Federal Reserve
Bank of Dallas

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

---

Michael Plante and Grant Strickler

Working Paper 1901

Research Department

https://doi.org/10.24149/wp1901

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

Michael Plante[†] and Grant Strickler[‡]

February 5, 2019

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 quality-related differentials experienced a major structural break in or around 2008, after which there was a marked reduction in their means and, in many cases, volatilities. Several factors explain these shifts, including 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. 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. We also show that these structural breaks can affect tests of stationarity for many price differentials.

Keywords: crude oil price differentials, oil, structural breaks, stationarity

JEL Codes: Q40, C22

---

[*] We thank Alexander Chudik and Kei-Mu Yi for helpful comments and suggestions on the work. 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. 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.

[†] Corresponding author: Michael Plante, Federal Reserve Bank of Dallas, 2200 N. Pearl St., Dallas, TX 75201, michael.plante@dal.frb.org.

[‡] Grant Strickler, Federal Reserve Bank of Dallas, 2200 N. Pearl St., Dallas, TX 75201, grant.strickler@dal.frb.org.

1 Introduction

Although crude oil is often considered a homogenous good, it can have a wide range of physical characteristics that generate price differentials between different crude oils. These differentials are important for many oil market participants. For refiners, they can impact profitability and influence investment decisions about specific equipment, such as cokers, that could make it more profitable to process lower grades of crude.[1] Oil producers are concerned about these differentials because of the impacts they can have on revenues earned from producing certain types of oil. They can also affect a government's choice of the benchmark used to set official selling prices.[2] 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 are changing over time for one type of crude relative to others.

In 1984, Morris Adelman famously wrote, "The world oil market, like the world ocean, is one great pool." Yet, the fact that large differentials exist between crude streams of different qualities suggests that statement is not quite true. If all crude oil streams were equally substitutable for one another, then in the long-run the size of the price differential between any pair of crude oils should be relatively small, reflecting primarily transportation costs. But, in general, this has historically not been the case. In a certain sense, then, these differentials also reflect the limits that exist on the 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 such quality-related differentials 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 different 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 quality-related differentials vary over time in response to changing market conditions, little has been said about whether they have been affected by structural breaks that have more permanently changed their average levels.[3]

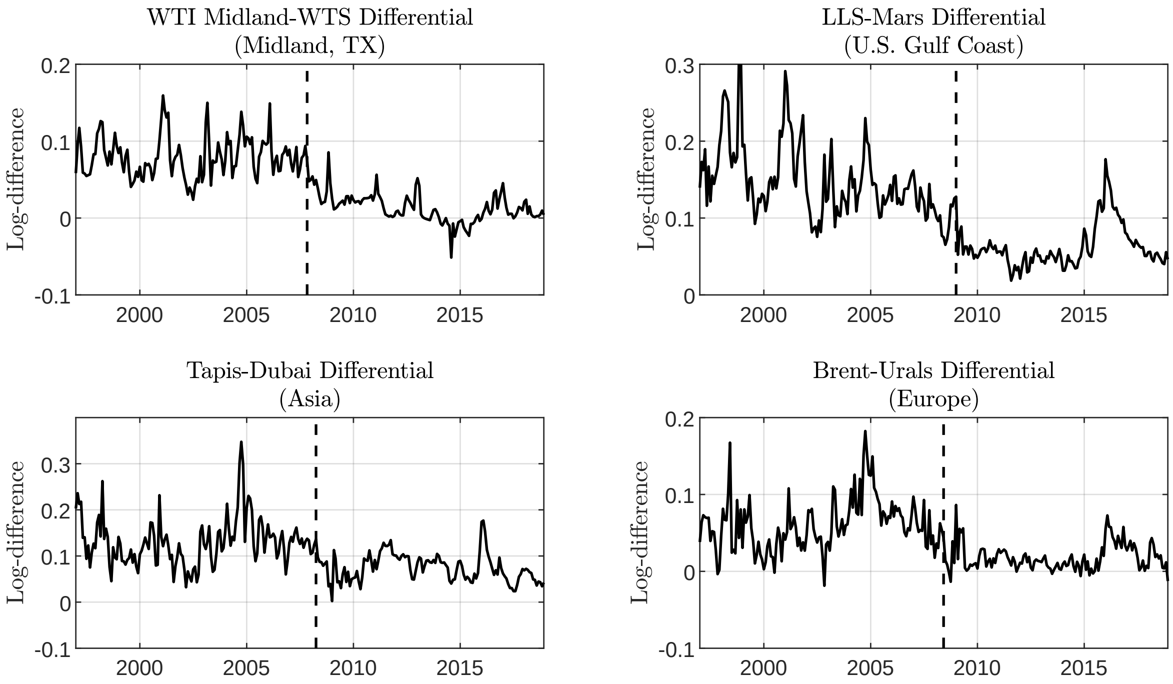
---

[1] This topic has received attention from trade press and market analysts since at least the early 2000s. See, for example, Evans and Mowler (2002), PIW (2005) and Piotrowksi (2009). More recently, the shale boom and IMO 2020 have both generated renewed interest in these issues.

[2] 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.

[3] Prior work in the literature has discussed the occurrence and importance of such breaks regarding differentials related to key benchmarks for light, sweet crude, such as West Texas Intermediate (WTI)

Figure 1.1: Oil price differentials in four areas



Notes: Figure plots differentials between a higher and lower grade crude oil from 1997 to 2018 in four areas of the world. 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.

To provide some motivation for our interest in structural breaks, we plot in Figure 1.1 examples of 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, considering the price of the 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 quality-related differentials, not shown here, strongly suggests the existence of structural breaks in many of their means, as well.

Our first contribution is to more rigorously and systematically document 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, a large number of these quality-related differentials—24 out of 25 cases to be

---

and Brent. See, for example, Buyuksahin et al. (2013), Borenstein and Kellogg (2014), Agerton and Upton (Forthcoming), and Scheitrum et al. (2018). In contrast to those works, our main focus is on price differentials between crude oils of differing qualities.

exact—experienced a significant break around 2007 and 2008. We then investigate how the means have shifted over time and find that most quality-related differentials have narrowed. In many cases, particularly those that experienced a break around 2008, the reduction in the mean has been accompanied by a major drop in volatility. After the break, the means and volatilities are often half their pre-break levels.

Price differentials between different types of oil reflect the limits to arbitrage that exist across crude oil quality, and a major decline in their average values points to a change in the extent of this arbitrage over time. More specifically, arbitrage opportunities for substituting between high and low-quality crudes that existed in the 1990s and parts of the 2000s have been reduced dramatically. We discuss several possible explanations for why this has happened. One factor is the continued global buildup of more complex refineries, which have the ability to increase the production of high-value petroleum products, such as gasoline and diesel, from low-grade crude oils. Another is the U.S. shale oil boom, which has unexpectedly boosted the supply of high-grade, light crude oil. This has reduced, on the margin, the need to have more complex refineries to process low-grade crude oils. We also discuss, but rule out, the possibility that changes in environmental regulations or an increase in demand for residual fuel oil, which is produced in greater abundance in lower grade crudes, could have led to the breaks. On the contrary, our analysis shows those two forces should have led to wider, not smaller, differentials over time.

Regarding the cluster of breaks around 2008, we find that the Great Recession played a role in this outcome. We first present some evidence that utilization of refining capacity to upgrade lower-quality crude oils was at relatively high levels prior to the Recession. We then show that there were some fairly significant capacity additions during the Recession, which occurred during a period of falling demand for petroleum products due to the Recession itself. Together, these two forces sharply reduced utilization rates for upgrading capacity in 2008 and 2009. Although demand for petroleum products has been growing since then, there has also continued to be significant capacity additions for upgrading low-quality crude oil and utilization rates have remained at more modest levels. This suggests the additional upgrading capacity has generally been sufficient to meet incremental demand, helping to keep quality-related differentials at relatively low levels.

We also investigated whether oil price differentials between crudes of the same type, for example, two light, sweet crude oils, have 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 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. A modest contribution on our part is to show that these breaks are more prevalent than previously documented in the literature.

Finally, it has long been known that structural breaks such as the ones we document can affect tests of stationarity and we show this is indeed the case for many of the oil price differentials we consider. For differentials constructed using daily data, we find 4 out of 38 differentials where a unit root test fails to reject the null of a unit root. With monthly data, we find 34 differentials out of a possible 59. Once we allow for a possible break in mean, the tests almost always reject the null of a unit root. Overall, these results show that some caution should be applied when using standard unit root tests to (log) oil price differentials.

We note that our work is connected with previous research papers, such as Weiner (1991), Sauer (1994), Gülen (1997) and Gülen (1999), 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. Our work differs from the previous literature in our focus on structural breaks in the long-run average size of quality-related differentials.[4]

The rest of the paper is organized as follows. A brief introduction to crude quality and oil price differentials is contained in Section 2. 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 some potential explanations for our findings while Section 6 provides evidence that the structural breaks we document can affect stationarity tests. 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 due to 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 different 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.

---

[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.

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 sell at a premium to other grades, while heavy, sour crude oils usually sell at a discount to all other grades. In this section we discuss why these physical characteristics generate such price differentials. Although it is not the focus of our paper, we briefly discuss how transportation costs, the direction of trade and infrastructure issues can also influence price differentials as these factors play a role in some of the results we present later.

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. This step is done by all refiners, from the simplest to most complex. In simplified terms, it is helpful to imagine that every crude oil can be distilled into three fractions: light products (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 products having the lowest densities and boiling points and the atmospheric residue having the highest density and boiling point.

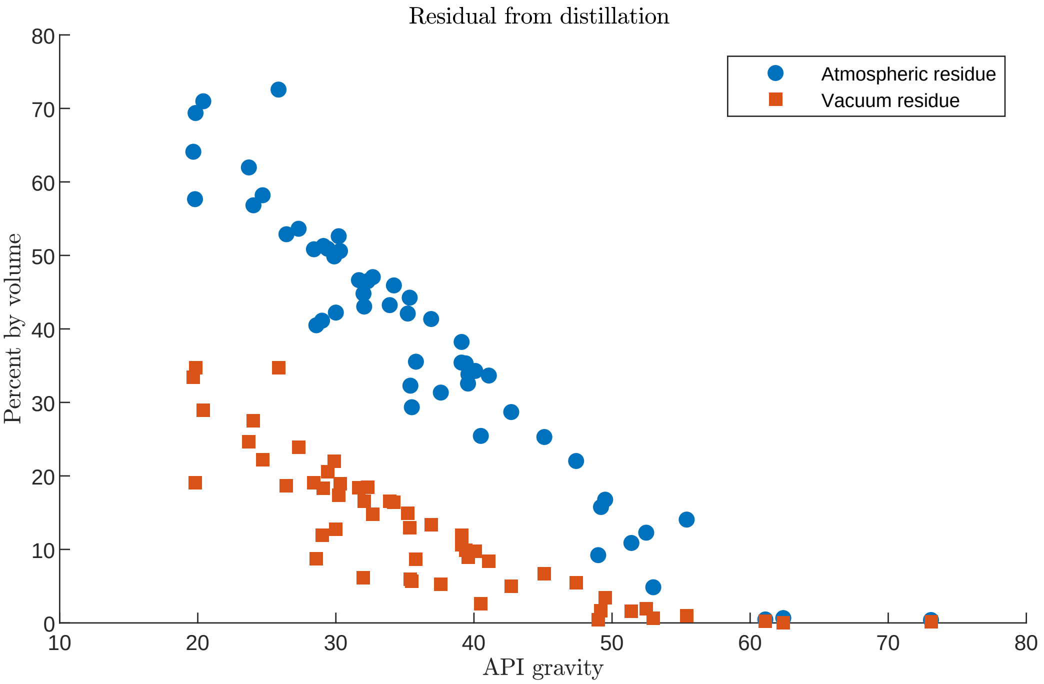
The API gravity of a crude is related to the proportions of the different cuts found in a specific crude oil. Light crudes, i.e. those with a high API gravity, tend to have greater proportions of gasoline and diesel than residual products, while medium and heavy crude oils usually contain greater amounts 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, and we now use some of these analyses to specifically discuss the relationship between API gravity and the residual content. As examples, 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 2.1 show the relationship between API gravity and the amount of atmospheric residue present for 54 crude oils.

One of the major differences between simple and more complex refineries is in the ability of the latter to transform the bottom of the barrel into other petroleum products. A simple

---

[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.

Figure 2.1: Heavy crude oils typically contain greater volumes of residual



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

refinery will have essentially no ability to do so. More complex refineries have additional machinery to convert the residual into higher valued petroleum products. This capital is often collectively referred to as secondary processing units, upgrading capacity or conversion capacity.

Moderately complex refineries will have a vacuum distillation unit (VDU), which further distills the residual from the CDU into vacuum gas oil (VGO) and vacuum residue, which is essentially residual fuel oil. They will also have crackers, machinery which processes the VGO into lighter products. The leftover, i.e. the residual fuel oil, is produced in greater concentrations in lower-quality crudes. The vacuum residue/residual fuel oil component of WTI is 9 percent, Brent 9.7 percent, Mars about 25 percent, and Maya 36.9 percent. The squares in Figure 2.1 show the relationship between API gravity and vacuum residue for a total of 54 crude oils.

The most complex refineries, in addition to having a VDU and crackers, will have a coker. This is an expensive piece of capital equipment that allows the refiner to break down the residual fuel oil left over from the VDU and transform it into lighter products and petroleum coke. The use of this equipment significantly reduces the amount of residual fuel oil produced from the refining process.

While the relative value of different refined products varies over time, 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 using crackers and cokers to increase the production of higher-valued products, while at the same time reducing the production of residual fuel oil from lower grade crude oils. As a result, the degree to which API gravity will play a role in price differentials should depend upon how costly this equipment is and, related to the cost, how widespread its use is. In the hypothetical case where the equipment were costless and available to everyone, API gravity would not contribute much to the formation of price differentials, since those differences would quickly be arbitraged away.

