineries increasingly complex

Notes: The conversion capacity ratio is conversion capacity divided by primary capacity. Sources: Eni World Oil Review 2018, Eni World Oil Review 2017, Eni World Oil and Gas Review 2016.

International Energy Agency reports also provide data on additions to global conversion capacity and desulfurisation capacity from 2006 to 2018. We plot both of those series in Figure 4. This data reinforces the findings of the Eni data, as it shows significant additions to conversion capacity.

Two other data sources are available on refining capacity but were not used for this discussion. The British Petroleum Statistical Review of World Energy has data on primary capacity but not upgrading capacity. The differences between the BP and Eni primary capacity series are relatively modest, so none of our conclusions are sensitive to using one series or the other. The Oil&Gas Journal Worldwide Refinery Survey also provides primary and conversion capacity numbers, but we have chosen not to use this data because participation is voluntary and it appears capacity in some important developing countries, particularly China, has been underreported.²⁶

Figure 4: IEA data shows substantial additions to upgrading capacity



Notes: Units are millions of barrels per day. The data come from various International Energy Agency Medium-Term Oil Market Reports and Market Report Series.

A referee raised an interesting question: Has there been some exogenous technological change or innovation that could explain the changes in the global refining sector just documented? Or do they reflect endogenous decisions made by refiners in response to market conditions? To an-

26. For example, the survey reported Chinese distillation capacity and coking capacity of 7 mb/d and 156,000 b/d at the start of 2013, respectively, but the International Energy Agency's Medium-Term Oil Market Report 2013 shows distillation capacity of 13.4 mb/d and coking capacity of 1.8 mb/d (see table on page 98).

swer this, we surveyed the academic literature, various general interest and trade publications, and government reports on the refining sector. While we found evidence that the sector has benefited from the adoption of new technology, such as computers and improved catalysts, we could identify no major innovation that could explain our breakpoint test results or the buildup in refinery upgrading capacity.²⁷

On the other hand, we found ample evidence that refinery operators are forward-looking and aware of many of the major trends previously discussed, to varying degrees. This includes expectations for continued growth in global demand for petroleum products and awareness of what areas would drive that growth, the skew in demand growth favoring transportation fuels, and major shifts in environmental regulations. Media reports show that while the industry was perhaps surprised by the size and persistence of the demand boom in the mid-2000s, the industry became aware of the boom as early as 2004 and knew of China's important role. We also found a debate about the direction of crude quality in the mid-2000s, the evolution of price differentials between light and heavy crude oil and about whether investment in upgrading capacity was warranted, particularly with regard to U.S. refiners, which are geographically distant from major centers of demand growth. However, we found no evidence that the sector anticipated the shale oil boom during the first decade of the 2000s.²⁸

Evolution of coking capacity in the U.S. Gulf: The data in Table 8 and Figure 4 provide information on upgrading capacity in the global refining sector. More detailed information is available for the U.S. refining sector from the U.S. Energy Information Administration. This data run from 1987 to 2018 and includes capacity data for several forms of upgrading capacity as well as data on how much feedstock was processed. This allows us to generate series on spare capacity and utilization rates, which are not available on the global level.

We use this data to focus specifically on U.S. Gulf Coast coking capacity. We focus on cokers because they provide the greatest ability to process lower grade crude oils. Data for the Gulf Coast is used for three primary reasons: (1) the area is a major global refining center; (2) it is home to a majority of U.S. coking capacity; and (3) it is fully integrated into global oil and fuel markets. As a result, utilization rate data for the U.S. Gulf Coast stand the best chance of being reflective of conditions elsewhere in the world. Additionally, our data set includes a number of differentials in the area.

