NBER WORKING PAPER SERIES

CRUDE OIL PRICE DIFFERENTIALS AND PIPELINE INFRASTRUCTURE

Shaun McRae

Working Paper 24170
http://www.nber.org/papers/w24170

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2017

I thank participants in the NBER Hydrocarbon Infrastructure and Transportation workshop, especially the organizers Jim Bushnell, Ryan Kellogg, and Erin Mansur, for many helpful comments and suggestions. I also thank the Sloan Foundation for their generous financial support. Jordán Mosqueda provided valuable research assistance. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Crude Oil Price Differentials and Pipeline Infrastructure
Shaun McRae
NBER Working Paper No. 24170
December 2017
JEL No. L71,L95,Q35,Q41

ABSTRACT

Crude oil production in the United States increased by nearly 80 percent between 2008 and 2016, mostly in areas that were far from existing refining and pipeline infrastructure. The production increase led to substantial discounts for oil producers to reflect the high cost of alternative transportation methods. I show how the expansion of the crude oil pipeline network reduced oil price differentials, which fell from a mean state-level difference of $10 per barrel in 2011 to about $1 per barrel in 2016. Using data for the Permian Basin, I estimate that the elimination of pipeline constraints increased local prices by between $6 and $11 per barrel. Slightly less than 90 percent of this gain for oil producers was a transfer from existing oil refiners and shippers. Refiners did not pass on these higher costs to consumers in the form of higher gasoline prices.

Shaun McRae
ITAM, Centro de Investigación Económica
Camino a Santa Teresa No. 930
Col. Héroes de Padierna
Ciudad de México
Mexico
shaun.mcrae@itam.mx

1 Introduction

Technological advances in oil production have reshaped the geography of oil markets in the United States. Areas where production had been low or declining, such as the Bakken shale region in North Dakota and the Permian Basin in western Texas, experienced rapid and unanticipated growth. However, the boom in oil production overwhelmed the capacity of existing infrastructure for transporting that oil. Oil producers received prices heavily discounted to the price at market hubs, reflecting the higher cost of transporting oil by pipeline alternatives such as rail or truck.

Price differentials created opportunities for investment in oil pipelines. Between 2010 and 2015, the total length of crude oil pipelines in the United States grew from 54,600 to 73,200 miles. Pipeline owners also reconfigured and expanded their existing networks to match the new market conditions. The new pipelines improved transportation access not only out of the Bakken and Permian regions but also from the Rocky Mountains, Great Plains and Midwest to the crude oil hub in Cushing, Oklahoma. A small fraction of these oil pipeline projects have been controversial, most notably the Keystone XL and the Dakota Access pipelines.[1]

Most discussion of pipeline costs and benefits misses the primary economic motivation for investment in transportation infrastructure: the reduction in trade costs. For pipeline supporters, the raw materials and labor used in construction are the primary benefits of the projects.[2] For opponents, the main cost of building a pipeline is the environmental risk associated with possible future oil spills. Both viewpoints ignore the economic effects. New oil pipelines reduce the cost of transporting oil out of expanding production regions to the world market. The cost reduction means that oil producers face a smaller discount to the world oil price for their output.

At an aggregate level, there appear to have been substantial benefits from the expansion of the oil pipeline network. The production-weighted discount of the oil price received by

[1]The proposed Keystone XL pipeline runs from Alberta, Canada to the United States Gulf Coast. In 2015, President Obama denied approval for the section from Alberta to Nebraska, although President Trump reversed this decision in 2017. Environmental groups campaigned against the pipeline due to concerns about greenhouse gas emissions from oil sands production in Canada. The Dakota Access project is a 1,172-mile crude oil pipeline from North Dakota to Illinois. The pipeline route passes close to the Standing River Sioux reservation. The tribe opposed its construction because of concerns about water contamination in the event of an oil spill (Crooks, 2017).

[2]For example, in signing a memorandum that advanced the Keystone and Dakota Access projects, President Trump discussed the 28,000 "great construction jobs" and a potential new requirement to build the pipelines out of United States steel (Lynch et al., 2017).

producers in the United States, relative to a world price benchmark, has fallen from $28 per barrel in September 2011 to $5 per barrel in December 2016. Volumes of oil transported by rail have fallen from a peak of 35 million barrels during October 2014 to 13 million barrels in December 2016. Although these aggregate changes appear correlated with the pipeline network expansion, there may be other contributing factors. In particular, world oil prices have halved since 2014, and oil production in the United States has fallen by 9 percent from its peak.

This paper decomposes the economic benefits from the expansion of the pipeline network. Additional pipeline capacity narrows the discount for oil producers from the world oil price. Higher prices for producers mean higher costs for refiners, so part of the benefit for producers is a transfer from oil refiners. There is also a transfer from oil shippers to oil producers, as shippers with access to the constrained pipeline capacity had earned profits from the price arbitrage. Two other components of the reduced discount represent overall welfare gains for society. First, there is a reduction in transportation costs for the oil carried by the new pipeline instead of a higher priced alternative. Second, there is an addition to oil production as the result of the increase in producer prices.

I quantify these economic effects at the level of an oil-producing region. The primary analysis focuses on the Permian Basin in western Texas and southeastern New Mexico. I also provide summary results for the Bakken region and Colorado. For each area, I assemble a dataset of monthly pipeline capacity, incorporating publicly available information on the construction and expansion of the pipeline network. I combine this data with information on oil production, refinery capacity, flow volumes, and producer prices. There are two sources for producer prices: monthly average wellhead prices at a state level (used for the Bakken and Colorado analyses) and daily crude oil prices at major trading hubs (used for the Permian Basin).

I use this data to construct a measure of excess production for each exporting region: the monthly oil output, less the crude oil input capacity of the local refineries, less the pipeline export capacity. I estimate the price discount as a function of this variable, then use these estimates to simulate the effect on producer prices of expansion in pipeline capacity. For the Permian Basin, a hypothetical pipeline that eliminates excess production would increase the producer price by between $6 and $11 per barrel. The results are used to calculate the increase in profits for oil producers and the reduction in profits for oil refiners and oil shippers. Most of the overall welfare gain is due to the elimination of higher shipping costs. However, these benefits are smaller in magnitude than the transfer of the surplus from refiners and shippers to oil producers. For the Permian Basin, slightly less than 90 percent of the gains for producers is a transfer from refiners and shippers.

Previous papers study the causes of price differentials in crude oil markets. Between 2011 and 2014, the benchmark price for the United States interior, known as WTI Cushing, traded at a substantial discount to international oil prices. Most closely related to this paper, Agerton and Upton (2017) test alternative explanations for the price differential between inland and coastal crude prices: pipeline constraints, refining constraints, and the crude oil export ban. They use the proportion of shipments by rail and barge to quantify the pipeline constraints and find that these explain most of the differential. McRae (2015) studies the strategic incentives for an integrated refinery and pipeline owner to delay the alleviation of pipeline constraints and profit from lower refinery input costs due to the discounted prices. Buyuksahin et al. (2013) test alternative explanations for the WTI discount, highlighting storage capacity constraints at Cushing as a contributing factor. Other papers have focused on the effect of a transportation constraint that was due to regulation, not infrastructure: the ban on exports of crude oil from the United States that Congress lifted in December 2015 (Brown et al., 2014; Melek and Ojeda, 2017).

This paper makes several contributions to this existing literature on price differentials in crude oil markets. Rather than focusing on the WTI Cushing discount and the infrastructure constraints at Cushing, I analyze the general phenomena of oil price differentials and their relationship to local and regional pipeline capacity. I use data on expansions to the oil pipeline network to infer the periods in which capacity constraints are binding. The pipeline data reveals changes in the marginal transportation method. As predicted by the theoretical model, the results show that switching the marginal barrel from rail to pipeline will narrow the price discount. This empirical approach also allows for a decomposition of the welfare effects of incremental changes to the pipeline network.