2.2 Environmental regulations and sulfur content

Sulfur is a pollutant, so environmental regulations in many countries require that various petroleum products meet strict specifications for sulfur content. Over time, these environmental regulations have become more stringent and more widespread. A series of policies regarding sulfur content in gasoline and diesel have been passed in the U.S., the EU and in China since 2000, for example, and regulations regarding the sulfur content in residual fuel oil used in shipping have also been tightened several times since then. Due to these policies, sulfur is usually removed at some stage of the refining process; doing so requires costly capital investments by refineries. These regulations, and the costs associated with compliance, lead to sour crude oils, i.e. those with high sulfur content, selling at a discount to sweet crude oils.

2.3 Transportation costs and infrastructure issues

Finally, although not the main focus of this paper, transportation costs, the direction of trade, and infrastructure issues can also play a role in price differentials. If an area is a net importer of a particular type of oil, say light, sweet oil, then the price of light, sweet crude in that area will build in the transportation cost associated with importing the marginal barrel. For example, up until the late 2000s, the price of light, sweet crude oil in the U.S., given by either Louisiana Light Sweet (LLS) or WTI, carried a premium over similar grade crudes produced in Europe, such as Brent. The literature previously pointed out that light, sweet crude differentials along the Gulf Coast (and in Cushing) have been subject to change as the shale boom has dramatically increased the supply of light, sweet crude in the area (see, for example, Buyuksahin et al. (2013) and Agerton and Upton (Forthcoming)).

Infrastructure issues, such as pipeline bottlenecks, can also affect price differentials between crude oils. This is true both for crudes of different types as well as the same stream of crude that is priced in more than one location. Over the last 10 years, these issues have become particularly prominent in the pricing of crude oil in Canada and the U.S. As production has grown in those two countries, large differentials have emerged at various points in time that have affected crude oil prices in Canada, the Bakken and the Permian Basin, to name a few locations.

3 Data and methodology

3.1 Prices

We work with a set of 12 crude oil prices. Table 3.1 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 Research (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, while 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 together crude oils of roughly the same characteristics. In reality, there is a continuum of quality. With that being said, our series include 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 is of reasonable quality.

All price series are daily and come from Bloomberg with the exception of 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 daily data for Urals is more limited and runs from 2002 to 2013. However we have monthly data for Urals available over the entire common sample which we also use in the analysis. 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, for two reasons. First, Bloomberg data for the current benchmark price, Western Canadian Select (WCS), only starts in 2008. Second, pipeline bottlenecks have had major impacts on differentials between WCS and other crude oils, making it difficult to discern longer-term trends that might be driven by quality-related factors. Given this, we have decided to exclude WCS prices from the analysis.

Finally, we also present some additional results using monthly averages. We do this for two reasons. First, for Urals crude we have a full sample and are able to show that our main conclusions regarding Urals-related differentials based on the daily data are not sensitive to the longer sample. Second, the use of monthly data allows us to expand the set of oil prices considered. Table A.3 in the appendix presents the full set of prices used for the monthly analysis. Additional details can be found in the data appendix.

Table 3.1: 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
Urals	31.5	1.44	Medium	Sour
Middle East/Asia				
Dubai	31.0	1.70	Medium	Sour
Oman	33.0	1.10	Medium	Sour
Tapis	44.6	0.03	Light	Sweet

3.2 Differentials

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

$$p_{ij,t} = \ln P_{i,t} - \ln P_{j,t}. \tag{3.1}$$

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 about 5400 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 it is 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 3.1. 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 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, the within-area differentials, being priced closer to each other, should be less affected by transportation costs or infrastructure issues and better reflect the role of arbitrage across quality. The across-area differentials, on the other hand, should reflect not only differences in quality but also arbitrage across space. For example, the LLS-Dubai differential builds in not only the effect of quality but also the fact that LLS may be influenced more heavily by local conditions in the United States Gulf Coast (USGC) market, while Dubai may be influenced a bit more by conditions in Asia. On the other hand, we expect the LLS-Mars differential, in general, to be heavily influenced by arbitrage across quality in the USGC.

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.

Table 3.2 shows some summary statistics for the series, based on a common sample starting from 1997 onwards.[7] In line with the previous literature, 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.[8] For example, the mean differential between LLS and HLS was only 1.5 percent while it was almost 23 percent for the LLS-Maya differential.

We also find that the greater the differences in API gravity and sulfur content, the more volatile the differential tends to be. One potential explanation for this is that the degree of substitutability between any two grades of crude is inversely related to the quality differences of those crudes. For example, LLS and HLS are quite similar and as a result should be highly substitutable for each other in the refining complex in the USGC (and elsewhere). This should ensure that their prices generally do not deviate too far from one another, minimizing volatility in the percent-differential. On the other hand, LLS and Maya are very different from each other and refiners who prefer to process one over the other are likely to be hesitant to switch back and forth over short periods of time. This could require large price swings to clear the market, which would lead to volatile differentials.

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. Due to the direction of trade at the start the sample, i.e. the Gulf Coast was a net importer of light crude, 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 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.

Table 3.3 presents summary statistics for the across-area differentials. The upper panel shows the summary statistics for differentials between crudes of different types, the lower panel for same types. In general, the statistics are similar in nature to the within-area differentials. On average, we find that the means are greater for those pairs of crudes with larger differences in their API gravity and sulfur content. However, a few differentials do

---

[7] Summary statistics for the full samples are available from the authors. They are generally not very different from the results in the table.

[8] See, for example, Bacon and Tordo (2005) and Giulietti et al. (2015).

Table 3.2: Oil price differentials within areas

Differential	API difference	Sulfur difference	Mean	Standard deviation
Midland, TX				
WTIM-WTS	5.0	-1.56	0.046	0.042
U.S. Gulf Coast				
LLS-HLS	2.0	0.05	0.015	0.016
LLS-Mars	6.8	-1.61	0.108	0.061
LLS-Maya	14.6	-2.94	0.227	0.109
HLS-Mars	4.8	-1.66	0.094	0.056
HLS-Maya	12.6	-2.99	0.212	0.102
Mars-Maya	7.8	-1.33	0.118	0.064
Europe / Atlantic Basin				
Brent-Urals	6.6	-1.03	0.043	0.036
Middle East / Asia				
Tapis-Oman	11.6	-1.07	0.093	0.053
Tapis-Dubai	13.6	-1.67	0.103	0.055
Oman-Dubai	2.0	-0.60	0.010	0.020

Notes: These statistics are based on a sample from January 1997 to December 2018. For Brent-Urals, the sample runs from January 2002 to November 2013.

not exhibit this property. The LLS and HLS differentials with Brent are both positive while the two WTI-LLS differentials have negative means.

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).[9] 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,

$$p_{ij,t} = c_{ij} + u_{ij,t}, \tag{3.2}$$

---

[9] Perron et al. (2006) provides a more general overview of structural breaks.

Table 3.3: Oil price differentials across areas

Crudes of different type				
Differential	API difference	Sulfur difference	Mean	Standard deviation
Light-medium differentials				
Tapis-Urals	13.1	-1.41	0.099	0.049
Tapis-Mars	15.7	-2.02	0.125	0.061
Brent-Oman	5.1	-0.69	0.040	0.044
Brent-Dubai	7.1	-1.29	0.050	0.047
Brent-Mars	9.2	-1.64	0.072	0.046
LLS-Oman	2.7	-0.66	0.078	0.066
LLS-Urals	4.2	-1.00	0.080	0.059
LLS-Dubai	4.7	-1.26	0.087	0.069
HLS-Oman	0.7	-0.71	0.062	0.062
HLS-Urals	2.2	-1.05	0.065	0.052
HLS-Dubai	2.7	-1.31	0.072	0.065
Light-heavy differentials				
Tapis-Maya	23.5	-3.35	0.244	0.098
Brent-Maya	17	-2.97	0.190	0.086
Medium-heavy differentials				
Oman-Maya	11.9	-2.28	0.150	0.075
Urals-Maya	10.4	-1.94	0.129	0.060
Dubai-Maya	9.9	-1.68	0.141	0.077
Crudes of similar type				
Light-light differentials				
WTIC-LLS	3.3	-0.10	-0.040	0.059
WTIM-LLS	3.3	-0.10	-0.057	0.076
LLS-Tapis	-8.9	0.41	-0.016	0.050
LLS-Brent	-2.4	0.03	0.037	0.045
HLS-Tapis	-10.9	0.36	-0.031	0.048
HLS-Brent	-4.4	-0.02	0.022	0.042
Medium-medium differentials				
Oman-Urals	1.5	-0.34	0.001	0.035
Oman-Mars	4.1	-0.95	0.032	0.049
Urals-Dubai	0.5	-0.26	0.011	0.034
Urals-Mars	2.6	-0.61	0.016	0.037
Dubai-Mars	2.1	-0.35	0.022	0.053

Notes: These statistics are based on a sample from January 1997 to December 2018. For Urals differentials, the sample runs from January 2002 to November 2013.

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.[10]

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. $0 \leq m \leq 5$. 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, $h = \epsilon T$. 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.[11]

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.[12] 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).[13] 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, that is 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

---

[10] 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 necessarily surprising given the nature of the breaks we are investigating. Given our interest in testing for breaks in the mean, and given our concern about potential misspecification of the model, we decided to work with the more parsimonious model in equation (3.2).

[11] For Urals differentials, where we have a smaller sample, we set the trimming parameter to 0.25 so that the minimum regime size is as close as possible to 3 years.

[12] 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.

[13] 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.

procedure. In the case of two breaks, the repartition process re-estimates the breaks using the subsamples $[1, k_2^0]$ and $[k_1^0, T]$. 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 for more than 1 break we then report the results for all of the cases up to and including the last break which is statistically significant at a 1 percent level. This occurs for only 2 cases.

4 Results

4.1 Identifying 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-light differentials, are introduced after.

We begin with the WTIM-WTS differential, which is a differential between a light, sweet crude and a light, sour crude. To formally test for the presence of a break, we use the procedure outlined in Section 3.3. The test identifies two breaks that are significant at a 1 percent level. The first 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. These dates and test statistics are listed at the top of the upper panel in Table 4.1. 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 or 1 break.

The middle portion of the upper panel in Table 4.1 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 it is only significant at a 5 percent level and hence not listed in the table.

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 2007 and 2008. The test also identifies a later break in the Brent-Urals differential.

We next consider the across-area differentials for different crude types, with the results presented in the bottom panel of Table 4.1. The test finds that all of the differentials, with just one exception, experienced a break around 2008. For the light-medium differentials, the test also identifies a few other breaks in the mid-2000s involving light crude in the USGC, and two breaks after 2010 involving Mars.

As shown in Table 4.1, a very large number of breaks occurred between 2007 and early 2009. An immediate question of interest to us was whether this break affected oil price differentials generally 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-light and medium-medium pairs. The results from those tests are shown in Table 4.2. 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. For the within-area differentials we find two breaks impacting the LLS-HLS differential after the start of the shale boom. For the light-light differentials in the across-area group, we find a set of breaks in the mid-2000s and another set during the shale boom. The test also identifies a set of potentially related breaks for the medium-medium differentials that include Mars crude.

4.2 Shifts in means across regimes

Tables 4.3, 4.4 and 4.5 display how the means have changed over regimes. The differentials are grouped in a similar manner to Tables 4.1 and 4.2. The final column shows how the means have changed from the initial to final regime. We do not exhaustively discuss all of the changes in the tables but instead highlight several key findings.

Quality-related differentials have shrunk over time: For the within-area differentials between crudes of different types, shown in Table 4.3, we find that most means have shrunk in half, at least. Substantial declines in the USGC on the order of 10 to 15 percentage points have occurred. In regards to the across-area differentials between crudes of different

Table 4.1: Breakpoint test results for crudes of different qualities

Part 1: Within-area differentials

Differential	Break 1	Break 2	Break 3	F-statistic 0 vs. 1	F-statistic 1 vs. 2	F-statistic 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	05/2007	06/2010	-	183.78	26.56	-
Middle East/Asia						
Tapis-Oman	05/2008	-	-	29.78	-	-
Tapis-Dubai	05/2008	-	-	39.15	-	-

Part 2: Across-area differentials

Differential	Break 1	Break 2	Break 3	F-statistic 0 vs. 1	F-statistic 1 vs. 2	F-statistic 2 vs. 3
Light-medium						
Tapis-Urals	06/2007	-	-	37.52	-	-
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	04/2009	04/2005	-	87.14	14.24	-
LLS-Dubai	12/2008	05/2005	-	116.83	14.39	-
HLS-Oman	11/2008	-	-	89.49	-	-
HLS-Urals	12/2008	03/2005	-	94.72	22.56	-
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	06/2008	-	-	18.27	-	-

Notes: Dates refer to the month of the last day of a given regime. The order of the breaks is determined by the test. 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.

Table 4.2: Breakpoint test results for crudes of similar type

Part 1: Within-area differentials

				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	-
Middle East/Asia						
Oman-Dubai	-	-	-	-	-	-

Part 2: Across-area differentials

				F-statistic		
Differential	Break 1	Break 2	Break 3	0 vs. 1	1 vs. 2	2 vs. 3
Light-light						
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
LLS-Brent	05/2011	01/2005	-	120.02	38.89	-
HLS-Tapis	05/2004	-	-	60.35	-	-
HLS-Brent	01/2005	08/2013	-	90.40	59.25	-
Medium-medium						
Oman-Mars#	01/2002	08/2013	-	9.39	36.70	-
Urals-Mars	11/2005	-	-	22.83	-	-
Dubai-Mars	08/2013	03/2002	11/2005	22.52	14.42	19.14
Oman-Urals	06/2010	-	-	36.59	-	-
Urals-Dubai	07/2010	-	-	12.45	-	-

Notes: Dates refer to the month of the last day of a given regime. The order of the breaks is determined by the test. 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 # means the test failed to reject 0 vs. 1 break at 1 percent significance but did so for the null of 1 vs. 2 breaks.

types, shown in Table 4.4, we also find that most of the differentials have experienced large declines over time.