The left panel of Figure 5 shows the evolution of coking capacity, feedstock inputs to cokers and the amount of spare capacity, which is simply capacity minus feedstock inputs. Capacity more than doubled from 1987 to 2004. The pace of capacity additions slowed down thereafter. No-

- 27. All major refining processes were developed well before our study period. See Enos (1958) and Enos (1962) for a historical overview. Based on interviews with refineries, Peterson and Mahnovski (2003) document incremental productivity gains due to improvements along several dimensions, including catalysts. Innovation in catalysts, though, seems to have often been an endogenous response to more stringent environmental regulations governing sulfur and lead content. Kerr and Newell (2003), for example, document the adoption of new catalyst technology that occurred in the 1980s and early 1990s in response to the U.S. ban on lead in gasoline. Government reports we surveyed, including Energy Information Administration (1990), Pirog (2007), Andrews et al. (2010) and Greenley (2019), did not mention any major innovations.
- 28. Peterson and Mahnovski (2003) provides a good overview of the issues of which the refining sector was aware as of the early 2000s. Issues related to demand growth were covered in, for example, Reuters (1990), Aizhu (2000), Reuters (2005b), and Pirog (2007). The oil market literature has documented the importance of unexpected demand shocks in the mid-2000s. See Kilian (2009) and various other papers. Both Jegarajah (2004) and Yong (2004) discuss China. Media articles on upgrading capacity, crude quality and price differentials include Reuters (2005a), Subrahmaniyan (2005), Petroleum Intelligence Weekly (2005b), Petroleum Intelligence Weekly (2005a) and, more recently, Koster (2017). van Schaik (2019) shows that large coker capacity additions continue outside the U.S.

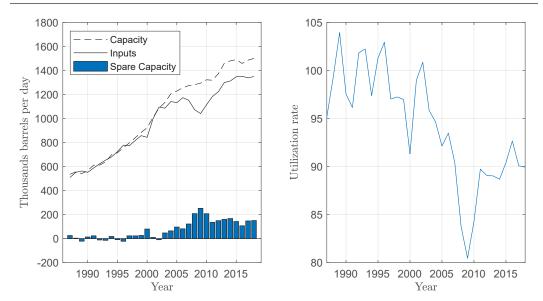


Figure 5: Coking capacity and utilization in the U.S. Gulf Coast

Capacity is the U.S. Energy Information Administration's calendar day capacity measure. Spare capacity is capacity minus fresh feed input to coking units. Utilization is fresh feed input divided by capacity. All data are for Petroleum Administration Defense District III.

tably, up through 2002 there was essentially no permanent spare capacity available in the system.²⁹ A small cushion emerged in 2003, grew modestly through 2007 and then expanded significantly during the Great Recession. By historical standards, a sizeable spare capacity cushion has remained in place since then.

The right panel of 5 shows the corresponding utilization rate for cokers. Through 2002, utilization rates frequently exceeded 100 percent; that has not occurred since 2003. During the recession, utilization rates declined sharply and have since remained at levels below those seen from 1987 to the mid-2000s.

5.2 Changes during the time of the Great Recession

Given the timing of the breaks, it is natural to ask about potentially important events between 2007 and 2009 that could impact the price differentials. In this section, we investigate the potential importance of the Great Recession, changes in supply, and speculation (i.e. the financialization of commodities).

5.2.1 The Great Recession

The impact of the Great Recession on global demand is easily seen in Figure 3, which shows a significant drop in total demand in both 2008 and 2009. Importantly for the econometric findings, those two years are the only periods where the demand for light petroleum products fell

29. Spare capacity is sometimes measured to be negative and utilization rates sometimes exceed 100 percent. This occurs because the capacity data is designed to take into account downtime at units. Utilization rates above 100 percent are therefore possible, although suggestive that the machinery is being pushed to its physical limits.

relative to residual fuel oil. Despite demand falling sharply, refinery capacity additions continued throughout the period, a natural outcome of the long-lead times associated with refinery expansions and the unexpected nature of the recession. Figure 4 shows this for upgrading capacity, with 2009 showing the largest increase for those years for which we have data.