This paper complements the literature on the distributional effects of changes in energy prices. Hausman and Kellogg (2015) study the gains and losses by sector of the drop in natural gas prices in the United States due to the availability of shale gas. They find that the shale revolution made natural gas producers are worse off overall, due to lower revenue for inframarginal production more than offsetting the profit for higher output. Borenstein and Kellogg (2014) study the passthrough of the WTI benchmark to gasoline. They find no evidence of lower retail prices, suggesting that oil refiners captured the full benefit of the wholesale price discount.

Finally, this paper contributes to the growing literature on the economic value of infrastructure investments.[3] Improved transportation and communication connections have been shown to increase economic growth through a reduction in trade costs. Donaldson (2012) finds that the expansion of the railroad network in India reduced trade costs, reduced price differentials between regions, and increased trade. Expansion of the railroad network in the United States increased agricultural land values due to improvements in market access (Donaldson and Hornbeck, 2016). Interstate highways in a city increased specialization in sectors with high weight-to-value ratios (Duranton et al., 2014) and increased employment (Duranton and Turner, 2012).

The next section provides descriptive evidence on the relationship between price differentials and pipeline investment. Section 3 describes a stylized model of price differentials and provides details of the empirical methodology. Section 4 provides the main analysis of the paper, focusing on the changes in price differentials for the Permian Basin caused by increases in oil supply and the subsequent expansion of pipeline infrastructure. Section 5 summarizes the results for other regions. Section 6 concludes.

2 Descriptive evidence on price differentials and pipeline infrastructure

After a long period of decline since the early 1970s, crude oil production in the United States rebounded in the ten years after 2007. Production rose from 5.0 million barrels per day in 2008 to 8.9 million barrels per day in 2016 (Table 1). The halving of oil prices at the end of 2014 caused only a small drop in production, which was higher in 2016 than in 2014 even though producer prices had dropped by 56 percent.

This production boom was due to the development of horizontal drilling, microseismic imaging, and hydraulic fracturing technologies. These innovations allowed for the profitable exploitation of oil and natural gas resources in shale formations (Kilian, 2016). Because of the new technology, much of the increase in oil production occurred in locations that are distant from existing major oil fields. These areas lacked both the refining capacity to process the crude oil locally as well as sufficient pipeline infrastructure to transport the oil

[3] Many papers focus on the value of infrastructure projects in developing countries. For electrification, Lipscomb et al. (2013) find large effects of electrification on development indicators in Brazil, while Burlig and Preonas (2016) find no medium-run development effects from a rural electrification program in India. Electrification reduced indoor air pollution (Barron and Torero, 2017) and increased female labor force participation (Dinkelman, 2011; Grogan and Sadanand, 2013). Nevertheless, household willingness to pay for electrification has been found to be below the cost of infrastructure provision (Lee et al., 2016).

to distant refineries.

The opportunity to transport oil from the new production regions led to the construction of many new oil pipelines. Total crude oil pipeline length within the United States increased from about 51,000 miles in 2008 to over 75,000 miles in 2016. The length of interstate oil pipelines increased by 37 percent between 2010 and 2016. Not all of this increase in pipeline length was the result of new construction. Some owners of natural gas, propane, or refined petroleum product systems converted these to carry crude oil. Also, there were several projects to expand the capacity of existing oil pipelines or to reverse the direction of flow. The increase in pipeline capacity led to more oil being transported by pipeline. Deliveries to refineries by pipeline increased from 7.3 to 10.2 million barrels per day between 2008 and 2016 (Figure 1).

When pipeline capacity is insufficient, oil shippers use more expensive alternatives such as rail and truck. Trainloads of oil pulling into refineries increased from 10,000 barrels per day in 2010 to 430,000 barrels per day in 2014. Truckloads more than doubled over the same period. These refinery statistics understate the growth of rail shipments, because shippers may transfer crude oil from trains to pipelines for final delivery. Overall rail shipments within the United States rose from 60,000 barrels per day in 2010 to 870,000 barrels per day in 2014.

In absolute terms, the growth in pipeline deliveries overshadowed the increase in rail shipments (Figure 1). After 2014, the use of rail declined. Refinery deliveries fell by about 25 percent between 2014 and 2016, and total rail shipments more than halved. Given that oil production and pipeline deliveries were still growing over this period, the rail decline is almost certainly due to the increased availability of cheaper pipeline infrastructure. In the two years after 2014, at the same time that rail deliveries were falling, pipeline length increased by 13 percent and pipeline deliveries by 8 percent.

The switch between pipeline and rail for oil transportation has a material effect on oil prices. There are many different oil price measures. Buyers and sellers of oil use benchmark prices for pricing oil shipments sold under long-term contracts. These contracts specify the price differential to the benchmark (Fattouh, 2011). The two most widely used oil price benchmarks are WTI Cushing and Brent.[4] Other widely-used price benchmarks within the United States are Light Louisiana Sweet ("LLS") and WTI Midland. The wellhead price

[4]The WTI Cushing price is an assessed price based on small transaction volumes at the pipeline hub in Cushing, Oklahoma. There is a liquid, globally-traded futures market linked to WTI. However, because the underlying commodity is delivered oil at Cushing, small changes in the supply and demand balance at Cushing can have substantial effects on WTI futures prices.

is the price per barrel received by oil producers and will be lower than international price benchmarks, with the difference reflecting marketing and transportation costs.

Price dispersion for wellhead oil prices in the United States peaked in 2011 and has been declining since then (Figure 2). The absolute price deviation is the non-negative difference between the mean wellhead price in a state and the mean for the entire country. I calculate the monthly average of these state-level price deviations, weighted by the oil production in each state. The mean absolute price deviation increased from \$2.16 per barrel in 2010 to \$6.02 per barrel in 2012, then declined to \$1.21 per barrel in 2016.

An alternative measure of producer price differentials is the difference between a benchmark price such as LLS and the wellhead price. The mean wellhead discount to LLS, weighted by state-level production, increased to over \$25 per barrel in 2011 and then declined to an average of \$6.33 per barrel in 2016. Both measures of price differentials follow a similar pattern.

Part of the explanation for the price differentials in 2011 and 2012 was the dislocation in oil markets due to the pipeline capacity constraints out of Cushing. The glut of oil accumulating at Cushing pushed down the WTI Cushing price relative to the LLS price on the Gulf Coast and the Brent benchmark price. In response, pipeline companies constructed a new pipeline and reversed and expanded an existing pipeline between Cushing and the Gulf Coast. These pipeline projects alleviated the constraints at Cushing. The price differential between WTI Cushing and LLS fell from \$17.41 per barrel in 2012 to \$3.67 per barrel in 2014.

The pipeline projects between Cushing and the Gulf Coast nearly eliminated the differential between the WTI Cushing and LLS price benchmarks. However, as described above, many other oil pipelines have been constructed, expanded, or reconfigured. These projects also affected the prices received by oil producers. Both measures of oil price differentials continued to decline, even after 2014 (Figure 2). This decline was coincident with the increase in pipeline shipments and decline in crude-by-rail.

3 Stylized model of pipeline expansion

In this section, I describe a simple model of price differentials in oil markets and how these are affected by the expansion of oil pipeline networks. This model can explain the observations about price differentials, pipeline availability, and oil shipments in Section 2.

3.1 Theoretical framework

Suppose there is a small isolated market with local oil production S (Figure 3). This market has a local refinery with demand for crude oil D^R. With no exports, the oil price in autarky will be p_0, with oil production equal to refinery consumption q_0.