Table 4.4 shows that the means of differentials involving light crude in the Gulf Coast, i.e. LLS and HLS, and medium crudes outside the U.S. have declined more substantially than the corresponding light-medium differentials involving Brent and Tapis. One reason is the LLS and HLS differentials experienced a set of breaks in 2004 and 2005 that did not affect light crude oils outside the U.S. We note that the breaks in 2004 and 2005 are widespread, in the sense that they also impacted LLS and HLS differentials involving light crude oil outside the U.S. This can be seen in Table 4.5. Also, the decline in the means of those LLS and HLS differentials after the Great Recession has typically been larger than corresponding differentials involving Brent or Tapis crude. A likely explanation for this is shale boom, which has dramatically increased the supply of light crude in the area.

Across-area Mars differentials experience reversals during shale boom: One exception to quality-related differentials shrinking are the differentials between light crude oils outside the U.S. and Mars crude. While both the Brent and Tapis differentials to Mars experienced a break in 2008 where their means declined, they then experienced another break after that which reversed most of the initial decline. Interestingly, we find evidence for a similar reversal in Mars differentials involving Oman, Dubai and Urals crude. The timing of most of the breaks, which occurred in 2013, suggests they are connected with changing market conditions on the USGC due to the shale boom. The literature has previously documented numerous breaks in light crude differentials due to the shale boom. A modest contribution on our part is to document these breaks in Mars differentials.

Light-light differentials involving U.S. crude have shifted dramatically: As shown in the upper portion of Table 4.5, across-area differentials between USGC light crude and light crude oil outside the U.S. have gone from being positive to near-zero or negative. This occurred in two steps, with a block of breaks in the mid-2000s and another set of breaks after 2010. Likewise, differentials between WTI and LLS have gone from being near 0 to negative over time. These findings reconfirm and expand upon some previous results in the literature regarding structural breaks affecting light, sweet crude differentials.

Table 4.3: Regression constant across regimes for crudes of different types

Within-area differentials

Differential	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	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					0.108						0.055											-0.106
LLS-Maya	0.312										0.151												-0.161
HLS-Maya	0.292										0.142												-0.150
Mars-Maya	0.158										0.083												-0.075
Brent-Urals	–					0.072					0.029			0.012				–					-0.060
Brent-Urals$^{(m)}$	0.060													0.018									-0.042
Tapis-Oman	0.116											0.069											-0.063
Tapis-Dubai	0.130											0.074											-0.056

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.

Table 4.4: Regression constant across regimes for crudes of different types

Across-area: Light-medium

Differential	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Change
Tapis-Urals	–					0.128					0.075							–					-0.053
Tapis-Urals$^{(m)}$	0.122											0.074					0.056						-0.066
Tapis-Mars	0.149											0.080			0.110								-0.039
Brent-Oman	0.055											0.024											-0.031
Brent-Dubai	0.069											0.030											-0.039
Brent-Mars	0.088											0.034						0.080					-0.008
LLS-Oman	0.117												0.030										-0.086
LLS-Urals	–					0.140			0.094				.027					–					-0.113
LLS-Urals$^{(m)}$	0.131									0.075					0.014								-0.117
LLS-Dubai	0.143								0.101				0.035										-0.107
HLS-Oman	0.098												0.021										-0.077
HLS-Urals	–					0.119			0.074				0.026					–					-0.094
HLS-Urals$^{(m)}$	0.113								0.081			0.040				0.005*							-0.108
HLS-Dubai	0.124								0.083				0.026										-0.098

Across-area: Light-heavy

Differential	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Change
Tapis-Maya	0.308										0.186												-0.121
Brent-Maya	0.245											0.140											-0.105

Across-area: Medium-heavy

Differential	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Change
Oman-Maya	0.192										0.113												-0.078
Dubai-Maya	0.210						0.119																-0.091
Urals-Maya	–					0.155						0.100						–					-0.055
Urals-Maya$^{(m)}$	0.219					0.131																	-0.088

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

Table 4.5: Regression constant across regimes for crudes of the same type

Light-light differentials

<table>
<tr><th>Differential</th><th>97</th><th>98</th><th>99</th><th>00</th><th>01</th><th>02</th><th>03</th><th>04</th><th>05</th><th>06</th><th>07</th><th>08</th><th>09</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>Change</th></tr>
<tr><td>WTIC-LLS</td><td colspan="9">0.002*</td><td colspan="4">-0.035</td><td colspan="3">-0.138</td><td colspan="6">-0.055</td><td>-0.057</td></tr>
<tr><td>WTIM-LLS</td><td colspan="10">-0.008</td><td colspan="4">-0.043</td><td colspan="8">-0.125</td><td>-0.117</td></tr>
<tr><td>LLS-HLS</td><td colspan="14">0.019</td><td colspan="4">-0.002*</td><td colspan="4">0.013</td><td>-0.007</td></tr>
<tr><td>LLS-Tapis</td><td colspan="8">0.015</td><td colspan="6">-0.019</td><td colspan="4">-0.067</td><td colspan="4">-0.027</td><td>-0.042</td></tr>
<tr><td>LLS-Brent</td><td colspan="8">0.075</td><td colspan="6">0.039</td><td colspan="8">-0.005*</td><td>-0.080</td></tr>
<tr><td>HLS-Tapis</td><td colspan="7">-0.002</td><td colspan="15">-0.045</td><td>-0.043</td></tr>
<tr><td>HLS-Brent</td><td colspan="8">0.056</td><td colspan="9">0.016</td><td colspan="5">-0.019</td><td>-0.075</td></tr>
</table>

Medium-medium differentials

<table>
<tr><th>Differential</th><th>97</th><th>98</th><th>99</th><th>00</th><th>01</th><th>02</th><th>03</th><th>04</th><th>05</th><th>06</th><th>07</th><th>08</th><th>09</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>Change</th></tr>
<tr><td>Oman-Urals</td><td colspan="5">–</td><td colspan="8">0.008</td><td colspan="4">-0.018</td><td colspan="5">–</td><td>-0.026</td></tr>
<tr><td>Oman-Urals$^{(\mathrm{m})}$</td><td colspan="15">0.007*</td><td colspan="7">-0.012</td><td>-0.019</td></tr>
<tr><td>Oman-Dubai</td><td colspan="22">0.010</td><td>–</td></tr>
<tr><td>Oman-Mars</td><td colspan="5">0.053</td><td colspan="12">0.014</td><td colspan="5">0.051</td><td>-0.002</td></tr>
<tr><td>Urals-Mars</td><td colspan="5">–</td><td colspan="4">-0.003*</td><td colspan="8">0.024</td><td colspan="5">–</td><td>+0.027</td></tr>
<tr><td>Urals-Mars$^{(\mathrm{m})}$</td><td colspan="5">0.053</td><td colspan="4">-0.010</td><td colspan="8">0.020</td><td colspan="5">0.060</td><td>+0.007</td></tr>
<tr><td>Dubai-Mars</td><td colspan="5">0.043</td><td colspan="4">-0.025</td><td colspan="8">0.014</td><td colspan="5">0.046</td><td>+0.003</td></tr>
<tr><td>Urals-Dubai</td><td colspan="5">–</td><td colspan="9">0.006*</td><td colspan="3">0.022</td><td colspan="5">–</td><td>+0.016</td></tr>
<tr><td>Urals-Dubai$^{(\mathrm{m})}$</td><td colspan="22">0.010</td><td>–</td></tr>
</table>

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 $^{(\mathrm{m})}$ refers to results based on monthly data.

4.3 Less volatile differentials

Table 4.3 shows that the means of almost all quality-related differentials have declined over time. Figure 1.1 is also suggestive that there may have been changes in the volatility of those differentials, particularly for those experiencing a break around 2008. We decided to investigate this a little deeper by comparing the means and standard deviations of the quality-related differentials before and after 2008. While the actual breakpoint for many series varies, 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 4.6, with the results for the within and across-area differentials in the top and bottom panels, respectively. 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.

4.4 Breaks in residual fuel oil differentials

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

We calculated differentials between the spot price of high-sulfur residual fuel oil and the following spot prices, all for delivery in the Gulf Coast: heating oil, gasoline, LLS and Mars. Figure 4.1 plots the monthly time series of these differentials since 1997. The left panel shows the heating oil and gasoline differentials to fuel oil, while the right shows the differentials involving crude oil. There is remarkable similarity between these and many of the differentials plotted in Figure 1.1. We note here that this is not just a Gulf Coast phenomenon: the chart looks very similar if one uses product prices for New York Harbor and replaces LLS and Mars prices with Brent and Dubai.

More formally, we ran breakpoint tests on the Gulf Coast fuel oil differentials using daily data and found that all of them experienced a break in their mean around the same time that many quality-related oil price differentials 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.

Table 4.6: Summary statistics pre and post-break

Part 1: Within-area differentials

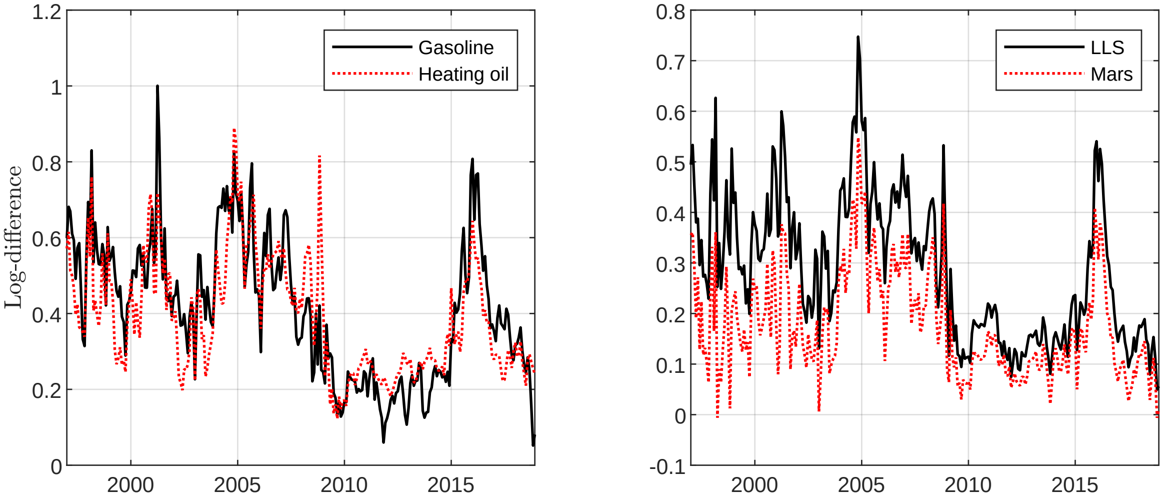
Differential	Pre-break Mean	Pre-break Standard deviation	Post-break Mean	Post-break Standard deviation	Ratio of mean (post/pre)	Ratio of std. dev. (post/pre)
Midland, TX						
WTIM-WTS	0.076	0.031	0.010	0.018	0.13	0.58
U.S. Gulf Coast						
LLS-Mars	0.147	0.052	0.063	0.029	0.45	0.56
LLS-Maya	0.299	0.089	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

Differential	Pre-break Mean	Pre-break Standard deviation	Post-break Mean	Post-break Standard deviation	Ratio of mean (post/pre)	Ratio of std. dev. (post/pre)
Light-medium						
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

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.

Figure 4.1: Residual fuel oil differentials



Notes: Figure plots log-differentials of spot prices for gasoline, heating oil, LLS and Mars relative to high sulfur fuel oil in the U.S. Gulf Coast using monthly data from January 1997 to December 2018.

4.5 Additional results using monthly data

We also repeated our analysis using monthly price data for a slightly larger set of crude oils. We add the following crude oils to our analysis: Algerian Saharan, Bonny Light, and Saudi Arabian Heavy prices to the US, Europe and Asia. The monthly data also provides an extended sample for Urals that starts before 1997 and runs up to the present. Additional details on the data and differentials can be found in the data appendix and in Tables A.3, A.4 and A.5 and the data appendix.

The breakpoint analysis uses a common sample that runs from January 1997 to December 2018. This gives us a total of 264 observations for each differential. The breakpoint results are shown in Tables A.6 and A.7. Tables A.8, A.9, and A.10 show how the means have shifted over time while Table A.11 shows the summary statistics for some of the monthly differentials pre and post-2008.

We only summarize our main findings. The interested reader is referred to the tables in the appendix for additional details. First, our main conclusions are not sensitive to the switch to monthly data. Although the statistical significance of some of the breaks is somewhat weaker than in the daily data, the breakpoint test identifies a large number of breaks affecting quality-related differentials between 2007 and 2009. More specifically, 38 out of 42 cases. Second, as Algerian Saharan and Bonny Light are light, sweet crude oils we are able to expand the number of tests looking for breaks in light-light differentials. We do not find any evidence of breaks around the Great Recession in differentials between those two crude oils and other light crudes. Consistent with our findings with the daily data, however, we do find evidence for two sets of breaks involving the differentials between the light, sweet crude oils in the U.S. Gulf Coast and both Algerian Saharan and Bonny Light. The first set occurs in late 2004 to early 2005, while the second set occurs after the start of the shale boom. Finally, the main conclusions regarding the Urals differentials hold in the monthly data which has the extended sample.

5 Discussion

The breakpoint test identifies when breaks occur but does not provide any information for why those breaks occurred. In this section, we bring to bear a host of data related to the refining sector and use this data to discuss, first, what factors can explain the breaks and, second, what role the Great Recession may have played in the timing of the breaks, i.e. why were there so many breaks in the means of quality-related differentials around 2008.

5.1 Factors behind the breaks

We are interested in explaining two key findings from the previous section: (1) most differentials between crudes of different types have shrunk since the start of our sample; (2) residual fuel oil differentials have also shrunk in a similar manner. In this section, we offer up several potential explanations related to longer-term changes in the oil market that could explain both of these findings and discuss if the available data support any of these possible explanations.