Is it possible to link the drop in demand and increase in capacity with the start of a new regime where 1) price differentials between light crude and other grades both shrink and become less volatile and 2) the same thing occurs to the price differentials between lighter petroleum products and residual fuel oil? To answer this, we make use of some simple, qualitative supply-demand diagrams to model the relative supply and relative demand of lighter petroleum products to residual fuel oil. This approach allows us to easily consider the impact of reduced demand and increased upgrading capacity on both product price differentials and oil price differentials.³⁰

Figure 6 plots hypothetical supply and demand curves for 2007 and 2009. Before discussing what the figure says, we provide some underlying assumptions about the curves. The relative demand curve is inelastic. We believe this is reasonable because it is very difficult to substitute away from gasoline or diesel towards residual fuel oil as a transportation fuel, due to the latter's physical properties. The supply curve is upward sloping as refiners can alter the mix of products produced by adjusting utilization rates of upgrading capacity, even if the relative supplies of different types of crude oil are fixed. Importantly, though, there comes a point at which utilization rates approach their maximum levels and capacity starts to be constrained. It then becomes increasingly difficult to make further changes to the product mix. Hence, the supply curve also has an extreme upward sloping portion.

 $\frac{P_{light}}{P_{residual}}$ D_{2009} D_{2007} S_{2007} D_{2007}

Figure 6: Price differentials lowered by Great Recession and capacity additions

In terms of demand shifters, we believe it is reasonable to assume the relative demand curve would shift to the right every year due to the growing global economy and the desire for

^{30.} Supply-demand diagrams for the relative supply and demand of high vs. low quality crude oil can also be used but the intuition requires more work to develop so we focus on the product side.

greater use of transportation fuels. Indeed, consumption data show the only exception to that trend occurred in 2008 and 2009. With regard to supply shifters, we expect increased upgrading capacity to shift the relative supply curve outward (see discussion in Section 2.1). We also expect the curve to shift out whenever the relative supply of light crude oil to low-grade oil increases (see Figure 2), although the data suggests that was not an important factor during this period.

The shifts in the curves from 2007 to 2009 reflect two important changes found in the data: (1) the relative demand for lighter products declined significantly due to the recession, i.e. the relative demand curve shifted to the left; (2) the amount of upgrading capacity continued growing in 2008 and 2009, i.e. the relative supply curve shifted to the right. If the shifts were big enough, the diagram predicts a large drop in the product price differential, which would then contribute to a large drop in the differentials between light crude and other grades of crude oil. The model would also predict a sharp drop in utilization rates for upgrading capacity. Figure 5 is in line with this, and media reports from this time confirm the impact of the Great Recession on coking utilization rates in the U.S. (see Evans, 2009).

Figure 6 also provides intuition on what it would take to move from a regime with relatively large and volatile price differentials to one with relatively small and less volatile price differentials. It requires moving away from the upward sloping portion of the supply curve to the relatively flatter portion of the supply curve and staying there. Or, in plainer English, moving away from a regime where upgrading capacity is relatively constrained, i.e. close to full utilization, to one where it is not. In line with the caveats we mentioned at the beginning of this section, absent a formal model and without more complete refining data, we cannot make strong claims about whether this has actually happened globally. However, Figure 5 provides suggestive evidence of such a change in coking capacity.

If the intuition form the model and our interpretation of the data is correct, then the sharp drop in demand due to the Great Recession, in essence, allowed capacity additions to catch up with demand. The fact that price differentials have remained smaller and less volatile since then suggests that global capacity additions after the Recession have been sufficient, in light of the other trends discussed in Section 5.1, to meet growing demand for gasoline and diesel, without leading to an oversupply of residual fuel oil.

5.2.2 Changes in supply

Most notably, global crude oil production grew little from 2005 to 2010. In regards to quality, table 7 shows some variation in the overall supply of different types of crude when comparing 2010 to 2005, but none of the changes are substantial. The monthly worldwide crude production including lease condensate series, from the U.S. Energy Information Administration, showed only one notable change around this time: A significant increase in OPEC production, primarily from Middle Eastern countries, which started in late 2007. However, that was fully reversed by the end of 2008 after the Great Recession hit.

5.2.3 The potential role of speculation

In surveying the oil price shock of 2007 and 2008, Hamilton (2009) discussed the potential role of speculation and the financialization of commodities. More generally, the impact of speculation on commodity markets has received a significant amount of attention in the literature. While this literature has covered many topics, we were unable to find any guidance on how speculation

or the financialization of commodities might affect the differentials between light crude and other types of crude oil. One likely reason is that outside of the two major benchmarks, Brent and WTI, futures markets for crude oil are generally not well developed and publicly available data is often quite limited.