The world price of oil p^W is higher than p_0. Suppose there are two methods available for transporting oil to the world market, pipeline and rail, with marginal costs $c^{pipe} < c^{rail}$. Both transportation methods have a limited capacity: K^{pipe} and K^{rail}. When there are exports, the oil producers in this market will receive p^W, less the marginal cost of the highest cost transportation method.

The local refinery demand, combined with the two capacity-constrained transportation methods, give an oil demand curve D^{RX}. The intersection of D^{RX} with S determines the oil price p_1. The pipeline will be used to capacity. Some rail will be used to export oil, but there will still be spare capacity. Oil production will be higher than under autarky, q_1 instead of q_0.

Now suppose there is an expansion of the pipeline to $K^{pipe'}$. This shifts out D^{RX} to $D^{RX'}$. With the expanded capacity, the demand curve crosses S at p_2, meaning that the pipeline can carry all exports with no need for rail. Local oil producers receive a higher price p_2 and their production increases to q_2.

Who are the winners and losers of this expansion of pipeline capacity? Local oil producers benefit from the higher price and are better off by the area $A+B+C+D$. However, the local oil refinery is worse off by the area A. It now pays the price p_2 for its input oil purchases, not p_1. Oil shippers who had access to the original pipeline capacity are also worse off by the area B. Before the pipeline expansion, they could buy oil at a price p_1 and sell it at a price p^W, less the pipeline charge c^{pipe}. They lose their profit of $c^{rail} - c^{pipe}$ for each barrel shipped.

The areas A and B represent transfers from the refinery and shippers to the local oil producers. In contrast, C and D represent overall welfare gains from the pipeline expansion. Eliminating the use of rail for transportation reduces shipping costs by C. Higher prices move oil producers along their supply curves, increasing production from q_1 to q_2 and profits by the area D.

Most of the benefits from greater pipeline capacity require the expansion to be large enough to eliminate the use of rail. For a smaller pipeline project, $D^{RX'}$ would still interest S at p_1. Such an expansion would produce welfare benefits from switching some oil transportation from rail to pipeline. Oil shippers with access to the pipeline capacity would capture the benefits of the lower costs. There would be no change to the welfare of the local oil refinery or oil producers.

The empirical methodology described in the next section will estimate the relationship between pipeline capacity and local prices, then use this to perform the decomposition illustrated in Figure 3 for a hypothetical pipeline.

3.2 Empirical details

Let p_t be the observed equilibrium price and q_t be the observed production quantity, for an oil-producing region in period t. There is a world market for oil with a price p_t^W in period t. I model the price differential, p_t^D:

$$p_t^D = p_t^W - p_t$$

This price differential depends on the relationship between oil production in the region and the available infrastructure to refine or export the oil. Let K_t^{ref} be the oil refining capacity in the region and K_t^{pipe} be the capacity of oil pipelines out of the region. Excess production, q_t^X, is defined as:

$$q_t^X = q_t - K_t^{ref} - K_t^{pipe} \tag{1}$$

The price differential p_t^D is a nonlinear function of the excess production q_t^X (Figure 4). This function is modeled as in Equation 2:

$$p_t^D = \beta_0 + \beta_1 q_t^X + \beta_2 I(q_t^X > 0) + \beta_3 I(q_t^X > 0) q_t^X + \varepsilon_t \tag{2}$$

For $q_t^X < 0$, the marginal barrel of oil production is exported by pipeline. Higher oil production may lead to higher shipping costs and so increased price differentials ($\beta_1 > 0$). In part, this may reflect the use of successively higher-cost pipelines. Furthermore, oil shippers sign long-term forward contracts with pipeline owners. Spot shipments that exceed the contracted quantity will pay significantly higher tariffs.

Once the refinery and pipeline capacity is exhausted ($q_t^X > 0$), additional oil production will have to be exported by higher cost transportation methods such as rail. These methods may require a fixed cost ($\beta_2 > 0$) representing the additional loading and unloading expenses. The marginal cost of shipment may also be higher than for pipelines (β_3).

One complication for the estimation of Equation 2 is that there may be multiple pipelines out of the region, potentially with different end locations and prices. Suppose some pipelines end at the world market with price p_t^W and other pipelines end at location V with price p_t^V. If V is the destination for the marginal barrel of oil, then the local oil price will be equal to p_t^V less the pipeline shipping costs. Equation (2) would then overstate the pipeline cost component of p_t^D, as the price differential $p_t^W - p_t^V$ would be attributed to the pipeline cost.

The effect of multiple pipeline destinations on price differentials can be accounted for by including a correction for $p_t^W - p_t^V$, as in Equation (3):

$$p_t^D = \beta_0 + \beta_1 q_t^X + \beta_2 I(q_t^X > 0) + \beta_3 I(q_t^X > 0) q_t^X + \beta_4 I(q_t^X < 0 \ \& \ V_{marg})(p_t^W - p_t^V) + \varepsilon_t \tag{3}$$

This correction is only required when the marginal barrel of additional oil production is being exported by pipeline ($q_t^X < 0$) and when the final destination for that oil is region V instead of the world market W ($V_{marg} = 1$). The coefficient β_4 represents the share of the price differential $p_t^W - p_t^V$ that is passed through to the price differential $p_t^W - p_t$, for only those periods with V as the marginal pipeline destination.

4 Oil pipelines and the Permian Basin

4.1 Background information

The Permian Basin is a geologic region in western Texas and southeastern New Mexico that encompasses several subbasins, including the Midland and Delaware Basins.[5] The gush of oil from Santa Rita No. 1 in 1923 initiated its growth into one of the major oil-producing regions in the United States. Midland and Odessa became boom towns servicing the oil industry, with the population of Midland growing from 1,400 to 23,000 in three years after the Spraberry discovery in 1949 (Owen, 1951). Despite an occasional resurgence in drilling in response to higher oil prices, Permian oil production entered a long decline, with the basin described in 1978 as being "in the late afternoon of life and the sunset is almost in view" (Stevens, 1978).

The sun rose again in the Permian with the twin innovations of horizontal drilling and hydraulic fracturing. Producers were relatively slow to exploit these technologies in the Permian, with their first use for oil production occurring in the Bakken and Eagle Ford

[5] The Wolfcamp shale formation appears in both of these subbasins. In the Midland Basin, it lies beneath the Spraberry Trend (Gaswirth et al., 2016). The complex, multilayered geology of the Spraberry-Wolfcamp formation comprises one of the largest oil fields in the world.

shale formations (Warren, 2014). Production of oil in the Permian Basin grew from 920,000 barrels per day in 2010 to over 2 million barrels per day in 2016 (Figure 6 and Table 3). Unlike in the Bakken and Eagle Ford, output continued to grow even after the large drop in world oil prices at the end of 2014.

The increase in oil production quickly exceeded the capacity of the existing infrastructure to refine the oil locally or export it by pipeline. There are three oil refineries located in or nearby the Permian Basin: Big Spring and El Paso in Texas, and Navajo in New Mexico (Figure 5).[6] Apart from local refineries, four major crude oil pipelines were configured to export oil out of the Permian Basin (Table 2). The Basin and Centurion pipelines had their delivery point at the Cushing oil hub. The Phillips pipeline ran north to the Borger refinery.[7] Only the West Texas Gulf pipeline traveled east and southeast to delivery points in Longview, Texas and Goodrich, Texas.