We first discuss whether changes in environmental regulations regarding sulfur emissions or recent trends in the use of residual fuel oil are possible explanations. We show that both of these factors cannot explain the narrowing of the differentials seen in the data and how both, on the contrary, should actually be contributing to a widening of differentials over time. We then discuss two other potential explanations: an increase in the ability of the global refining sector to process low-quality crude oil and the U.S. shale boom. Our analysis of the data leads us to conclude that these latter factors are plausible explanations for the breaks we have documented.

Environmental regulations tightened: A weakening of environmental regulations regarding sulfur emissions, particularly for residual fuel oil, would be a straight forward explanation for key findings (1) and (2). However, in general, standards regarding sulfur have actually been tightening over time, which should be pushing apart the quality-related differentials we have considered in this paper. These standards have been applied in several large consuming countries to a variety of petroleum products, including gasoline, diesel and residual fuel oil. Of particular note, requirements for residual fuel oil have been tightened several times since 2008. For example, the International Maritime Organization 2015 marine fuel rule set 0.1% as the maximum sulfur content allowed for emission control areas. The IMO plans to implement stiffer sulfur regulations on residual fuel used for shipping from a 3.5% sulfur content maximum to 0.5% starting in 2020.

Use of residual fuel oil declining over time: An increase in the demand for residual fuel oil relative to other petroleum products since 2008 is another explanation that could potentially explain key findings (1) and (2). However, the tightening of environmental regulations discussed in the previous paragraph has certainly worked against that trend. Moreover, consumption data does not seem to support this argument. We show this visually in Figure 5.1, where we plot annual changes in the consumption of residual fuel oil and all other petroleum products from 1997 to 2017. This data comes from various Annual Statistical Supplements from the International Energy Agency. Use of residual fuel oil has declined almost every year since 1997 while use of other petroleum products has been increasing at a relatively rapid pace, with the exception of the Great Recession. We will return to this point when we discuss the timing of the breaks.

Increasingly complex refining sector: Cokers and crackers increase 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, an increasingly complex refining sector is a natural candidate for explaining both the breaks in quality-related oil price differentials and the residual fuel oil price differentials.

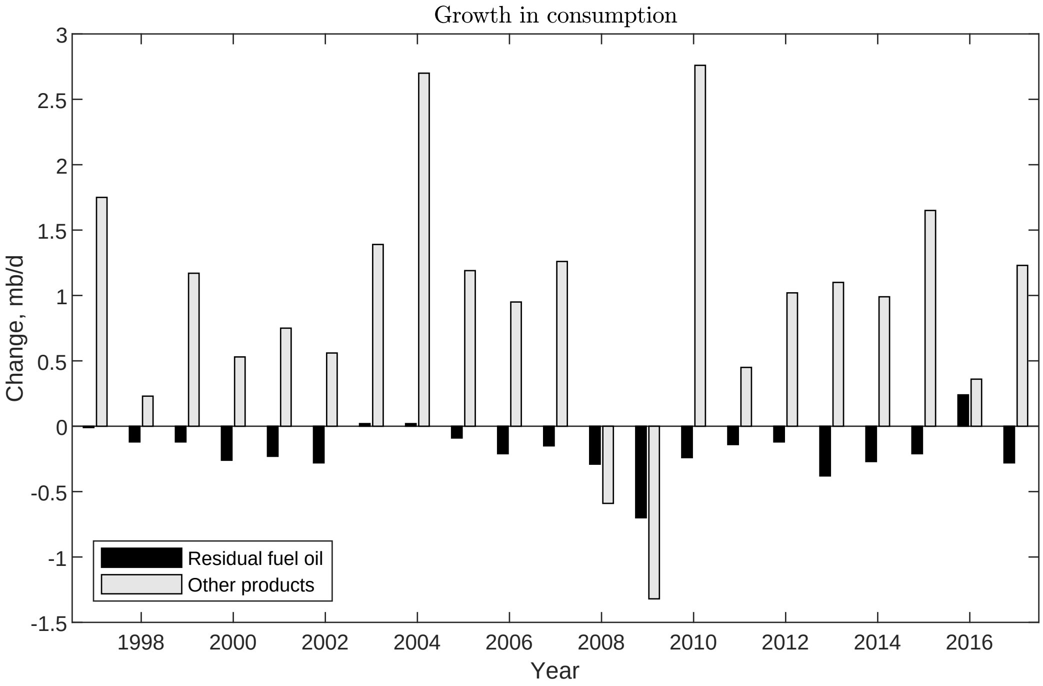
It has long been noted that the refining sector in the U.S. has become increasingly complex over time. Publicly available data for the rest of the world is limited but it also shows an increasing ability of the refining sector worldwide to convert medium and heavy crude oils into high-valued petroleum products. Our main discussion is based on data from Eni's World Oil Review and World Oil and Gas Review publications. Several other data sources are discussed afterwards.

Table 5.1 shows the data from Eni on global refining capacity, as well as two measures of how complex the refining sector is overall.[14] The second column shows data on primary capacity, which is crude distillation capacity and condensate splitters. The third column

---

[14] It is possible to string together a slightly longer, more complete time series using older versions of Eni publications. However, the data appears to have been revised several times, most recently in 2015. As a result, while the numbers in Table 5.1 are comparable to each other, the longer time series one can put together using older reports are not comparable, strictly speaking. We report the longer time series in the appendix.

Figure 5.1: Residual fuel oil consumption declining over time



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

shows data on conversion capacity, which measures how much cracking and coking capacity is available.[15] The fourth 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.[16] Unlike the conversion capacity data, the NCI reflects not only the amount of upgrading capacity available but also the amount of desulfurisation capacity available.

Since the year 2000, primary capacity has been growing at about 1 percent per year, on average, or 0.9 million barrels per day (mb/d). Conversion capacity has been growing 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 values of the NCI. 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 have occurred in Asia, where the conversion capacity ratio has risen from 36

---

[15] Conversion capacity is fluid catalytic cracking equivalent. Details on the calculation can be found in Eni's World Oil Review 2018.

[16] Johnston (1996) provides a good introduction to the index and how it is calculated.

Table 5.1: Global refineries increasingly complex

Year	Primary capacity (mb/d)	Conversion capacity (mb/d)	Conversion capacity ratio (percent)	Complexity Ratio 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

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.

percent in 2000 to 66 percent in 2017, with the NCI rising from 7.0 to 9.7.

Publicly available data on refining capacity is also available from several other sources but not reported in the table. The first is the British Petroleum (BP) Statistical Review of World Energy, which provides annual data on primary refining capacity as well as crude throughput (the amount of crude processed by refiners), but not data on conversion capacity. For those years where there is data from both Eni and BP, we find the differences between the two primary capacity series are relatively modest, usually less than one percent, so none of our conclusions are sensitive to using one series or the other. Some data on world distillation capacity are also available from International Energy Agency (IEA) Medium-Term Oil Market Reports and Market Report Series. This data are generally close to both Eni and BP's data on primary capacity.

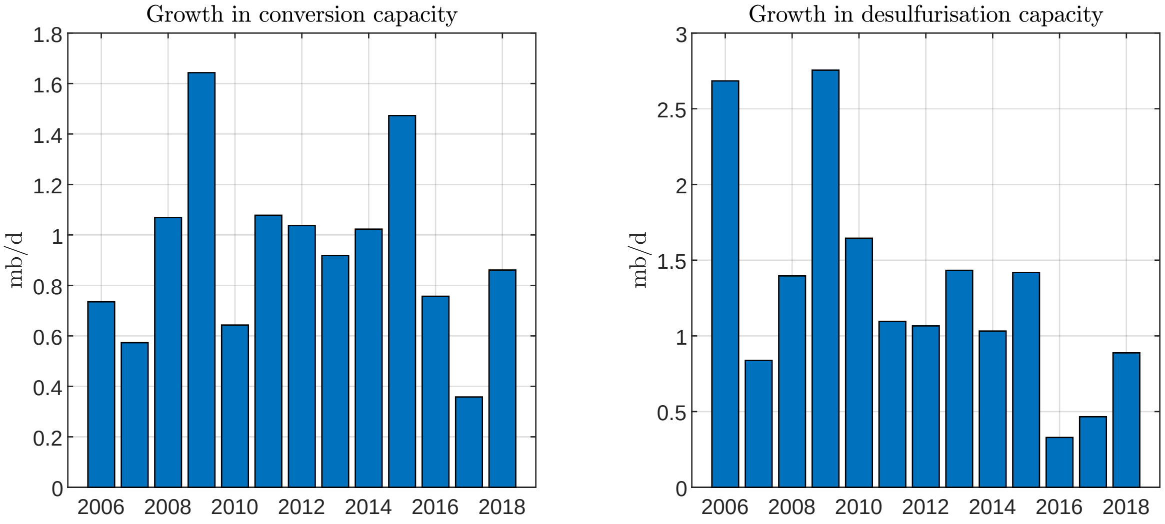
The IEA reports also provide data on additions to global conversion capacity and desulfurisation capacity from 2006 to 2017, as well as a forecast for 2018. We plot both of those series in Figure 5.2. This data reinforces the findings of the Eni data, as it shows significant additions to conversion capacity. We also note here that there were some fairly large capacity additions in 2008 and 2009, a point we will come back to when we discuss the timing of the breaks.

Another source is the Oil&Gas Journal Worldwide Refinery Survey, which provides both primary capacity and conversion capacity numbers. Ideally, we would prefer to use this data as it is available at an annual frequency and begins earlier than other data series we have available. We do not, however, because participation in the survey is voluntary and it appears that the survey is not accurately measuring capacity in some important developing countries, particularly China.[17]

---

[17] The discrepancies appear to be large enough to be important for our discussion. The survey reported that at the start of 2011 that China's crude distillation capacity was little under 7 mb/d. Wu (2011), however,

Figure 5.2: 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 which are publicly available.

Unexpected growth in light oil production: At the same time the global refining sector has increased its ability to transform low-quality crude oil into gasoline and diesel, the production of light, sweet crude has unexpectedly increased since the late 2000s due in large part to the U.S. shale boom. In Table 5.2, we show data from Eni on global production of ultra light and light crude oil. Ultra light crude is defined as any crude oil with an API gravity of 50 or higher while light crude has an API gravity from 35 up to 50. Since 2010, the production of ultra light is up almost 1 million barrels per day, while light is up about 2 million barrels per day. These increases are important because, as shown in Figure 2.1, lighter oil naturally produces less residual fuel oil than medium and heavy crude oils. In the case of ultra light oil, the amount of residual produced can be extremely low, even when run through a simple refinery with no conversion capacity. All else equal, this increase in supply, therefore, would reduce the spread between light crude and other crude, while also reducing spreads between residual fuel oil and other, higher-valued petroleum products.

Although it is taken for granted that the shale boom was an unexpected event, for our purposes it is worth pointing out how surprising the production increases have been in hindsight. The Energy Information Administration, in its 2010 Annual Energy Outlook, forecasted that U.S. crude production would rise from 5.3 mb/d to just 5.8 mb/d by 2017.

---

brings additional data to bear and reports distillation capacity over 10 mb/d. Likewise, the survey reports distillation capacity of 7 mb/d and coking capacity of 156,000 b/d at the start of 2013 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).

The International Energy Agency, in its Medium-term Oil and Gas Markets 2010 report, predicted that U.S. crude production would have declined by almost 0.4 mb/d from 2010 to 2015. While it is difficult to construct a true counterfactual, we believe it very unlikely that there would have been a commensurate increase in light oil production somewhere outside the U.S. had the shale boom not happened.

Table 5.2: Light crude production up since 2005

Year	Ultra light	Light	Other
2000	1.46	20.71	46.49
2005	1.94	19.81	53.10
2010	2.46	20.43	52.11
2015	3.43	21.98	56.35
2016	3.42	21.56	56.73
2017	3.40	22.48	55.82

Notes: Units are millions of barrels per day. Eni defines ultra light as crude oil with API gravity of 50 or above while light crude oil has an API gravity from 35 up to but not including 50. Sources: Eni World Oil Review 2018, Eni World Oil Review 2017, Eni World Oil and Gas Review 2016.

5.2 The timing of the breaks and the Great Recession

The previous discussion focused on longer-term trends that can explain why quality-related differentials have fallen over time but was silent on why there was a cluster of breaks around 2008. In fact, the timing of those breaks is a bit of a puzzle when considering some of those longer-term trends. For example, neither the shale boom nor environmental regulations could have played a major role in 2008. Due to the timing, it is natural to wonder about a potential connection with the Great Recession. In this section, we introduce some evidence for such a connection, making use of some previously introduced data as well as some additional refinery data.

Based on the data we have available, utilization rates for refining capacity, both primary and conversion, were high in the years preceding the Great Recession. Some conversion capacity data even suggested capacity may have even been constrained. The Recession played a role in the timing of the breaks because it significantly reduced demand for petroleum products, particularly for products besides residual fuel oil. This can be seen in Figure 5.1. At the same time, as we have shown in Figure 5.2, global conversion capacity experienced very large increases in 2008 and 2009. As a result of both factors, utilization rates of both primary and conversion capacity fell sharply during the Recession. Since then, utilization rates generally have remained lower than before, suggesting the additional capacity that has come online after the Recession has been sufficient to process new supplies of medium and heavy crude and meet the growing demand for gasoline and diesel without significantly increasing the production of residual fuel oil. As a result, the Great Recession played a role in the timing of the breaks by allowing conversion capacity to catch up with demand, as it were.