With that being said, several futures contracts for Oman and Dubai crude have traded since the first decade of the 2000s and other contracts have come into existence after 2010, for example for Mars crude. However, open interest for those contracts is a small fraction compared with WTI and Brent-based contracts. Furthermore, none of those contracts have experienced a sustained, robust growth in open interest. While this, of course, does not rule out a role for speculation, it suggests the forces at play in WTI and Brent futures markets are much less important when it comes to futures contracts for other crude oils.

6. CONCLUSION

Crude oil can vary significantly in some key physical properties, making various streams imperfect substitutes for each other and leading to the existence of price differentials among crude oils. In a certain sense, these differentials reflect the limits to arbitrage that exist across crude oil quality. In this paper, we documented that a large number of differentials between crude oils of different types have experienced structural breaks where their means have become smaller over time. In particular, we show that many price differentials between different grades of crude oil experienced a major break around the time of the Great Recession.

We document several fundamental, long-lasting changes in the oil market that we believe are consistent with the changes seen in these price differentials. One is the fact that the global refining sector has become increasingly complex over time, as upgrading capacity additions have increased the ability of the sector to transform lower-grade crude oil into high-valued petroleum products. The other is the shale boom, which has unexpectedly increased the relative supply of light crude oil, reducing, on the margin, the need for such complex refineries. This narrowing of the differentials has occurred despite the fact that increasingly stringent environmental regulations on sulfur content in petroleum products and a persistent decline in the use of residual fuel oil should be pushing them apart.

We believe a number of possible avenues exist for future research. For one, our paper has focused on changes in the long-run means of crude oil price differentials. More sophisticated time-series analysis could try to disentangle the structural factors behind the short-run dynamics of those differentials. Additionally, further empirical analysis, which would be possible with better data, would also help answer a number of questions about the quantitative importance of various factors affecting the differentials.

ACKNOWLEDGMENTS

This work has benefitted greatly from comments by the three anonymous referees as well as comments by participants at The Society for Nonlinear Dynamics and Econometrics 2019 conference, the 42nd IAEE International Conference, the Commodity and Energy Markets Association Annual Meeting 2019, the 3rd Annual JPMCC International Commodities Symposium, the EIA 2019 Workshop on Financial and Physical Energy Market Linkages, and the 37th USAEE/IAEE North American Conference. We also thank Julie Harris and John Powell at the U.S. Energy Information Administration for useful discussions on refineries and U.S. refining data, as well as Mark

Agerton, Nathan Balke, Alexander Chudhik, Reinhard Ellwanger, Ben Gilbert, Lutz Kilian, Kei-Mu Yi and Mine Yucel for helpful comments. Any errors or omissions are our own. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