By the middle of 2012, the volume of Permian crude exceeded the capacity of the local refineries and the available pipelines (Figure 6).[8] The excess production required the use of rail and truck to deliver the crude oil production to market. The volume of oil transported by rail within PADD 3 increased throughout 2012 and peaked in early 2013 (bottom graph of Figure 7). Most rail loading terminals in the PADD 3 region are in or near the Permian Basin (Figure 5), with the majority of the rail unloading terminals located at refineries on the Gulf Coast.[9] This increase in rail volumes during 2012 is consistent with the model presented in Section 3.1. As the supply curve for oil production shifted out, the quantity exceeded the refinery and pipeline capacity, leading to the use of more expensive forms of transportation.

As described by the model in Section 3.1, the spot price of oil for producers in the Permian Basin equals the world price of oil, less the transportation cost for the marginal barrel exported. The price differential between LLS and WTI Midland reached its maximum at the end of 2012, with an average price difference exceeding $30 per barrel for the month (middle graph of Figure 7). In part, this difference reflected the differential between the LLS and WTI Cushing prices, due to constraints on the infrastructure between those two

[6] Their total capacity is 297,000 barrels per day, an increase from 273,000 barrels per day in 2008 as the result of capacity upgrades at Navajo in 2008–09 and Big Spring in 2014–15.

[7] There are two refineries in northern Texas: Borger and Valero McKee. Separate Phillips pipelines connect both Cushing and Odessa to the Borger refinery.

[8] As discussed in the next section, I modify the nameplate capacity to reflect real-world operating constraints. All graphs and analysis use the adjusted numbers.

[9] As shown on the map, there are also three rail loading terminals located in the Eagle Ford oil production region of southern Texas. The public data does not allow me to distinguish between crude oil shipped out of the Eagle Ford and Permian basins.

locations. However, this does not explain the full differential, because the price difference between WTI Midland and WTI Cushing also increased to over $10 per barrel.

The low price of oil in the Permian Basin relative to the nearby market hubs encouraged investment in new pipeline infrastructure. By the end of 2016, six new pipelines were in service, resulting in the pipeline capacity out of the region more than doubling. Four carried Permian crude directly to the Gulf Coast refinery complex. The Cactus pipeline connected to new infrastructure constructed in the Eagle Ford region. The PELA pipeline delivered oil to Baton Rouge, Louisiana. Three of the projects (Bridgetex, Cactus, Permian Express II) were greenfield developments. Sunoco reversed the flow of its Amdel pipeline to carry oil from west-to-east instead of east-to-west. Magellan also changed the flow direction of its Longhorn pipeline and converted it back from refined products to crude oil. The PELA project comprised some new segments and some reconfiguration of existing pipelines.

The Amdel and Longhorn pipeline projects led to the refinery and pipeline capacity exceeding Permian production in the second part of 2013 (top graph in Figure 7). This period coincided with low price differentials and low rail volumes.

Nevertheless, the continuing growth in Permian crude volumes, coupled with the extra time required to complete greenfield pipeline projects, led to the second period of excess production during 2014. Rail volumes within PADD 3 were higher, although these did not reach their level of early 2013. The price differential between LLS and WTI Midland increased to a mean of $10.53 during 2014 (Table 3). Unlike in 2012, by 2014 the infrastructure constraints between Cushing and the Gulf Coast had been resolved. Because the differential between LLS and WTI Cushing was only a few dollars per barrel, the Cushing–Midland differential closely tracked the LLS–Midland differential.

The completion of new pipeline projects during 2014 and 2015 once again relaxed the pipeline infrastructure constraints. By the middle of 2015, excess production had fallen below -0.2 million barrels per day. The differentials between LLS, WTI Midland, and WTI Cushing were around $2 per barrel. Rail volumes in PADD 3 had fallen close to zero.

One notable feature of the data is the spike in rail volumes within PADD 3 during 2011, even though there was sufficient pipeline capacity to export all production (Figure 7). At this time, the marginal barrel out of the Permian Basin would be sent by pipeline to Cushing, meaning that the WTI Midland price was equal to the WTI Cushing price less the pipeline shipping cost. The unprecedented price differential between WTI Midland and LLS during 2011 was entirely due to the price differential between WTI Cushing and LLS, as during that year there were not yet infrastructure constraints out of Midland. This differential created an opportunity to send crude oil by train directly to the Gulf Coast and receive the LLS price less the rail cost.

4.2 Data

The essential explanatory variable for the analysis is excess production q_t^X, defined in Equation (1). I obtain monthly crude oil production q_t for the Permian Basin from EIA (2017). Section 4.1 provides the identity of the refineries and pipelines that are used to calculate the aggregate K_t^{ref} and K_t^{pipe}. Refinery capacity, in barrels per calendar day, is from annual EIA reports on refinery capacity. Pipeline capacity (including the in-service dates for new or reconfigured pipelines) is compiled from industry reports and news sources, particularly RBN Energy (2017) and Genscape (2015).

The maximum amount of oil that can be processed by an oil refinery may differ from its nominal capacity due to unplanned maintenance or outages, unanticipated changes in oil input quality, and so on. I approximate the typical input volumes of the three refineries using the average capacity utilization rate in the Inland Texas refining district, multiplied by the nominal capacity in barrels per calendar day.[10] Over the period 2008 to 2016, this average capacity utilization was 92 percent.

Similarly, the maximum amount of oil that can be transported by a pipeline may differ from its nominal capacity due to downtime for maintenance or repairs, or differences in characteristics of the oil. For example, reducing the API of crude oil from 40 to 35 (that is, making the oil more viscous) reduces pipeline throughput by over three percent (Genscape, 2015). I use proprietary data on pipeline volumes from Genscape to calculate the maximum flow observed, in all available months of data, for four pipelines: Basin, Centurion, Longhorn, and Bridgetex (Genscape, 2016a,b). In aggregate, the highest observed volumes are equal to 94 percent of their theoretical maximum. I apply this adjustment factor to the nominal capacity of all pipelines in the data.

Estimation of Equation (3) requires information on the periods in which the pipeline destination for the marginal barrel of oil is Cushing rather than the Gulf Coast. Using the Genscape data on volumes, I observe a steady increase in flows through the Basin and Centurion pipelines (with a delivery point at Cushing) until February 2013. Flows are

[10] There are two measures of oil refinery capacity: barrels per stream day and barrels per calendar day (`https://www.eia.gov/tools/glossary/index.php?id=B`). Barrels per stream day is the maximum amount that refiners can process under optimal conditions over a 24-hour period. Barrels per calendar day is a measure based on typical operating conditions.

always lower than the pipeline capacity. This evidence suggests a regime shift at that time, coincident with the completion of pipelines to the Gulf Coast (Table 2).[11]

4.3 Empirical results

Table 4 shows results of the estimation of Equations (2) and (3) using data for the Permian Basin. The dependent variable in all regressions is the monthly mean difference between the LLS and WTI Midland prices. Adding the term based on the price differential between WTI Cushing and LLS (β_4) increases the explanatory power of the regression (Column 2 of Table 4). This variable controls for the large price differentials during 2011 at a time when the pipeline capacity constraints out of the Permian Basin were not binding.

Adding year fixed effects to Equation (3) reverses the sign of β_3 and slightly reduces the magnitude of the other coefficients (Column 3 of Table 4). The year controls allow for changes in market structure over the ten years so that the effect of capacity constraints is estimated using only within-year variation in q_t^X.

The remaining columns provide additional robustness checks. Column 4 fixes the value of β_4 at -1, implying that there is full passthrough of the WTI Cushing to LLS price differential to the WTI Midland to LLS price differential during those periods when pipelines to Cushing receive the marginal oil output. Column 5 uses an alternative calculation for q_t^X in which there is no adjustment of refinery capacity based on average capacity utilization in the region.