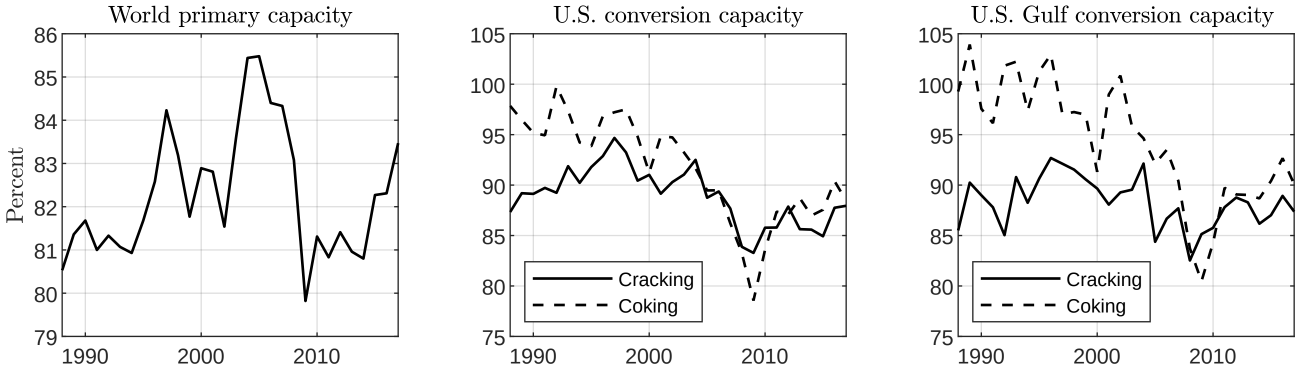
Turning to the details of our data, we construct a utilization rate for world primary capacity using data from the British Petroleum (BP) Statistical Review of World Energy on crude throughput and primary capacity. Unfortunately, no publicly available data exist that allow us to construct a world utilization rate for either cracking or coking capacity. However, data is available from the U.S. Energy Information Administration for the U.S. refining sector from 1987 to 2017. As the U.S. refining sector is the largest in the world, composed of profit-maximizing firms, and fully integrated with the global fuel market, we believe it should be reflective to some extent of conditions elsewhere. We use data on fresh feed input to cokers along with annual data on coking capacity to construct the utilization rate for coking capacity. Likewise, we use equivalent data for catalytic crackers and hydrocrackers to construct a utilization rate for cracking capacity. We do so for both the U.S. as a whole as well as the Gulf Coast. We consider the Gulf Coast not only because we have price differential data for the area, but also because it is home to a substantial portion of U.S. conversion capacity and has seen the largest increase in coking capacity over the sample period. The data appendix provides full details on the data and the calculations.

Figure 5.3 plots the utilization rates. The BP data reflect the boom in demand that occurred before the Great Recession, as well as the effects of the Recession itself. We note the utilization rates for both coking and cracking capacity exhibit what appear to be structural breaks right around the time of the Great Recession. Before, the utilization rates were relatively high and, for coking capacity, the data is suggestive of some potential capacity constraints early in the sample. Utilization rates for coking capacity in the U.S. Gulf Coast were even higher, with rates exceeding 100 percent several years in the 1990s and in 2002.[18] During the recession, however, utilization rates for conversion capacity declined sharply and since then they have remained at levels below those seen from 1987 to the mid-2000s. We show in the appendix that inputs into U.S. crackers and cokers dropped sharply during the Great Recession while there were some modest additions to coking capacity. Outside the U.S., we have shown in Figure 5.2 that there were some extremely large capacity additions to global conversion capacity in 2009, which would also could have weighed on utilization

---

[18] The EIA capacity data is designed to take into account downtime at units and, as a result, allows for the possibility of utilization rates above 100 percent. However, episodes such as that suggest the machinery is being pushed to its physical limits.

Figure 5.3: Utilization rates for refining capacity



Notes: World primary capacity utilization is calculated as crude throughput divided by world primary capacity, which is a nameplate capacity figure. Cracking and coking utilization is calculated as fresh feed input divided by the Energy Information Administration's measure of calendar day capacity.

rates. For both the U.S. and the rest of the world, further conversion capacity additions have been made in the post-Recession period.

Our interpretation of all this data is that the Great Recession, in essence, allowed capacity additions to catch up with demand. While the demand for gasoline and diesel has been growing due to an expanding world economy since the Recession, conversion capacity additions appear to have been sufficiently large to meet that incremental demand growth and keep both quality-related crude oil differentials and residual-fuel oil differentials at relatively low levels.

6 Additional results

6.1 Unit root testing

Previous works in the literature have tested for the stationarity of oil price differentials using unit root tests, for example Fattouh (2010), Giulietti et al. (2015) and Agerton and Upton (Forthcoming). As discussed in Perron (1989) and many papers since then, the structural breaks identified so far can influence the results one gets from such tests. We investigated the importance of this issue using both the daily and monthly data. To conserve space, we relegate the full set of tables with results to the appendix and summarize our main findings here.

Our procedure is straight forward. We run an Augmented Dickey Fuller (ADF) test on each differential, considering cases where the optimal lag length is chosen using the Akaike Information Criterion (AIC) or Schwarz Information Criterion (SIC) tests. We report results for both as we have found that the results can be sensitive to how the lag length is chosen. A differential is flagged by us any time one of the tests fails to reject the null of a unit root at the 1 percent significance level. For those flagged cases, we then perform an ADF breakpoint unit root test which searches for the break that minimizes the intercept break t-statistic, trimming 15 percent of the sample.

To summarize, using daily data we find that some variant of the test fails to reject the null of a unit root at a 1 percent significance level for only 4 out of 38 differentials. Of these, three involve Urals crude, for which we only have a more limited data set. In all cases, the unit root breakpoint test that uses the SIC to determine the lag length overwhelmingly rejects the null of a unit root. The problem is more widespread with monthly data, where we have fewer observations and the time aggregation involved with going from daily observations to monthly averages has the potential to make the differentials appear more persistent (Taylor (2001)). With the monthly data, we flag 34 out of 59 differentials. As with the daily data, the null of a unit root is rejected in all cases when using the SIC test to determine the lag length and allowing for a single break.

Overall, these results show that some caution should be applied when using standard unit root tests to (log) oil price differentials. Many of these differentials have experienced one or more structural breaks in their means, and these breaks can generate misleading results from the unit root tests if not taken into account by the modeler, particularly for monthly data.

7 Conclusion

Crude oil can vary significantly in some key physical properties, making them 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 quality-related differentials experienced a major break in and around the time of the Great Recession.

Our analysis points to several explanations for why quality-related differentials have narrowed over time. One is the fact that the global refining sector has become increasingly complex over time, as capacity additions have increased the ability of the sector to transform lower-grade crude oil into high-valued petroleum products. At the same time, the shale boom has unexpectedly increased the production of light crude oil, reducing, on the margin, the need for complex refineries. We also considered, but ruled out, the possibility that a relaxation of environmental regulations on sulfur emissions or a shift towards greater use of residual fuel oil could explain our findings. In fact, both of those forces should be contributing to wider differentials, as environmental regulations have been tightened and use of residual fuel oil has been falling out of favor in many parts of the world.

We find a connection with the Great Recession and the timing of the breaks due to the Great Recession's substantial impact on demand for petroleum products and the fact that there were significant conversion capacity additions during the Recession. The reduction in demand appears to have allowed capacity additions to catch up with demand, which had been strong for several years prior to the Recession. The available data on utilization rates for conversion capacity shows that since then those rates have remained below their pre-Recession highs, suggesting recent conversion capacity additions have been sufficient to process supplies of medium and heavy crude oil and meet growing demand for gasoline and diesel while minimizing the production of residual fuel oil.

A number of possible avenues suggest themselves 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. One could also consider setting up a theoretical model of the global refining sector to explore how theory suggests that changes in that sector should affect oil price differentials. We leave these for future research.

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 (Forthcoming): "Decomposing Crude Price Differentials: Domestic Shipping Constraints or the Crude Oil Export Ban?" *The Energy Journal*, https://doi.org/10.2139/ssrn.2942989.

Andrews, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858, https://doi.org/10.2307/2938229.

Bacon, R. and S. Tordo (2005): "Crude oil price differentials and differences in oil qualities: A statistical analysis," ESMAP Technical Paper.

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.

——— (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., T. K. Lee, J. T. Moser, and M. A. Robe (2013): "Physical Markets, Paper Market and the WTI-Brent Spread," *The Energy Journal*, 34, 129–151, https://doi.org/10.5547/01956574.34.3.7.

Energy Intelligence Research (2009): *The International Crude Oil Market Handbook 2009*, Energy Intelligence Research.

Evans, B. and J. Mowler (2002): "Switch to sour crude haunts US 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.

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.

——— (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.

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.

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

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

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

Perron, P. et al. (2006): "Dealing with structural breaks," *Palgrave Handbook of Econometrics*, 1, 278–352.

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

PIW (2005): "US refining trends ride on shifting spreads," *Petroleum Intelligence Weekly*.

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

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.

Taylor, A. M. (2001): "Potential Pitfalls for the Purchasing-Power-Parity Puzzle? Sampling and Specification Biases in Mean-Reversion Tests of the Law of One Price," *Econometrica*, 69, 473–498, https://doi.org/10.3386/w7577.

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.

Wu, K. (2011): "Capacity, complexity expansions characterize China's refining industry past, present, future," *Oil&Gas Journal*.

A Appendix (not for publication)

A.1 Oil price data

Algerian Saharan - HAVER Mnemonic Q830AGS@OGJ. API gravity of 44.0 and sulfur content of 0.1, as reported in OPEC Annual Statistical Bulletins. HAVER reports the original source as OPEC.

Bonny Light - HAVER Mnemonic Q830NGBL@OGJ. API gravity of 36.7 and sulfur content of 0.1, as reported in OPEC Annual Statistical Bulletins.

Brent - Bloomberg Ticker EUCRBRDT. Bloomberg states the API gravity is greater than 35 while sulfur content is less than 1 percent. Table 3.1 reports API gravity and sulfur content from Platts. This price is FOB at Sullom Voe terminal in the Shetland Islands, UK.

Dubai - Bloomberg Ticker PGCRDUBA. API gravity of 31, sulfur content 1.7. This is FOB at Dubai, UAE.

Heavy Louisiana Sweet - Bloomberg Ticker USCRHLSE. API gravity of 33.7, sulfur content 0.39. This is a spot price at Empire, LA and is FOB.

Louisiana Light Sweet - Bloomberg Ticker USCRLLS. API gravity of 35.7, sulfur content 0.44. This is a spot price at St. James, LA and is FOB.

Mars - Bloomberg Ticker USCRHLSE. API gravity of 33.7, sulfur content 0.39. This is a spot price at Empire, LA and is FOB.

Maya - Bloomberg Ticker LACRMAUS. API gravity of 21.1, sulfur content 3.38. Price is derived from the formula for Maya sales to the U.S. Gulf Coast: [0.4*(WTS+3.5% Fuel Oil)] + [0.1*(LLS + Dated Brent)]. Price is FOB.

Oman - Bloomberg Ticker PGCROMAN. API gravity of 33, sulfur content 1.11. This is FOB at Muscat, Oman.

Saudi Heavy - Bloomberg Tickers PGCRAHUS, PGCRAHEU, and PGCRARHV for US, Europe and Asia, respectively. API gravity of 27, sulfur content 2.8. Asia price is FOB at Yanbu or Ras Tanura.

Tapis-Bloomberg Ticker APCRTAPI. API gravity of 44.6 and sulfur content of 0.028. Loading port is reported as Kertih, Malaysia.

Urals-HAVER Mnemonic is P922URL@INTDAILY for daily and Q830URU@OGJ for monthly. API gravity of 31.7, sulfur content 1.35. HAVER reports the original source for monthly data as OPEC.

WTI Cushing-Bloomberg Ticker USCRWTIC. API gravity of 39, sulfur content 0.34. This is a spot price at Cushing, OK and is FOB.

WTI Midland-Bloomberg Ticker USCRWTIM. API gravity of 39, sulfur content 0.34. This is a spot price at Midland, TX and is FOB.

West Texas Sour-Bloomberg Ticker USCRWTSM API gravity 34, sulfur content 1.9. This is a spot price at Midland, TX and is FOB.

A.2 Unit root tests using daily data

In this section of the appendix we present a full set of results for the unit root tests that use daily data. Results for the unit root tests using monthly data are presented in the next section.

Previous works in the literature have tested for the stationarity of oil price differentials using unit root tests, for example Fattouh (2010), Giulietti et al. (2015) and Agerton and Upton (Forthcoming). As discussed in Perron (1989) and many papers since then, the structural breaks identified so far can influence the results one gets from such tests. *A priori*, our expectation is that the differentials should be stationary once we have taken into account structural breaks. Our reasoning behind this expectation is similar to the logic of the previous literature (and the law of one price literature): arbitrage across locations and across types of oil should prevent a differential from becoming exceptionally large in either direction for significant periods of time. However, changes in the nature of that arbitrage could generate breaks of the type we have documented, and not taking them into account could lead to the appearance of non-stationarity.

Our procedure to test for unit roots is straight forward. We run an Augmented Dickey Fuller (ADF) test on each differential, considering cases where the optimal lag length is chosen using the AIC or SIC. We report results for both as we have found that the results can be sensitive to how the lag length is chosen. A differential is flagged by us any time one of the tests fails to reject the null of a unit root at the 1 percent significance level. For those cases flagged, we then perform an ADF breakpoint unit root test which searches for the break that minimizes the intercept break t-statistic, trimming 15 percent of the sample.

The results are reported in Tables A.1 and A.2. As discussed in the main text, we only find a handful of cases where one of the unit root tests is unable to reject the null of a unit root. This occurs for 4 out of 38 differentials. The null of a unit root is rejected in all cases when using a breakpoint unit root test and the SIC test to pick the lag length.

Table A.1: Unit root test results for daily data

	AIC		SIC	
Differential	ADF	ADF (BP)	ADF	ADF (BP)
WTIM-WTS	-4.20	-7.63	-4.48	-9.80
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
LLS-HLS	-7.10	-8.42	-9.34	-10.87
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
LLS-Mars	-3.84	-6.22	-4.66	-8.14
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
LLS-Maya	-3.49	-5.43	-3.51	-5.66
	(<0.01)	(<0.01)	(<0.01)	<(0.01)
HLS-Mars	-4.76	-6.76	-4.94	-7.82
	(<0.01)	(<0.01)	(<0.01)	<(0.01)
HLS-Maya	-3.68	-6.23	-4.14	-6.23
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Mars-Maya	-5.47	-7.71	-6.07	-8.83
	(<0.01)	(<0.01)	(<0.01)	<0.01
Brent-Urals	**-2.96**	-6.52	-4.11	-7.29
	(.04)	(<0.01)	(<0.01)	(<0.01)
Tapis-Oman	-6.60	-7.96	-7.06	-8.45
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Tapis-Dubai	-6.24	-7.94	-6.83	-8.47
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Oman-Dubai	-4.39	-4.95	-4.81	-6.83
	(<0.01)	(<0.01)	(<0.01)	(<0.01)

Notes: For each differential, the first row shows the test statistics for the Augmented Dickey-Fuller (ADF) and the ADF breakpoint (ADF BP) tests. The second row shows the p-value for the test. Bold text identifies a case where the null of a unit root would not be rejected at a one percent significance level.