REFERENCES

- Adelman, M. A. (1984). "International oil agreements." *The Energy Journal* 5: 1–9. https://doi.org/10.5547/issn0195-6574-ej-vol5-no3-1.
- Agerton, M. and G.B. Upton Jr. (2019). "Decomposing crude price differentials: Domestic shipping constraints or the crude oil export ban?" *The Energy Journal* 40: 155–172. https://doi.org/10.5547/01956574.38.1.mage.
- Aizhu C. (2000). "China's Sinopec maps out 20 pct refinery expansion." Reuters News, December 14, 2000.
- Andrews, D.W.K. (1991). "Heteroskedasticity and autocorrelation consistent covariance matrix estimation." *Econometrica* 59: 817–858. https://doi.org/10.2307/2938229
- Andrews A. and R. Pirog and M.F. Sherlock (2010). "The U.S. oil refining industry: Background in changing markets and fuel policies." *Congressional Research Service Report*, November, 2010.
- Bachmeier, L.J. and J.M. Griffin (2006). "Testing for market integration: crude oil, coal, and natural gas." *The Energy Journal* 27: 55–71. https://doi.org/10.5547/ISSN0195-6574-EJ-Vol27-No2-4.
- Bacon, R. and S. Tordo (2005). "Crude oil price differentials and differences in oil qualities: A statistical analysis." ESMAP Technical Paper, October, 2005.
- Bai, J. (1997). "Estimating multiple breaks one at a time." Econometric Theory, 13: 315–352. https://doi.org/10.1017/s0266466600005831.
- Bai, J. and P. Perron (1998). "Estimating and testing linear models with multiple structural changes." *Econometrica*, 66: 47–78. https://doi.org/10.2307/2998540.
- Bai, J. and P. Perron (2003). "Critical values for multiple structural change tests" *The Econometrics Journal*, 6: 72–78. https://doi.org/10.1111/1368-423x.00102.
- Borenstein, S. and R. Kellogg (2014). "The incidence of an oil glut: Who benefits from cheap crude oil in the Midwest?" *The Energy Journal*, 35: 15–33. https://doi.org/10.5547/01956574.35.1.2.
- Buyuksahin, B. and T.K. Lee, J.T. Moser, and M.A. Robe (2013). "Physical markets, paper markets and the WTI-Brent spread." *The Energy Journal*, 34: 129–151. https://doi.org/10.5547/01956574.34.3.7.
- Cook, L.J. (2006). "Canadian oil sloshes in; rather than moving north from the Gulf, crude is coming south from Alberta's sands." Houston Chronicle, March 4, 2006.
- Dipaola, A. (2019). "Arcane pollution rule may cost Saudis and neighbors billions." Bloomberg, March, 26 2019.
- Energy Information Administration (1990). "The U.S. Petroleum Refining Industry in the 1980s." Energy Information Administration, 1990.
- Energy Intelligence Research (2009). "The International Crude Oil Market Handbook 2009." Energy Intelligence Research, Houston.
- Enos, J. L. (1958). "A measure of the rate of technological progress in the petroleum refining industry." *The Journal of Industrial Economics* 6: 180–197. https://doi.org/10.2307/2097628.
- Enos, J.L. (1962). "Invention and innovation in the petroleum refining industry." *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton University Press, Princeton 1962. https://doi.org/10.1515/9781400879762-012.
- Evans, B. (2009). "Valero coker closure part of growing trend; U.S. refiners battling narrow crude spread through next year: analyst." *Platts Oilgram News*, 87. https://doi.org/10.1016/B978-0-7020-2933-2.00011-3.
- Evans, B. and J. Mowler (2002). "Switch to sour crude haunts U.S. refiners." Platts Oilgram News, 80.
- Fattouh, B. (2010). "The dynamics of crude oil price differentials." *Energy Economics* 32: 334–342. https://doi.org/10.1016/j.eneco.2009.06.007.
- Ghoshray, A. and T. Trifonova (2014). "Dynamic adjustment of crude oil price spreads." *The Energy Journal* 35: 119–136. https://doi.org/10.5547/01956574.35.1.7.
- Giulietti, M., A.M. Iregui, and J. Otero (2015). "A pair-wise analysis of the law of one price: Evidence from the crude oil market." *Economics Letters* 129: 39–41. https://doi.org/10.1016/j.econlet.2015.02.002.
- Greenley H.L. (2019). "The world oil market and U.S. policy: Background and select issues for Congress." *Congressional Research Service Report*, April, 2019.
- Gregory A.W. and B.E. Hansen (1996). "Residual-based tests for cointegration in models with regime shifts." *Journal of Econometrics* 70: 99–126. https://doi.org/10.1016/0304-4076(69)41685-7

Gülen, G. (1997). "Regionalization in the world crude oil market." *The Energy Journal* 18: 109–126. https://doi.org/10.5547/issn0195-6574-ej-vol18-no2-6.

Gülen, G. (1999). "Regionalization in the world crude oil market: Further evidence." *The Energy Journal* 20: 125–139. https://doi.org/10.5547/issn0195-6574-ej-vol20-no1-7.

Hamilton, J.D. (2009). "Causes and consequences of the oil shock of 2007–08." National Bureau of Economic Research, 2009. https://doi.org/10.3386/w15002.

Hammoudeh, S.M., B.T. Ewing, and M.A. Thompson (2008). "Threshold cointegration analysis of crude oil benchmarks." *The Energy Journal* 29: 79–95. https://doi.org/10.5547/issn0195-6574-ej-vol29-no4-4.

Jegarajah, S. (2004). "Crude oil rises as consumer report raises world demand forecast." Bloomberg, March 12, 2004.