In all specifications, there is a positive and statistically significant coefficient on q_t^X, meaning that higher oil production in the Permian Basin reduces the price received by oil producers relative to the world price. In all but the first specification, there is a positive and statistically significant coefficient on $I(q_t^X > 0)$, meaning that there is a further reduction in the price received by Permian oil producers when their production exceeds the available refinery and pipeline capacity. An increase in q_t^X from -150,000 to -50,000 barrels per day would reduce the price received by Permian producers by \$1.61 per barrel.[12] A further increase in q_t^X to 50,000 barrels per day, meaning the production exceeds the refinery and pipeline capacity, would reduce the price received by \$11.15 per barrel.[13] As illustrated in Figure 3, once production in a region exceeds the refining and pipeline capacity, there

[11] I formalize this analysis using weekly data for the Basin, Centurion, Longhorn, and Bridgetex pipelines. I regress the weekly change in deliveries to the Gulf Coast on the weekly change in total pipeline flows. For 2015, the results show that Gulf Coast deliveries absorb 85 percent of the weekly change in volumes.

[12] Using the estimates from Column 3, the effect of a production increment of 0.1 million barrels is $\beta_1(0.1) = 1.61$.

[13] Again using the estimates from Column 3, the price effect of this production increase is $\beta_1(0.1) + \beta_2 + \beta_3(0.05) = 11.15$.

is a discrete jump in price differentials (equivalently, a drop in the price received by local producers) that reflects the higher marginal cost of alternative methods of transportation.

There are few observations in the data for which q_t^X is positive. Price differentials are relatively noisy during these months. This data limitation makes it difficult to estimate β_3, the differential slope term for positive values of q_t^X. Therefore, estimates of this coefficient are imprecise and sensitive to small changes in the specification (Table 4).

Figure 8 plots the data used in the estimation, with the dependent variable shown on the vertical axis. Triangles indicate those months in which the marginal barrel of oil production is sent by pipeline to Cushing, meaning that the WTI Cushing–LLS differential affects the price differential. The line of best fit uses the estimates of Column 3 in Table 4, setting that the WTI Cushing–LLS price differential to zero. It illustrates the discrete jump in price differentials once production exceeds refining and pipeline capacity.

I use the estimates from Table 4 to simulate the effects of a hypothetical pipeline project completed during September 2014.[14] This pipeline is assumed to have a capacity to transport 100,000 barrels per day of oil out the Permian Basin, sufficient to reduce q_t^X to zero. Similar to the calculation above, this hypothetical pipeline would increase oil prices for local producers by between \$5.72 and \$10.86 per barrel (first row of Table 5). The columns of Table 5 are based on the results for alternative specifications provided in Columns 2 to 5 of Table 4.

Based on the stylized model in Figure 3, I decompose the effects of this price increase on oil producers, refiners, and shippers. The higher oil price reduces the profits of the oil refiners in the Permian Basin region by \$2.7 million per day, due to the higher cost for their crude oil inputs (Column 3 of Table 5). Oil shippers who had access to the existing pipelines are notably worse off from the new pipeline construction. Their profits fall by \$13 million per day, reflecting the diminished opportunity for buying cheap Permian oil and selling it at a higher price at Cushing or in the Gulf Coast. The losses for refiners and shippers are gains for oil producers, in the form of higher oil revenues.

There are overall welfare gains as the result of the hypothetical pipeline. These comprise the reduction in transportation costs for the oil that is shipped by pipeline instead of rail or truck, valued at \$0.95 million per day.[15] Based on an assumed price elasticity of oil supply of 1.0, the additional profit from higher oil production due to the higher oil price is about \$1 million per day.[16]

[14]The choice of month and year determines the baseline oil production and oil prices.

[15]This calculation assumes that there is no market power in the market for crude-by-rail transportation. If there are markups in the cost of rail, then part of the reduction in transportation costs will be a transfer from rail operators to oil producers, not an overall welfare gain for society.

[16]Using Texas data from 1990 to 2007, Anderson et al. (forthcoming) find that output from existing wells

Across the four model specifications, oil producers are better off by between \$9.8 and \$19.2 million per day as the result of the hypothetical pipeline project. Slightly less than 90 percent of this benefit is a transfer from oil refiners and shippers (final row of Table 5). There is limited variation in the composition of producer gains across the four specifications.

4.4 Additional robustness checks

In the welfare results for oil refiners in Table 5, higher costs for crude oil reduce refinery profits one-for-one. This relationship will be more complicated if refineries vary their production quantities in response to changes in oil prices, or if they pass higher costs on to end consumers. In this section I show that neither effect will change the qualitative results in Section 4.3.

Manufacturing firms increase input purchases, and potentially total output, in response to lower input prices. Therefore it is plausible that the refineries in the Permian Basin increased their crude oil purchases and their refined product output during the periods of low WTI Midland prices. One reason this may not happen is if refineries are already operating at or near the capacity of their facilities. At a national level, this appears to have been the case. Overall refinery capacity utilization in the United States has been above 80 percent in every year since 1985 and is relatively unaffected by oil prices. It was 90.4 percent in 2014 and 89.7 percent in 2016, despite refinery acquisition costs being 56 percent lower in the latter period.[17].

The aggregate trends match the local or refinery-level data for refineries in the Permian Basin (Figure 9). There is no significant relationship between the price discount in the Permian Basin and the refinery capacity utilization in the inland Texas region, based on monthly data from 2007 to 2016 (left graph of Figure 9). The lack of relationship is in spite of an oil price discount exceeding \$20 per barrel in eleven of the months. The same result holds for the two Texas refineries in the Permian Basin, based on monthly refinery-level data for 2014 to 2016 (right graph of Figure 9). There is a slight negative relationship between the price differential and the refinery utilization: lower relative input prices imply lower, not higher, input volumes. There is no evidence that the refineries adjusted their purchases or output in response to the drop in input prices.

does not respond to oil prices, but drilling of new wells has a price elasticity of about 0.7. Newell and Prest (2017) use more recent data that includes shale oil wells. They estimate larger drilling elasticities: approximately 1.2 for conventional and 1.6 for unconventional wells.

[17]Refinery utilization is from the EIA Refinery Utilization and Capacity data, available at `https://www.eia.gov/dnav/pet/pet_pnp_unc_dcu_nus_a.htm`. Refinery input costs are from the EIA Refinery Acquisition Cost of Crude Oil data, available at `https://www.eia.gov/dnav/pet/pet_pri_rac2_dcu_nus_m.htm`

The second possibility is that changes in regional oil prices did not affect refinery profits because they were passed on to consumers through lower or higher gasoline prices. In that case, part of the benefit of the pipeline expansion for oil producers would be a transfer from gasoline consumers.

The oil price differential between LLS and WTI Midland does not influence the gasoline price differential between Midland and Houston (left graph of Figure 10). Over the period 2013 to 2016, the difference in oil prices fell from 75 cents per gallon to almost zero. At the same time, the difference in gasoline prices showed no apparent trend, and the Houston gasoline price was mostly lower than the Midland price. For El Paso, gasoline prices were 25 cents per gallon lower than Houston at the start of 2013 when the oil price differential was highest (right graph of Figure 10). Nevertheless, the subsequent relationship is less clear. In some periods, such as the end of 2014, the oil price declines at the same time as the gasoline price increases.