Table A.2: Unit root test results for daily data

	AIC		SIC	
Differential	ADF	ADF (BP)	ADF	ADF (BP)
Tapis-Urals	-6.04	-7.39	-6.09	-7.97
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Tapis-Mars	-7.97	-8.87	-7.97	-11.42
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Brent-Oman	-6.64	-7.28	-9.04	-9.75
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Brent-Dubai	-5.65	-6.58	-7.92	-9.31
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Brent-Mars	-6.91	-7.12	-11.25	-12.05
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
LLS-Oman	-4.36	-6.55	-5.42	-6.25
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
LLS-Urals	**-1.90**	**-3.61**	**-3.02**	-8.36
	(0.33)	**(0.05)**	**(0.03)**	(<0.01)
LLS-Dubai	-4.02	-6.25	-4.93	-6.23
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
HLS-Oman	-5.39	-6.89	-6.35	-10.85
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
HLS-Urals	**-2.52**	**-4.14**	**-2.71**	-9.44
	(0.11)	**(0.01)**	**(0.07)**	(<0.01)
HLS-Dubai	-5.21	-6.67	-5.80	-10.78
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Tapis-Maya	-4.46	-6.20	-5.46	-7.12
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Brent-Maya	-4.31	-6.61	-5.14	-6.98
	(<0.01)	(<0.01)	(<0.01)	(<0.01)

	AIC		SIC	
Differential	ADF	ADF (BP)	ADF	ADF (BP)
Oman-Maya	-6.01	-7.22	-7.08	-8.43
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Urals-Maya	-4.92	-5.78	-6.03	-6.63
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Dubai-Maya	-5.85	-6.93	-6.88	-8.07
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
WTIC-LLS	-3.55	-5.20	-3.91	-6.25
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
WTIM-LLS	**-3.31**	-4.83	**-3.30**	-5.57
	(**0.01**)	(<0.01)	(**0.02**)	(<0.01)
LLS-Tapis	-7.72	-9.94	-9.54	-12.56
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
HLS-Tapis	-8.80	-11.01	-12.06	-13.66
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
LLS-Brent	-4.14	-7.62	-4.95	-13.40
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
HLS-Brent	-5.20	-8.52	-6.17	-14.45
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Oman-Urals	-5.59	-9.43	-8.44	-10.93
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Oman-Mars	-8.56	-10.51	-12.60	-13.04
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Urals-Dubai	-6.02	-8.90	-8.85	-9.59
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Urals-Mars	-6.46	-6.52	-6.46	-10.57
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Dubai-Mars	-7.52	-9.15	-9.18	-12.91
	(<0.01)	(<0.01)	(<0.01)	(<0.01)

Notes: For each differential, the first row shows the test statistics for the Augmented Dickey-Fuller (ADF) and the ADF breakpoint (ADF BP) tests. The second row shows the p-value for the test. Bold text identifies a case where the null of a unit root would not be rejected at a one percent significance level.

A.3 Additional results using monthly data

In this section we present results based on monthly data which includes a larger set of crude oil prices and a longer sample for Urals crude. Table A.3 lists all of the crude oils considered plus their properties. The additional crude oils are Algerian Saharan, Bonny Light, and Saudi Arabian Heavy. For Saudi Arabian Heavy, there are specific prices for the US, Europe and Asia and we only consider within-area differentials using the relevant Saudi Heavy price. Summary statistics for the within-area and across-area differentials are presented in Tables A.4 and A.5, respectively.

We report breakpoint results for crude oils of different quality in Tables A.6 and A.7. For the breakpoint test we set the trimming parameter to 0.15, which sets the minimum regime length at roughly 3 years. The procedure used to estimate the long-run variance-covariance matrix is the same as with the daily data. We report results for a statistical significance of 5 percent. We have chosen to be somewhat less restrictive with the monthly data as we have significantly few observations but the results are not overly sensitive to choosing a more stringent level of significance. The evolution of the means can be found in Tables A.8, A.9, and A.10. Summary statistics pre and post-break can be found in A.11. Unit root test results are shown in Tables A.12 and A.13.

Table A.3: 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
Saudi Heavy to US (SHU)	27.0	2.80	Medium	Sour
Europe/Atlantic Basin				
Algerian Saharan (Saharan)	44.0	0.10	Light	Sweet
Bonny Light (Bonny)	36.7	0.10	Light	Sweet
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

Table A.4: Oil price differentials within areas

Differential	API difference	Sulfur difference	Mean	Standard deviation
Midland, TX				
WTIM-WTS	5.0	-1.56	0.046	0.040
U.S. Gulf Coast				
LLS-HLS	2.0	0.05	0.015	0.014
LLS-Mars	6.8	-1.61	0.110	0.059
LLS-SHU	8.7	-2.36	0.195	0.115
LLS-Maya	14.6	-2.94	0.230	0.107
HLS-Mars	4.8	-1.66	0.094	0.054
HLS-SHU	6.7	-2.41	0.180	0.109
HLS-Maya	12.6	-2.99	0.213	0.100
Mars-SHU	1.9	-0.75	0.086	0.069
Mars-Maya	7.8	-1.33	0.120	0.060
SHU-Maya	5.9	-0.58	0.032	0.056
Europe / Atlantic Basin				
Saharan-Brent	5.9	-0.31	0.009	0.015
Saharan-Bonny	7.4	0.00	-0.004	0.016
Saharan-Urals	12.5	-1.34	0.050	0.037
Saharan-SHE	17.0	-2.70	0.147	0.091
Brent-Bonny	1.4	0.31	-0.013	0.016
Brent-Urals	6.6	-1.03	0.040	0.035
Brent-SHE	11.1	-2.39	0.138	0.087
Bonny-Urals	5.2	-1.34	0.053	0.036
Bonny-SHE	9.7	-2.70	0.151	0.083
Urals-SHE	4.5	-1.36	0.098	0.070
Middle East / Asia				
Tapis-Oman	11.6	-1.07	0.093	0.049
Tapis-Dubai	13.6	-1.67	0.103	0.051
Tapis-SHA	17.6	-2.77	0.157	0.086
Oman-Dubai	2.0	-0.60	0.010	0.018
Oman-SHA	6.0	-1.70	0.063	0.056
Dubai-SHA	4.0	-1.10	0.053	0.053

Notes: These statistics are based on a sample from January 1997 to December 2018.

Table A.5: Oil price differentials across areas

Crudes of different type

Differential	API difference	Sulfur difference	Mean	Standard deviation
Light-medium differentials				
Tapis-Urals	13.1	-1.41	0.093	0.053
Tapis-Mars	15.7	-2.02	0.126	0.053
Saharan-Oman	11.0	-1.00	0.049	0.043
Saharan-Dubai	13.0	-1.60	0.059	0.045
Saharan-Mars	15.1	-1.95	0.081	0.046
Brent-Oman	5.1	-0.69	0.040	0.038
Brent-Dubai	7.1	-1.29	0.050	0.041
Brent-Mars	9.2	-1.64	0.072	0.041
Bonny-Oman	3.7	-1.00	0.053	0.039
Bonny-Dubai	5.7	-1.60	0.063	0.041
Bonny-Mars	7.8	-1.95	0.085	0.041
LLS-Oman	2.7	-0.66	0.077	0.062
LLS-Urals	4.2	-1.0	0.077	0.066
LLS-Dubai	4.7	-1.26	0.087	0.065
HLS-Oman	0.7	-0.71	0.062	0.057
HLS-Urals	2.2	-1.05	0.062	0.059
HLS-Dubai	2.7	-1.31	0.072	0.060
Light-heavy differentials				
Tapis-Maya	23.5	-3.35	0.244	0.094
Saharan-Maya	22.9	-3.28	0.199	0.089
Brent-Maya	17.0	-2.97	0.191	0.084
Bonny-Maya	15.6	-3.28	0.204	0.082
Medium-heavy differentials				
Oman-Maya	11.9	-2.28	0.151	0.070
Urals-Maya	10.4	-1.94	0.151	0.070
Dubai-Maya	9.9	-1.68	0.141	0.072
Crudes of similar type				
Light-light differentials				
WTIC-LLS	3.3	-0.10	-0.040	0.058
WTIM-LLS	3.3	-0.10	-0.057	0.074
LLS-Tapis	-8.9	0.41	-0.016	0.043
LLS-Saharan	-8.3	0.34	0.028	0.040
LLS-Brent	-2.4	0.03	0.037	0.041
LLS-Bonny	-1.0	0.34	0.024	0.047
HLS-Tapis	-10.9	0.36	-0.031	0.040
HLS-Saharan	-10.3	0.29	0.014	0.036
HLS-Brent	-4.4	-0.02	0.022	0.037
HLS-Bonny	-3.0	0.29	0.009	0.042
Medium-medium differentials				
Oman-Urals	1.5	-0.34	0.000	0.034
Oman-Mars	4.1	-0.95	0.032	0.039
Urals-Dubai	0.5	-0.26	0.010	0.036
Urals-Mars	2.6	-0.61	0.032	0.044
Dubai-Mars	2.1	-0.35	0.022	0.044

Notes: These statistics are based on a sample from January 1997 to December 2018.

Table A.6: Breakpoint test results for crudes of different qualities

Part 1: Within-area differentials

				F-statistic		
Differential	Break 1	Break 2	Break 3	0 vs. 1	1 vs. 2	2 vs. 3
Midland, TX						
WTIM-WTS	11/2007	02/2013	-	115.19	10.22	-
U.S. Gulf Coast						
LLS-Mars	01/2009	12/2001	-	43.47	23.30	-
LLS-SHU	05/2006	09/2009	02/2002	32.53	17.98	15.45
LLS-Maya	05/2007	-	-	61.36	-	-
HLS-Mars	04/2008	12/2001	-	37.87	15.62	-
HLS-SHU	05/2005	02/2009	02/2002	31.53	31.28	14.01
HLS-Maya	05/2007	-	-	63.43	-	-
Mars-Maya	04/2007	-	-	41.97	-	-
SHU-Maya	09/2000	-	-	9.11	-	-
Europe/Atlantic Basin						
Saharan-Urals	06/2008	-	-	46.03	-	-
Saharan-SHE	06/2007	-	-	14.93	-	-
Brent-Urals	06/2008	-	-	31.96	-	-
Brent-SHE	02/2007	-	-	16.43	-	-
Bonny-Urals	10/2007	01/2004	-	33.93	13.27	-
Bonny-SHE	07/2007	-	-	14.97	-	-
Middle East/Asia						
Tapis-Oman	04/2008	-	-	14.01	-	-
Tapis-Dubai	04/2008	-	-	21.03	-	-
Tapis-SHA	02/2009	-	-	10.07	-	-

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	05/2008	07/2012	-	30.10	11.68	-
Tapis-Mars	01/2008	04/2011	-	15.37	10.93	-
Saharan-Oman	04/2008	-	-	15.67	-	-
Saharan-Dubai	04/2008	-	-	21.93	-	-
Saharan-Mars	01/2002	-	-	11.72	-	-
Brent-Oman	04/2008	-	-	11.10	-	-
Brent-Dubai	04/2008	-	-	12.95	-	-
Brent-Mars[#]	01/2008	08/2013	-	7.02	32.04	-
Bonny-Oman[#]	06/2004	04/2008	-	7.18	15.96	-
Bonny-Dubai	06/2008	-	-	10.67	-	-
Bonny-Mars[#]	04/2008	07/2013	-	5.88	19.94	-
LLS-Oman	11/2008	-	-	47.92	-	-
LLS-Urals	05/2009	-	-	51.09	-	-
LLS-Dubai	12/2008	04/2005	-	45.17	13.59	-
HLS-Oman	10/2008	-	-	48.43	-	-
HLS-Urals	03/2007	04/2012	-	57.55	16.50	-
HLS-Dubai	10/2008	03/2005	-	48.57	16.74	-
Light-heavy						
Tapis-Maya	05/2007	-	-	34.67	-	-
Saharan-Maya	07/2007	-	-	29.95	-	-
Bonny-Maya	07/2007	-	-	26.37	-	-
Brent-Maya	05/2007	-	-	31.47	-	-
Medium-heavy						
Oman-Maya	05/2007	-	-	23.18	-	-
Dubai-Maya	03/2002	-	-	13.30	-	-
Urals-Maya	02/2002	-	-	14.53	-	-

Notes: Dates refer to the last month of a given regime. The order of the breaks is determined by the test. The critical values are 8.58, 10.13 and 11.14 for tests of 0 or 1 break, 1 or 2 breaks, and 2 or 3 breaks, respectively. These reflect a significance level of 5 percent. [#]: The test rejects the null of 1 break vs. 2 but fails to reject the null of 0 vs. 1.