Johnston, D. (1996). "Refining report: Complexity index indicates refinery capability, value." Oil&Gas Journal March 18, 1996.

Kemp, J. (2009). "Saudi move is bid to realign oil market." Reuters, October 29, 2009.

Kerr, R. and R.G. Newell (2003). "Policy-induced technology adoption: Evidence from the U.S. lead phasedown." The Journal of Industrial Economics 51: 317–343. https://doi.org/10.1111/1467-6451.00203.

Khrennikova, D. and E. Mazneva (2018). "Russian oil set to lose billions in ship-fuel overhaul." *Bloomberg*, October 29, 2018.

Kilian, L. (2009). "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market." *American Economic Review* 99: 1053–1069. https://doi.org/10.1257/aer.99.3.1053.

Koster, F. (2017). "Market eye: Switch to lighter crude squeezes US fuel oil supply." The Oil Daily, August 25, 2017.

National Energy Board (2018a). "Canada's Energy Future 2018: An Energy Market Assessment." *National Energy Board*, 2018.

National Energy Board (2018b). "Canada's Energy Future 2018: Oil Sands Production Supplement." *National Energy Board*, 2018.

Perron P. (1989). "The great crash, the oil price shock, and the unit root hypothesis." *Econometrica* November: 1361–1401. https://doi.org/10.2307/1913712.

Perron P. (2006). "Dealing with structural breaks." Palgrave Handbook of Econometrics 1: 278-352.

Peterson, D.J. and S. Mahnovksi (2003). "New forces at work in refining: Industry views of critical business and operations trends." *RAND Monographs*, 2003.

Petroleum Intelligence Weekly (2005a). "Skeptics query." Petroleum Intelligence Weekly, February 21, 2005.

Petroleum Intelligence Weekly (2005b). "U.S. refining trends ride on shifting spreads." *Petroleum Intelligence Weekly*, June 6, 2005.

Piotrowksi, M. (2009). "Tight differentials for heavy/sour crudes hurt complex refiners." The Oil Daily, March 26, 2009.

Pirog R. (2007). "Petroleum refining: Economic performance and challenges for the future." *Congressional Research Service Report*, March, 2007.

Platts (2018). "Methodology and Specifications Guide: Crude Oil." Platts.

Reuters (1990). "Asia-Pacific oil demand seen rising 30 pct 1988-95." Reuters News, September 19, 1990.

Reuters (2005b). "New U.S. oil refineries unlikely—Exxon." Reuters News, February 15, 2005.

Reuters (2005a). "Refining capacity to stay tight, profits high." Reuters News, July 25, 2005.

Ripple, R. and P.R. Wilamoski (1995). "Is the world oil market one great pool?: revisited." *OPEC Review*, 19: 283–292. https://doi.org/10.1111/j.1468-0076.1995.tb00530.x.

Samuelson, P.A. (1952). "Spatial price equilibrium and linear programming." American Economic Review 42: 283-303.

Sauer, D. G. (1994). "Measuring economic markets for imported crude oil." The Energy Journal 15: 107–124. https://doi.org/10.5547/issn0195-6574-ej-vol15-no2-6.

Scheitrum, D.P., C.A. Carter, and C. Revoredo-Giha (2018). "WTI and Brent futures pricing structure." *Energy Economics* 72: 462–469. https://doi.org/10.1016/j.eneco.2018.04.039.

Scott, N. (2006). "Market for Canadian crude oil is rapidly picking up momentum." The Wall Street Journal, April 19, 2006.

Subrahmaniyan, N. (2005). "China's \$3 billion refinery expansions may cut costs." Bloomberg, June 20, 2005.

Tuttle, R. (2019). "Ship-fuel rule to keep refiners running hard as 2020 approaches." Bloomberg, April 13, 2019.

van Schaik, J. (2019). "Lower fuel oil supply helps ease IMO 2020 shift." The Oil Daily, April 22, 2019.

Weiner, R.J. (1991). "Is the world oil market 'One Great Pool'?" *The Energy Journal* 12: 95–107. https://doi.org/10.5547/issn0195-6574-ej-vol12-no3-7.

Yong, L. (2004). "China's demand for African oil helps boost gas prices." Bloomberg, May 6, 2004.