I estimate Equation (4) using weekly gasoline price data for three cities in Texas: Houston, El Paso, and Midland, with the latter two being in the Permian Basin. The results provide a decomposition of the weekly change in the gasoline prices in the Permian Basin.

$$\Delta P_t^{retail} = \sum_{\tau=0}^{3} \Delta P_{t-\tau}^{LLS} + \Delta P_{t-\tau}^{WTIMidland} + \varepsilon_t \qquad (4)$$

In this equation, ΔP_t^{retail} is the weekly change in the mean of the daily average gasoline prices in the corresponding city, using information from Gas Buddy. ΔP_t^{LLS} is the weekly change in the daily average LLS price. The estimation equation includes three lags of this variable. $\Delta P_t^{WTIMidland}$ and its lags are defined similarly for the WTI Midland price. An additional specification adds the weekly change in the WTI Cushing price and its lags.

Changes in the LLS price are the primary determinant of changes in gasoline prices in all three cities (Table 6). For the three cities, the coefficients on the weekly change in LLS and its first two lags are statistically significantly different from zero. The magnitude of the estimates is similar across the three cities. There is a statistically significant negative coefficient on the change in the WTI Midland price for the Houston and El Paso regressions, though the magnitude is less than that of the LLS terms. For the regressions that also include the WTI Cushing price, the only statistically significant coefficients are those for the change in the LLS price and its lags. The WTI Midland and WTI Cushing coefficients are statistically insignificant and either small in magnitude or negative.

The results demonstrate that retail gasoline prices in Texas did not respond to the geographical variation observed for crude oil prices. Borenstein and Kellogg (2014) reported a similar finding for gasoline prices in the Midwest at a time when WTI Cushing traded at a substantial discount to LLS. They argued that the lack of constraints in refined product pipelines meant that Gulf Coast refineries were setting the marginal price for gasoline. Based on the results in Table 6, the same argument applied for gasoline prices within Texas between 2013 and 2016. This conclusion implies that gasoline consumers were unaffected by oil price changes caused by additions to pipeline capacity.

5 Analysis for other oil-producing regions

Many regions contributed to the 3.9 million barrels per day increase in the United States oil production between 2008 and 2016. Production in the Bakken region increased by nearly 1 million barrels per day over the period (Table 1). The state of Colorado, which shares the Niobrara Region with Kansas, Nebraska, and Wyoming, increased its production by 240,000 barrels per day. The largest increase occurred in the Eagle Ford region of southern Texas, where production rose from 50,000 barrels per day in 2008 to a peak of 1.7 million barrels per day in March 2015.[18] In this section I repeat the analysis in Section 4 using information on prices and pipelines in the Bakken region and Colorado.

Since 2008, midstream companies have constructed three oil pipelines out of the Denver-Julesburg Basin in Colorado. The White Cliffs pipeline connects Platteville, Colorado to Cushing. It was commissioned in 2009 and expanded in 2014 and 2016 to a capacity of 215,000 barrels per day.[19] In 2015, the Northeast Colorado Lateral pipeline connected to the Pony Express pipeline from Guernsey, Wyoming to Cushing, allowing the export of an additional 90,000 barrels per day out of Colorado. In late 2016, the new Saddlehorn pipeline from Platteville to Cushing was commissioned, increasing capacity by a further 190,000 barrels per day.[20].

By the end of 2015, the capacity of the new pipelines exceeded oil production in Colorado (top graph of Figure 12).[21] Between 2013 and 2015, oil production exceeded refinery and

[18] Production data from the EIA Drilling Report. Eagle Ford is not separately analyzed in this paper because there is no public data available about producer prices in the region.

[19] `https://www.semgroupcorp.com/Operations/RoseRock/Transportation/Pipelines/WhiteCliffs.aspx`

[20] `http://okenergytoday.com/2016/09/saddlehorn-pipeline-platteville-cushing-operational/`

[21] The figure shows monthly output for Colorado. Most oil production in the state is in the Denver-Julesburg Basin, in the southern part of the Niobrara Region. Total Niobrara production in December 2016 was 418,000 barrels per day, compared to 294,000 barrels per day for Colorado.

pipeline capacity. The differential between the WTI Cushing benchmark and the Colorado wellhead price was high during this period (bottom graph of Figure 12), creating a $2.80 per barrel discount for producers. The new pipeline projects increased prices for Colorado oil producers, just as shown for Permian producers.

For the Bakken region, production first exceeded the available refining and pipeline capacity in 2010 (top graph of Figure 11). This condition continued through the end of 2016, in spite of the construction of new pipelines and a refinery and the expansion of existing facilities. In 2016, there were three pipeline routes out of the Bakken Region: south to Wyoming, north to Canada, and east to Minnesota.[22] Capacity along all of these routes increased. With the completion of the Butte Loop in 2014 and other projects, maximum volumes on the Butte pipeline system to Wyoming rose from 92,000 barrels per day in 2007 to 260,000 barrels per day in 2016. Double H is a smaller pipeline to Wyoming, commissioned in 2015.[23]. The pipeline company Enbridge increased the capacity on its Line 81 to Minnesota to 214,000 barrels per day. Its new Line 26 to Manitoba, Canada, provided a further 145,000 barrels per day of pipeline capacity. The Dakota Prairie Refinery, with a processing capacity of 20,000 barrels per day, was the first new oil refinery to be built in the United States since the 1970s. All of these projects increased the total Bakken refining and pipeline capacity from 232,000 barrels per day in 2007 to 833,600 barrels per day in 2016.[24]

Unlike the Permian and Colorado results, the price differential for North Dakota oil producers is uncorrelated with excess production (bottom graph of Figure 11). One reason is that excess production has been positive since 2011. There is almost no variation in the data between positive and negative excess production. Transportation options out of the Bakken are relatively more complex. Pipeline and rail alternatives connect to other pipeline systems that may also be congested, and rail can potentially deliver oil to refineries across North America.[25] Compared to the Permian Basin, this makes it difficult to determine the output price and transportation cost for the marginal barrel.

[22] The Dakota Access Pipeline, from North Dakota to Illinois, began operating in 2017.

[23] `https://www.kindermorgan.com/pages/business/products_pipelines/doubleH.aspx`

[24] Information on infrastructure projects in North Dakota is primarily from the North Dakota Pipeline Authority: `https://northdakotapipelines.com/about-us/`.

[25] Covert and Kellogg (2017) develop a model in which oil shippers choose between pipeline and rail to arbitrage price differences between regions. Rail has the advantages of not requiring a long-term commitment and providing the flexibility to serve multiple locations. They calibrate their model using data for the Bakken region to show that lower rail transportation costs reduce investment in pipeline capacity.

6 Conclusion

Large investments in transportation infrastructure have reshaped the North American crude oil pipeline network over the past decade. The reconfigured system provides access to world markets for the near-doubling in oil production since 2008. As quantified in this paper, these new pipelines nearly eliminated the discount to world oil prices that many United States producers faced between 2011 and 2014, leading to substantial improvements in oil producer surplus. Most of the gain came at the expense of oil shippers and oil refiners who benefited from the price differentials due to infrastructure constraints.

This analysis focuses purely on the price effects of pipeline investment. However, there may be changes in negative externalities that affect the overall welfare outcome. Clay et al. (2017) show that the external costs from air pollution and greenhouse gas emissions are almost twice as large for rail as they are for pipeline transportation of crude oil. They also report costs associated with spills and accidents that are six times greater for rail than for pipelines. A complete accounting of the welfare change from pipeline investments would also incorporate this reduction in the externalities from rail transportation.

References

Agerton, Mark and Gregory Upton, "Decomposing Crude Price Differentials: Domestic Shipping Constraints or the Crude Oil Export Ban?," Technical Report 2017.

Anderson, Soren T., Ryan Kellogg, and Stephen W. Salant, "Hotelling Under Pressure," *Journal of Political Economy*, forthcoming.

Barron, Manuel and Maximo Torero, "Household electrification and indoor air pollution," *Journal of Environmental Economics and Management*, 2017.