Table A.7: Breakpoint test results for crudes of similar type

Part 1: Within-area differentials

					F-statistic			
Differential	Break 1	Break 2	Break 3	Break 4	0 vs. 1	1 vs. 2	2 vs. 3	3 vs. 4
U.S. Gulf Coast								
LLS-HLS	01/2011	06/2014	-	-	11.80	14.65	-	-
Mars-SHU	02/2002	12/2009	-	-	25.76	31.48	-	-
Europe/Atlantic Basin								
Saharan-Brent	05/2000	-	-	-	27.93	-	-	-
Saharan-Bonny	12/2001	-	-	-	64.88	-	-	-
Brent-Bonny	03/2005	07/2014	-	-	18.77	32.24	-	-
Urals-SHE	01/2001	-	-	-	13.55	-	-	-
Middle East / Asia								
Oman-Dubai	-	-	-	-	-	-	-	-
Oman-SHA	-	-	-	-	-	-	-	-
Dubai-SHA	-	-	-	-	-	-	-	-

Part 2: Across-area differentials

					F-statistic			
Differential	Break 1	Break 2	Break 3	Break 4	0 vs. 1	1 vs. 2	2 vs. 3	3 vs. 4
Light-light								
WTIC-LLS	04/2010	11/2006	04/2001	07/2013	10.92	100.91	22.73	14.05
WTIM-LLS	01/2011	10/2006	04/2014	04/2001	14.58	243.76	30.84	12.53
LLS-Tapis	01/2005	-	-	-	35.71	-	-	-
LLS-Saharan	12/2010	02/2005	-	-	45.90	18.70	-	-
LLS-Brent	05/2011	12/2004	-	-	50.11	22.28	-	-
LLS-Bonny	02/2005	03/2011	-	-	56.57	16.17	-	-
HLS-Tapis	04/2004	-	-	-	38.41	-	-	-
HLS-Saharan	12/2004	07/2013	-	-	52.33	19.68	-	-
HLS-Brent	12/2004	07/2013	-	-	45.47	33.25	-	-
HLS-Bonny	01/2005	-	-	-	87.49	-	-	-
Medium-medium								
Oman-Mars$^{\#}$	01/2002	08/2013	-	-	8.58	26.69	-	-
Urals-Mars	07/2013	01/2002	10/2005	-	- 17.50	11.38	11.67	-
Dubai-Mars	08/2013	-	-	-	11.92	-	-	-
Oman-Urals	06/2010	-	-	-	10.97	-	-	-
Urals-Dubai	-	-	-	-	-	-	-	-

Notes: Dates refer to the last month of a given regime. The order of the breaks is determined by the test. The critical values are 8.58, 10.13, 11.14 and 11.83 for tests of 0 or 1 break, 1 or 2 breaks, 2 or 3 breaks, and 3 or 4 breaks, respectively. These reflect a significance level of 5 percent. $^{\#}$: The test rejects the null of 1 break vs. 2 but fails to reject the null of 0 vs. 1.

Table A.8: Regression constant across regimes for crudes of different types

Within-area differentials

Differential	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Change
WTIM-WTS	0.080											0.025					0.003*						-0.077
LLS-Mars	0.179					0.126							0.062										-0.117
HLS-Mars	0.161					0.108						0.055											-0.106
LLS-Maya	0.312										0.151												-0.161
HLS-Maya	0.292										0.142												-0.150
LLS-SHU	0.343					0.240				0.167				0.102									-0.241
HLS-SHU	0.325					0.223			0.168				0.095										-0.230
SHU-Maya	-0.018*				0.043																		0.061
Mars-Maya	0.158										0.083												-0.075
Saharan-Urals	0.073											0.023											-0.050
Saharan-SHE	0.208										0.091												-0.117
Brent-Urals	0.061												0.018										-0.043
Brent-SHE	0.198										0.087												-0.111
Bonny-Urals	0.055							0.103				0.035											-0.020
Bonny-SHE	0.204											0.101											-0.103
Tapis-Oman	0.117											0.069											-0.048
Tapis-Dubai	0.131											0.074											-0.057
Tapis-SHA	0.195												0.109										-0.086

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.

Table A.9: Regression constant across regimes for crudes of different types

Across-area: Light-medium

Differential	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Change
Tapis-Urals	0.122											0.074					0.055						-0.067
Tapis-Mars	0.150											0.080		0.110									-0.040
Saharan-Oman	0.067											0.029											-0.038
Saharan-Dubai	0.082											0.035											-0.047
Saharan-Mars	0.121			0.069																			-0.052
Brent-Oman	0.055											0.024											-0.031
Brent-Dubai	0.069											0.030											-0.039
Brent-Mars	0.088											0.033						0.080					-0.008
Bonny-Oman	0.048								0.099				0.040										-0.008
Bonny-Dubai	0.080												0.045										-0.035
Bonny-Mars	0.098											0.054						0.089					-0.009
LLS-Oman	0.117												0.030										-0.087
LLS-Urals	0.121												0.020										-0.101
LLS-Dubai	0.143								0.100				0.035										-0.108
HLS-Oman	0.098												0.021										-0.077
HLS-Urals	0.108										0.047					0.003*							-0.105
HLS-Dubai	0.124								0.083				0.027										-0.097

Across-area: Light-heavy

Differential	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Change
Tapis-Maya	0.308												0.189										-0.119
Saharan-Maya	0.258											0.146											-0.112
Brent-Maya	0.246											0.141											-0.105
Bonny-Maya	0.255											0.156											-0.099

Across-area: Medium-heavy

Differential	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Change
Oman-Maya	0.192										0.113												-0.079
Dubai-Maya	0.209			0.119																			-0.089
Urals-Maya	0.219			0.130																			-0.089

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.

Table A.10: Regression constant across regimes for crudes of the same type

Light-light differentials

<table>
<tr><th>Differential</th><th>97</th><th>98</th><th>99</th><th>00</th><th>01</th><th>02</th><th>03</th><th>04</th><th>05</th><th>06</th><th>07</th><th>08</th><th>09</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>Change</th></tr>
<tr><td>WTIC-LLS</td><td colspan="4">0.008*</td><td colspan="6">-0.005</td><td colspan="3">-0.038</td><td colspan="4">-0.139</td><td colspan="5">-0.055</td><td>-0.050</td></tr>
<tr><td>WTIM-LLS</td><td colspan="4">-0.005</td><td colspan="6">-0.011</td><td colspan="4">-0.042</td><td colspan="3">-0.164</td><td colspan="5">-0.097</td><td>-0.092</td></tr>
<tr><td>LLS-HLS</td><td colspan="14">0.019</td><td colspan="4">-0.002*</td><td colspan="4">0.012</td><td>-0.007</td></tr>
<tr><td>LLS-Tapis</td><td colspan="8">0.015</td><td colspan="14">-0.035</td><td>-0.050</td></tr>
<tr><td>LLS-Saharan</td><td colspan="8">0.062</td><td colspan="6">0.029</td><td colspan="8">-0.008*</td><td>-0.070</td></tr>
<tr><td>LLS-Brent</td><td colspan="8">0.074</td><td colspan="6">0.039</td><td colspan="8">-0.005*</td><td>-0.079</td></tr>
<tr><td>LLS-Bonny</td><td colspan="8">0.070</td><td colspan="6">0.015</td><td colspan="8">-0.018*</td><td>-0.088</td></tr>
<tr><td>HLS-Tapis</td><td colspan="7">-0.002*</td><td colspan="15">-0.045</td><td>-0.043</td></tr>
<tr><td>HLS-Saharan</td><td colspan="8">0.044</td><td colspan="9">0.008</td><td colspan="5">-0.022</td><td>-0.066</td></tr>
<tr><td>HLS-Brent</td><td colspan="8">0.056</td><td colspan="9">0.017</td><td colspan="5">-0.019</td><td>-0.075</td></tr>
<tr><td>HLS-Bonny</td><td colspan="8">0.051</td><td colspan="14">-0.015</td><td>-0.066</td></tr>
<tr><td>Saharan-Brent</td><td colspan="3">0.024</td><td colspan="19">0.006</td><td>-0.018</td></tr>
<tr><td>Brent-Bonny</td><td colspan="8">-0.004*</td><td colspan="10">-0.023</td><td colspan="4">-0.007</td><td>-0.003</td></tr>
</table>

Medium-medium differentials

<table>
<tr><th>Differential</th><th>97</th><th>98</th><th>99</th><th>00</th><th>01</th><th>02</th><th>03</th><th>04</th><th>05</th><th>06</th><th>07</th><th>08</th><th>09</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>Change</th></tr>
<tr><td>Mars-SHU</td><td colspan="6">0.166</td><td colspan="8">0.086</td><td colspan="8">0.039</td><td>-0.127</td></tr>
<tr><td>Oman-Mars</td><td colspan="5">0.053</td><td colspan="12">0.014</td><td colspan="5">0.052</td><td>-0.001</td></tr>
<tr><td>Oman-SHA</td><td colspan="22">0.063</td><td>–</td></tr>
<tr><td>Urals-Mars</td><td colspan="5">0.053</td><td colspan="4">-0.010*</td><td colspan="8">0.020</td><td colspan="5">0.060</td><td>+0.007</td></tr>
<tr><td>Urals-SHE</td><td colspan="4">0.189</td><td colspan="18">0.077</td><td>-0.112</td></tr>
<tr><td>Dubai-Mars</td><td colspan="17">0.014</td><td colspan="5">0.046</td><td>+0.032</td></tr>
<tr><td>Dubai-SHA</td><td colspan="22">0.053</td><td>–</td></tr>
</table>

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.

Table A.11: Summary statistics pre and post-break

Part 1: Within-area differentials

	Pre-break		**Post-break**			
Differential	Mean	Standard deviation	Mean	Standard deviation	Ratio of mean (Post/pre)	Ratio of std. dev. (post/pre)
Midland, TX						
WTIM-WTS	0.076	0.028	0.01	0.016	0.13	0.57
U.S. Gulf Coast						
LLS-Mars	0.149	0.048	0.062	0.028	0.42	0.58
LLS-SHU	0.270	0.102	0.105	0.040	0.39	0.39
LLS-Maya	0.299	0.086	0.141	0.054	0.47	0.63
HLS-Mars	0.128	0.049	0.054	0.022	0.42	0.45
HLS-SHU	0.250	0.100	0.097	0.033	0.39	0.33
HLS-Maya	0.279	0.083	0.133	0.049	0.48	0.59
Mars-Maya	0.151	0.056	0.079	0.038	0.52	0.68
Europe/Atlantic Basin						
Saharan-Urals	0.070	0.035	0.023	0.020	0.33	0.57
Saharan-SHE	0.195	0.092	0.089	0.046	0.46	0.50
Brent-Urals	0.058	0.036	0.018	0.017	0.31	0.47
Brent-SHE	0.183	0.088	0.084	0.044	0.46	0.50
Bonny-Urals	0.070	0.038	0.033	0.017	0.47	0.45
Bonny-SHE	0.194	0.085	0.099	0.039	0.51	0.46
Middle East/Asia						
Tapis-Oman	0.114	0.053	0.069	0.031	0.61	0.58
Tapis-Dubai	0.128	0.053	0.074	0.031	0.58	0.58
Tapis-SHA	0.196	0.091	0.110	0.049	0.56	0.54

Part 2: Across-area differentials

	Pre-break		**Post-break**			
Differential	Mean	Standard deviation	Mean	Standard deviation	Ratio of mean (Post/pre)	Ratio of std. dev. (post/pre)
Light-medium						
Tapis-Urals	0.119	0.058	0.063	0.022	0.53	0.38
Saharan-Oman	0.065	0.045	0.029	0.031	0.45	0.69
Saharan-Dubai	0.079	0.046	0.035	0.030	0.44	0.65
Brent-Oman	0.053	0.042	0.024	0.025	0.45	0.60
Brent-Dubai	0.067	0.045	0.030	0.024	0.45	0.53
Bonny-Oman	0.065	0.044	0.039	0.025	0.60	0.57
Bonny-Dubai	0.079	0.046	0.044	0.024	0.56	0.52
LLS-Oman	0.116	0.049	0.030	0.037	0.26	0.76
LLS-Urals	0.121	0.049	0.023	0.037	0.19	0.76
LLS-Dubai	0.130	0.050	0.035	0.037	0.27	0.74
HLS-Oman	0.097	0.048	0.021	0.035	0.22	0.73
HLS-Urals	0.101	0.046	0.015	0.033	0.15	0.71
HLS-Dubai	0.111	0.050	0.027	0.035	0.24	0.70
Light-heavy						
Tapis-Maya	0.296	0.089	0.181	0.052	0.61	0.58
Saharan-Maya	0.248	0.082	0.141	0.057	0.57	0.70
Brent-Maya	0.236	0.078	0.136	0.053	0.58	0.68
Bonny-Maya	0.236	0.078	0.151	0.050	0.64	0.64
Medium-heavy						
Oman-Maya	0.183	0.072	0.112	0.045	0.61	0.63

Notes: The pre-break sample runs from Jan. 1997 to Dec. 2008. The post-break sample runs from Jan. 2009 to July 2018.

Table A.12: Unit root test results for monthly data

	AIC		SIC	
Differential	ADF	ADF (BP)	ADF	ADF (BP)
WTIM-WTS	**-1.89**	-7.17	**-2.90**	-6.76
	(0.34)	(<0.01)	**(0.05)**	(<0.01)
LLS-HLS	-6.09	-7.03	-6.09	-7.03
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
LLS-Mars	**-2.08**	-4.40	**-2.53**	-5.27
	(.25)	(<0.01)	(**0.11**)	(<0.01)
LLS-SHU	**-2.15**	-4.71	**-2.51**	-4.29
	(0.22)	(<0.01)	(**0.11**)	(<0.01)
LLS-Maya	**-2.03**	-5.62	**-2.96**	-5.62
	(0.28)	(<0.01)	**(0.04)**	<(0.01)
HLS-Mars	**-2.15**	**-3.48**	**-2.15**	-5.69
	(0.23)	**(0.07)**	**(0.23)**	<(0.01)
HLS-SHU	**-1.90**	-4.83	**-2.42**	-6.11
	(0.33)	(<0.01)	**(0.14)**	<(0.01)
HLS-Maya	**-3.02**	-5.77	**-3.02**	-5.77
	(0.03)	(<0.01)	**(0.03)**	(<0.01)
Mars-Maya	-4.68	-7.77	-4.68	-7.77
	(<0.01)	(<0.01)	(<0.01)	<0.01
SHU-Maya	-5.21	-5.16	-5.21	-5.16
	(<0.01)	(<0.01)	(<0.01)	<0.01
Saharan-Urals	-3.46	-5.79	-3.99	- 5.79
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Saharan-SHE	**-2.44**	-4.76	-3.55	-4.76
	(**0.13**)	(<0.01)	(<0.01)	(<0.01)
Brent-Urals	**-3.08**	**-4.13**	-4.23	-5.77
	(0.03)	**(0.012)**	(<0.01)	(<0.01)
Brent-SHE	**-2.04**	-5.00	-3.70	-5.00
	(0.27)	(<0.01)	(<0.01)	(<0.01)
Bonny-Urals	**-2.68**	-6.23	-4.89	-6.23
	(0.08)	<(0.01)	(<0.01)	(<0.01)
Bonny-SHE	**-2.46**	-5.00	-3.85	-5.00
	(0.13)	(<0.01)	(<0.01)	(<0.01)
Tapis-Oman	-3.72	-4.49	- 3.72	- 6.75
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Tapis-Dubai	-4.08	-7.01	-5.74	-7.01
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Tapis-SHA	-3.66	-4.23	-4.74	-5.54
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Oman-Dubai	-5.08	-5.22	-5.09	-4.68
	(<0.01)	(<0.01)	(<0.01)	(<0.01)

Notes: For each differential, the first row shows the test statistics for the Augmented Dickey-Fuller (ADF) and the ADF breakpoint (ADF BP) tests. The second row shows the p-value for the test. Bold text identifies a case where the null of a unit root would not be rejected at a one percent significance level.