Borenstein, Severin and Ryan Kellogg, "The Incidence of an Oil Glut: Who Benefits from Cheap Crude Oil in the Midwest?," *Energy Journal*, 2014, *35* (1), 15–34.

Brown, Stephen PA, Charles Mason, Alan Krupnick, and Jan Mares, "Crude Behavior: How Lifting the Export Ban Reduces Gasoline Prices in the United States," 2014.

Burlig, Fiona and Louis Preonas, "Out of the Darkness and Into the Light? Development Effects of Rural Electrification," Technical Report 2016.

Buyuksahin, Bahattin, Thomas K Lee, James T Moser, and Michel A Robe, "Physical markets, paper markets and the WTI-Brent spread," *Energy Journal*, 2013, *34* (3).

Clay, Karen, Akshaya Jha, Nicholas Muller, and Randall Walsh, "The External Costs of Transporting Petroleum Products by Pipelines and Rail: Evidence From Shipments of Crude Oil from North Dakota," Working Paper 23852, National Bureau of Economic Research September 2017.

Covert, Thomas R. and Ryan Kellogg, "Crude by Rail, Option Value, and Pipeline Investment," Working Paper 23855, National Bureau of Economic Research September 2017.

Crooks, Ed, "Work restarts on Dakota Access oil pipeline," *Financial Times*, February 9 2017.

Dinkelman, Taryn, "The effects of rural electrification on employment: New evidence from South Africa," *The American Economic Review*, 2011, *101* (7), 3078–3108.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," 2012.

_ **and Richard Hornbeck**, "Railroads and American economic growth: A "market access" approach," *The Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

Duranton, Gilles and Matthew A Turner, "Urban growth and transportation," *Review of Economic Studies*, 2012, *79* (4), 1407–1440.

_ **, Peter M Morrow, and Matthew A Turner**, "Roads and Trade: Evidence from the US," *Review of Economic Studies*, 2014, *81* (2), 681–724.

EIA, "Drilling Productivity Report May 2017," Technical Report 2017.

Fattouh, Bassam, *An anatomy of the crude oil pricing system* 2011.

Gaswirth, Stephanie B., Kristen R. Marra, Paul G. Lillis, Tracey J. Mercier, Heidi M. Leathers-Miller, Christopher J. Schenk, Timothy R. Klett, Phuong A. Le, Marilyn E. Tennyson, Sarah J. Hawkins, Michael E. Brownfield, Janet K. Pitman, and Thomas M. Finn, "Assessment of undiscovered continuous oil resources in the Wolfcamp shale of the Midland Basin," Technical Report, U.S. Geological Survey Fact Sheet 2016-3092 2016.

Genscape, "Mid-Continent Pipelines Report: Frequently Asked Questions," Technical Report 2015.

_ , "Gulf Coast Pipeline Report," 2016. Data from 2009 to 2016.

_ , "Mid-Continent Pipelines Report," 2016. Data from 2009 to 2016.

Grogan, Louise and Asha Sadanand, "Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua," *World Development*, 2013, *43* (Supplement C), 252 – 265.

Hausman, Catherine and Ryan Kellogg, "Welfare and Distributional Implications of Shale Gas," *Brookings Papers on Economic Activity*, 2015.

Kilian, Lutz, "The impact of the shale oil revolution on US oil and gasoline prices," *Review of Environmental Economics and Policy*, 2016, *10* (2), 185–205.

Lee, Kenneth, Edward Miguel, and Catherine Wolfram, "Experimental Evidence on the Demand for and Costs of Rural electrification," Technical Report, National Bureau of Economic Research 2016.

Leroux, Bailey, "Pipe Dreams," *PB Oil and Gas*, November 2012. `https://pboilandgasmagazine.com/pipe-dreams/`.

Lipscomb, Molly, Ahmed Mobarak, and Tania Barham, "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil," *American Economic Journal: Applied Economics*, 2013, *5* (2), 200–231.

Lynch, David J., Barney Jopson, and Ed Crooks, "Trump ends Obama block on Keystone XL and Dakota Access pipelines," *Financial Times*, January 24 2017.

McRae, Shaun, "Vertical Integration and Price Differentials in the U.S. Crude Oil Market," *Working paper*, 2015.

Melek, Nida Çakır and Elena Ojeda, "Lifting the US Crude Oil Export Ban: Prospects for Increasing Oil Market Efficiency," *Economic Review*, 2017.

Newell, Richard G. and Brian C. Prest, "The Unconventional Oil Supply Boom: Aggregate Price Response from Microdata," Working Paper 23973, National Bureau of Economic Research October 2017.

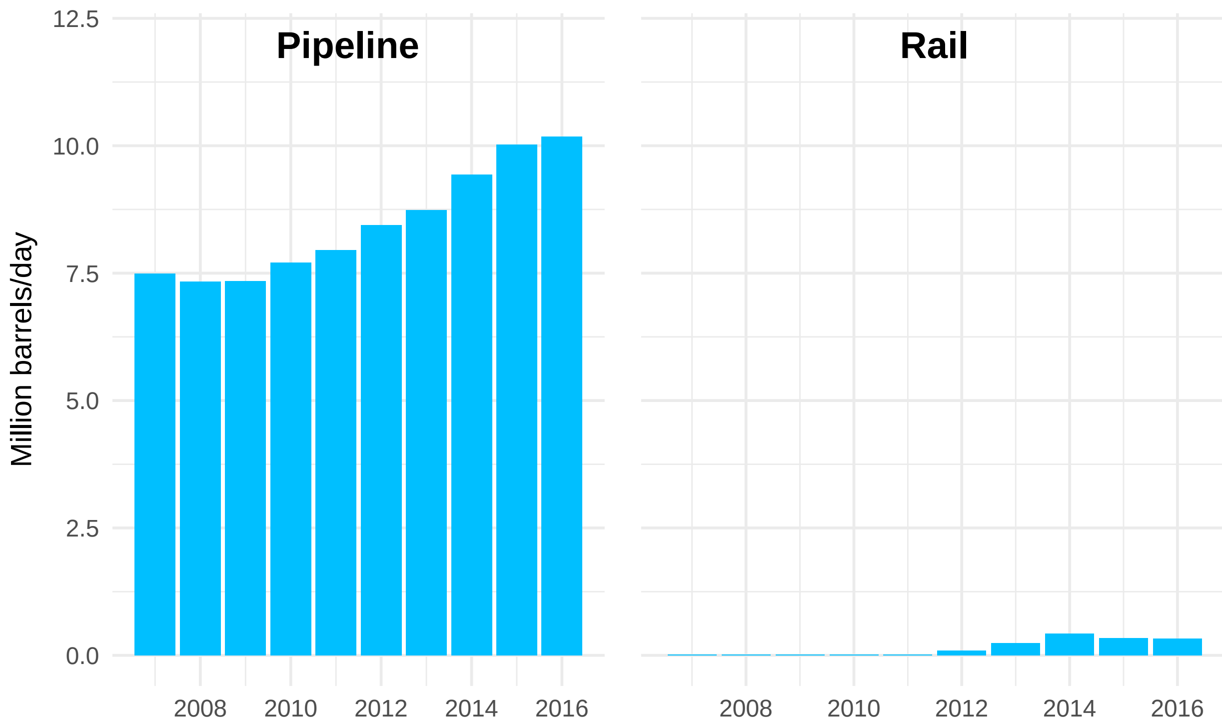
Owen, Russell, "New Boom Town," *New York Times*, April 23 1951, p. 184.

RBN Energy, "With a Permian Well, They Cried More, More, More," Technical Report, A RBN Energy Drill Down Report 2017.