Table A.13: Unit root test results for monthly data

	AIC		SIC	
Differential	ADF	ADF (BP)	ADF	ADF (BP)
Tapis-Urals	-3.75	-4.83	-3.75	-8.75
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Tapis-Mars	-3.61	-7.17	- 6.62	-7.61
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Saharan-Oman	-4.72	-5.60	-5.81	-6.63
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Saharan-Dubai	-4.36	-5.42	-5.49	-6.56
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Saharan-Mars	**-3.27**	-4.97	-5.87	-6.75
	(0.02)	(<0.01)	(<0.01)	(<0.01)
Brent-Oman	-4.72	-5.45	-5.79	-6.44
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Brent-Dubai	-4.13	-4.91	-5.12	-5.91
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Brent-Mars	**-3.08**	-5.01	-6.15	-6.81
	(0.03)	(<0.01)	(<0.01)	(<0.01)
Bonny-Oman	-4.95	-7.08	-6.68	-7.08
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Bonny-Dubai	-4.43	-6.85	-6.17	-6.85
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Bonny-Mars	**-3.10**	-7.49	-7.00	-7.49
	(0.03)	(<0.01)	(<0.01)	(<0.01)
LLS-Oman	**-1.85**	-5.74	-3.88	-5.74
	(0.35)	(<0.01)	(<0.01)	(<0.01)
LLS-Urals	**-2.40**	-5.08	**-2.97**	-5.08
	(0.14)	(<0.01)	**(0.04)**	(<0.01)
LLS-Dubai	-3.58	-5.49	-3.58	-5.49
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
HLS-Oman	**-3.23**	-6.91	-4.61	-6.48
	(0.02)	(<0.01)	(<0.01)	(<0.01)
HLS-Urals	**-1.44**	-5.83	-3.50	-7.24
	(0.56)	(<0.01)	(<0.01)	(<0.01)
HLS-Dubai	**-3.10**	-6.53	-4.26	-6.17
	(0.03)	(<0.01)	(<0.01)	(<0.01)
Tapis-Maya	**-2.30**	-6.39	-4.69	-6.39
	(0.17)	(<0.01)	(<0.01)	(<0.01)
Saharan-Maya	**-2.42**	-5.98	-4.47	5.98
	(0.14)	(<0.01)	(<0.01)	(<0.01)
Brent-Maya	-4.37	-5.95	-4.37	-5.95
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Bonny-Maya	-4.58	-5.99	-4.34	-5.60
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Oman-Maya	-5.58	-6.98	-5.58	-6.98
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Urals-Maya	-4.20	-7.04	-5.67	-6.87
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Dubai-Maya	-5.16	-6.29	-4.88	-5.94
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
WTIC-LLS	**-2.35**	**-2.86**	**-2.35**	-4.29
	(0.16)	**(0.21)**	**(0.16)**	(<0.01)
WTIM-LLS	**-1.17**	**-2.55**	**-2.32**	-4.63
	(0.69)	**(0.32)**	**(0.17)**	(<0.01)
LLS-Tapis	**-2.50**	-8.49	-6.73	-8.49
	(0.12)	(<0.01)	(<0.01)	(<0.01)
LLS-Saharan	**-2.12**	-4.96	**-3.17**	7.51
	(0.24)	(<0.01)	**(0.02)**	(<0.01)
LLS-Brent	**-1.37**	-6.67	**-2.82**	-6.67
	(0.60)	(<0.01)	**(0.06)**	(<0.01)
LLS-Bonny	**-1.58**	-4.56	**-2.62**	-5.77
	(0.49)	(<0.01)	**(0.09)**	(<0.01)
HLS-Tapis	**-2.70**	-9.77	-7.94	-9.77
	(0.07)	(<0.01)	(<0.01)	(<0.01)
HLS-Saharan	**-2.35**	-8.56	-3.87	-8.56
	(0.16)	(<0.01)	(<0.01)	(<0.01)
HLS-Brent	**-1.57**	-7.46	**-3.45**	-7.46
	(0.50)	(<0.01)	**(0.01)**	(<0.01)
HLS-Bonny	**-2.04**	-7.12	**-2.70**	-9.45
	(0.27)	(<0.01)	**(0.08)**	(<0.01)
Oman-Urals	-8.77	-9.25	-8.77	-9.25
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Oman-Mars	-3.80	-6.18	-8.03	-8.47
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Urals-Dubai	-5.46	-8.57	-8.58	-8.57
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Urals-Mars	**-2.17**	**-1.67**	-7.41	-7.71
	(0.22)	**(0.68)**	(<0.01)	(<0.01)
Dubai-Mars	**-3.12**	-5.11	-6.99	-7.37
	(0.03)	(<0.01)	(<0.01)	(<0.01)

Notes: For each differential, the first row shows the test statistics for the Augmented Dickey-Fuller (ADF) and the ADF breakpoint (ADF BP) tests. The second row shows the p-value for the test. Bold text identifies a case where the null of a unit root would not be rejected at a one percent significance level.

B Additional refinery data

B.1 U.S. refinery data

The U.S. refinery data comes from the Energy Information Administration. We use the "Downstream processing of fresh feed input" series and the "Downstream charge capacity" series to construct utilization rates for U.S. conversion capacity. The fresh feed input series is available at https://www.eia.gov/dnav/pet/pet˙pnp˙dwns˙dc˙nus˙mbblpd˙m.htm. The charge capacity data is available at https://www.eia.gov/dnav/pet/pet˙pnp˙capchg˙dcu˙nus˙a.htm. There is a single series available for the input and capacity for cokers. For the cracker data we combine the input and capacity series for catalytic cracking and catalytic hydrocracking. The units for the capacity data is barrels per calendar day, which means the capacity series is adjusted to take into account normal downtime at those units. As a result, utilization rates at or above 100 percent are theoretically possible. Capacity data is unavailable in 1996 and 1998. For those two years, we average the capacity data from the preceding and following year to construct an estimate. Tables B.1 and B.2 show the full time series for the U.S. and the U.S. Gulf Coast.

Table B.1: U.S. conversion capacity data

	Cracking			Coking		
Year	Capacity	Input	Utilization	Capacity	Input	Utilization
1987	6058	5316	87.7	1347	1265	93.9
1988	6235	5445	87.3	1394	1364	97.8
1989	6172	5505	89.2	1394	1345	96.5
1990	6322	5635	89.1	1425	1356	95.2
1991	6458	5794	89.7	1499	1423	94.9
1992	6554	5849	89.2	1459	1456	99.8
1993	6520	5990	91.9	1555	1514	97.4
1994	6491	5856	90.2	1635	1540	94.2
1995	6498	5964	91.8	1677	1574	93.9
1996	6513	6049	92.9	1708	1654	96.8
1997	6529	6181	94.7	1739	1691	97.2
1998	6711	6257	93.2	1797	1752	97.5
1999	6893	6234	90.4	1854	1758	94.8
2000	7004	6375	91.0	1962	1790	91.2
2001	7081	6312	89.1	2070	1963	94.8
2002	7097	6408	90.3	2148	2035	94.7
2003	7138	6498	91.0	2173	2026	93.2
2004	7231	6689	92.5	2247	2060	91.7
2005	7258	6441	88.7	2296	2054	89.5
2006	7275	6501	89.4	2330	2085	89.5
2007	7451	6533	87.7	2359	2034	86.2
2008	7456	6254	83.9	2390	1990	83.3
2009	7440	6196	83.3	2429	1908	78.6
2010	7339	6295	85.8	2388	1996	83.6
2011	7482	6419	85.8	2397	2094	87.4
2012	7318	6430	87.9	2499	2177	87.1
2013	7569	6481	85.6	2596	2303	88.7
2014	7651	6548	85.6	2687	2337	87.0
2015	7707	6545	84.9	2686	2352	87.6
2016	7718	6773	87.8	2651	2396	90.4
2017	7751	6817	87.9	2689	2379	88.5

Notes: Units for capacity and inputs are in thousands of barrels per day. Utilization is inputs divided by capacity multiplied by 100.

Table B.2: U.S. Gulf Coast conversion capacity data

	Cracking			Coking		
Year	Capacity	Input	Utilization	Capacity	Input	Utilization
1987	2827	2460	87.0	535	509	95.1
1988	2955	2527	85.5	556	552	99.3
1989	2882	2601	90.2	542	563	103.9
1990	2979	2651	89.0	566	552	97.6
1991	3036	2666	87.8	612	588	96.1
1992	3132	2663	85.0	612	623	101.8
1993	3077	2794	90.8	639	653	102.2
1994	3071	2710	88.2	698	680	97.4
1995	3084	2796	90.7	715	724	101.3
1996	3158	2927	92.7	756	778	102.9
1997	3232	2977	92.1	797	773	97.0
1998	3298	3019	91.5	840	817	97.2
1999	3364	3047	90.6	884	857	97.0
2000	3474	3115	89.7	923	843	91.3
2001	3508	3089	88.1	1009	999	99.0
2002	3563	3181	89.3	1086	1095	100.9
2003	3614	3236	89.5	1133	1086	95.8
2004	3640	3354	92.1	1206	1141	94.6
2005	3642	3073	84.4	1229	1132	92.1
2006	3616	3134	86.7	1255	1173	93.5
2007	3627	3181	87.7	1274	1152	90.4
2008	3646	3009	82.5	1282	1073	83.7
2009	3661	3117	85.1	1294	1041	80.5
2010	3780	3242	85.8	1322	1114	84.3
2011	3775	3315	87.8	1318	1183	89.7
2012	3777	3353	88.8	1373	1223	89.1
2013	3910	3452	88.3	1459	1299	89.0
2014	4035	3477	86.2	1479	1312	88.7
2015	4084	3554	87.0	1490	1347	90.4
2016	4062	3612	88.9	1458	1351	92.7
2017	4073	3559	87.4	1485	1338	90.1

Notes: Units for capacity and inputs are in thousands of barrels per day. Utilization is inputs divided by capacity multiplied by 100.

B.2 Other refinery data

Table B.3 shows the full time series available from the International Energy Agency for refinery capacity additions. Table B.4 shows the full time series available from Eni reports along with the source year for each observation.

Table B.3: Refinery capacity additions based on International Energy Agency data

Year	Primary	Conversion	Desulfurisation
2006	1.256	0.735	2.684
2007	0.786	0.573	0.838
2008	1.125	1.069	1.396
2009	2.236	1.643	2.755
2010	1.026	0.643	1.645
2011	0.652	1.078	1.096
2012	0.555	1.037	1.066
2013	0.547	0.918	1.433
2014	0.832	1.023	1.032
2015	1.440	1.473	1.419
2016	0.033	0.757	0.329
2017	0.623	0.358	0.466
2018	1.269	0.861	0.888

Notes: Units are growth rates in millions of barrels per day. The International Energy Agency tables list the data as Refining Capacity Additions and Expansions, Upgrading Capacity Additions, and Desulphurisation Capacity Additions. The sources are Medium-Term Oil Market Reports for 2006 - 2009, Medium-Term Oil & Gas Markets 2010 - 2011, Medium-Term Oil Market Reports for 2012 - 2016, and Oil Market Reports for 2017 and 2018.

Table B.4: Full time series from Eni publications

Year	Primary capacity	Conversion capacity	Conversion ratio	Nelson Complexity	Source
1995	76.3	27.5	0.36	7.9	Eni 2013
1996	77.8	28.0	0.36	N/A	Eni 2008
1997	80.0	29.6	0.37	N/A	Eni 2006
1999	82.1	31.2	0.38	N/A	Eni 2008
2000	83.2	31.6	0.38	7.9	Eni 2018
2001	84.7	33.9	0.40	N/A	Eni 2006
2003	85.5	36.7	0.43	N/A	Eni 2008
2005	87.3	37.5	0.43	8.2	Eni 2018
2007	89.2	40.1	0.45	N/A	Eni 2008
2010	92.4	43.4	0.47	8.7	Eni 2018
2011	93.4	46.7	0.50	7.8	Eni 2012
2012	94.0	47.9	0.51	8.0	Eni 2013
2013	95.1	49.4	0.52	8.0	Eni 2014
2014	96.3	48.1	0.50	9.0	Eni 2015
2015	96.5	50.2	0.52	9.1	Eni 2016
2016	98.1	52.0	0.53	9.3	Eni 2017
2017	98.7	53.3	0.54	9.3	Eni 2018

Notes: Units for capacity are millions of barrels per day. Conversion ratio is conversion capacity divided by primary capacity. Eni 2018 refers to the World Oil Review 2018 volume 1. The other reports are titled World Oil & Gas Review.