Stevens, William K., "Texas Oilfield Is Revived By the Increase in Prices," *New York Times*, April 15 1978, p. 461.

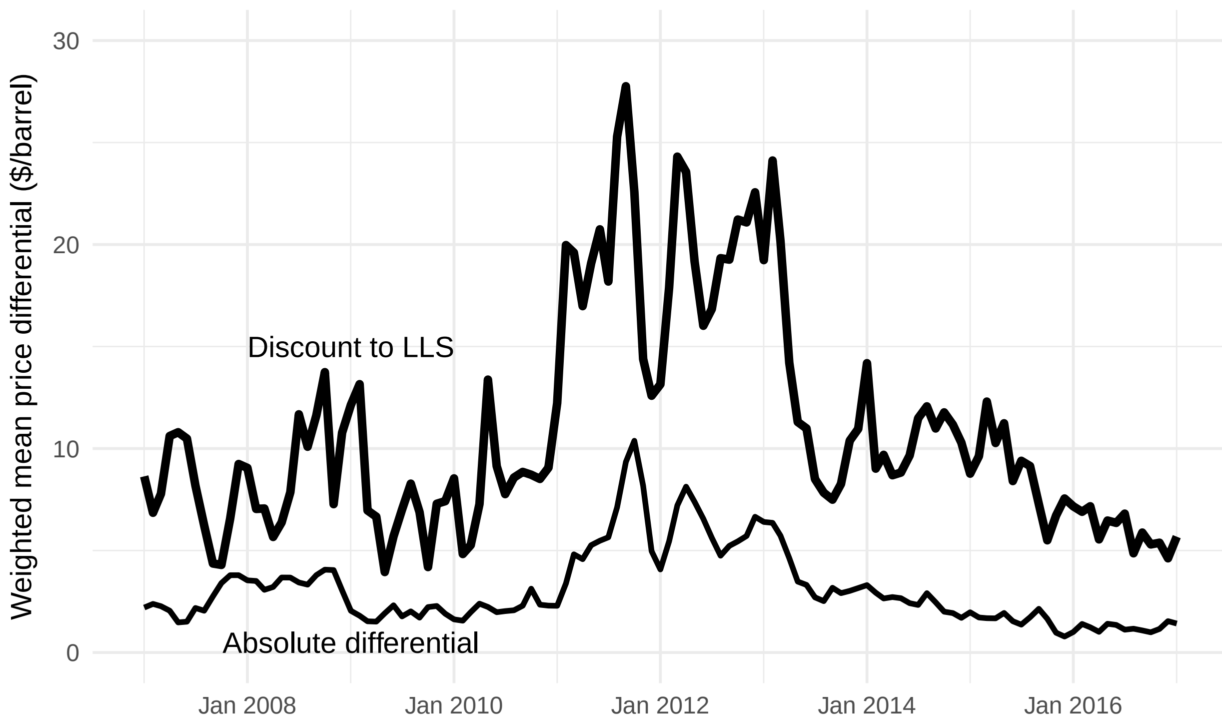
Warren, Jennifer, *The Chronicles of an Oil Boom: Unlocking the Permian Basin*, Concept Elemental, 2014.

Figure 1: U.S. refinery deliveries of crude oil by method of transportation, 2007–16



Notes: Data from the EIA "Refinery Receipts of Crude Oil by Method of Transportation" report, available at `https://www.eia.gov/dnav/pet/pet_pnp_caprec_dcu_nus_a.htm`

Figure 2: U.S. wellhead crude oil price differentials, 2004–16



Notes: State-level wellhead oil price data is from the EIA "Domestic Crude Oil First Purchase Prices by Area" report, available at `https://www.eia.gov/dnav/pet/pet_pri_dfp1_k_m.htm`. LLS price data is from Bloomberg. Absolute wellhead price differential is the mean of the absolute difference between the state-level wellhead price and the national average wellhead price. Discount to LLS is the mean difference between LLS and the state-level wellhead price. Both means are weighted by the oil production in each state.

Figure 3: Stylized model of the welfare effects of oil pipeline expansions

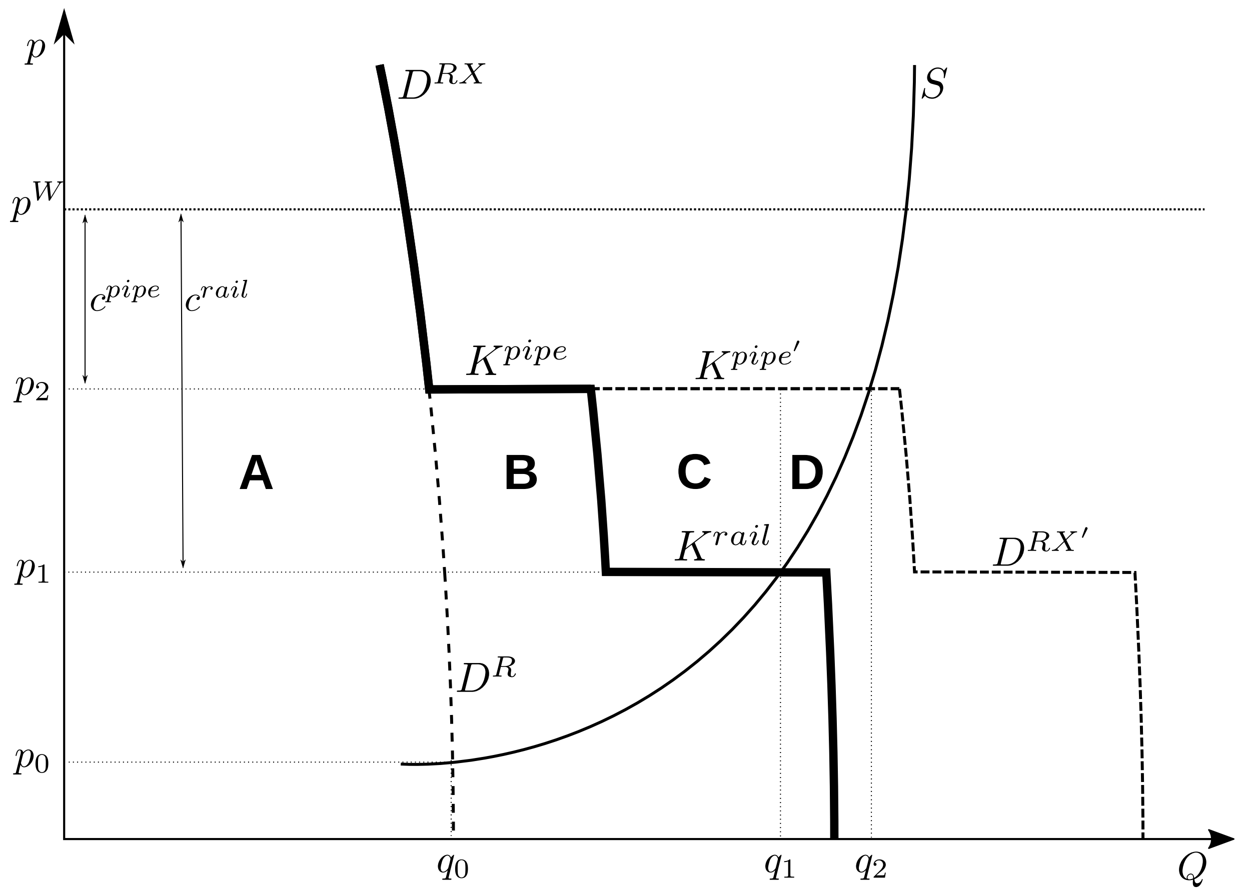


Figure 4: Empirical model of price differentials and infrastructure

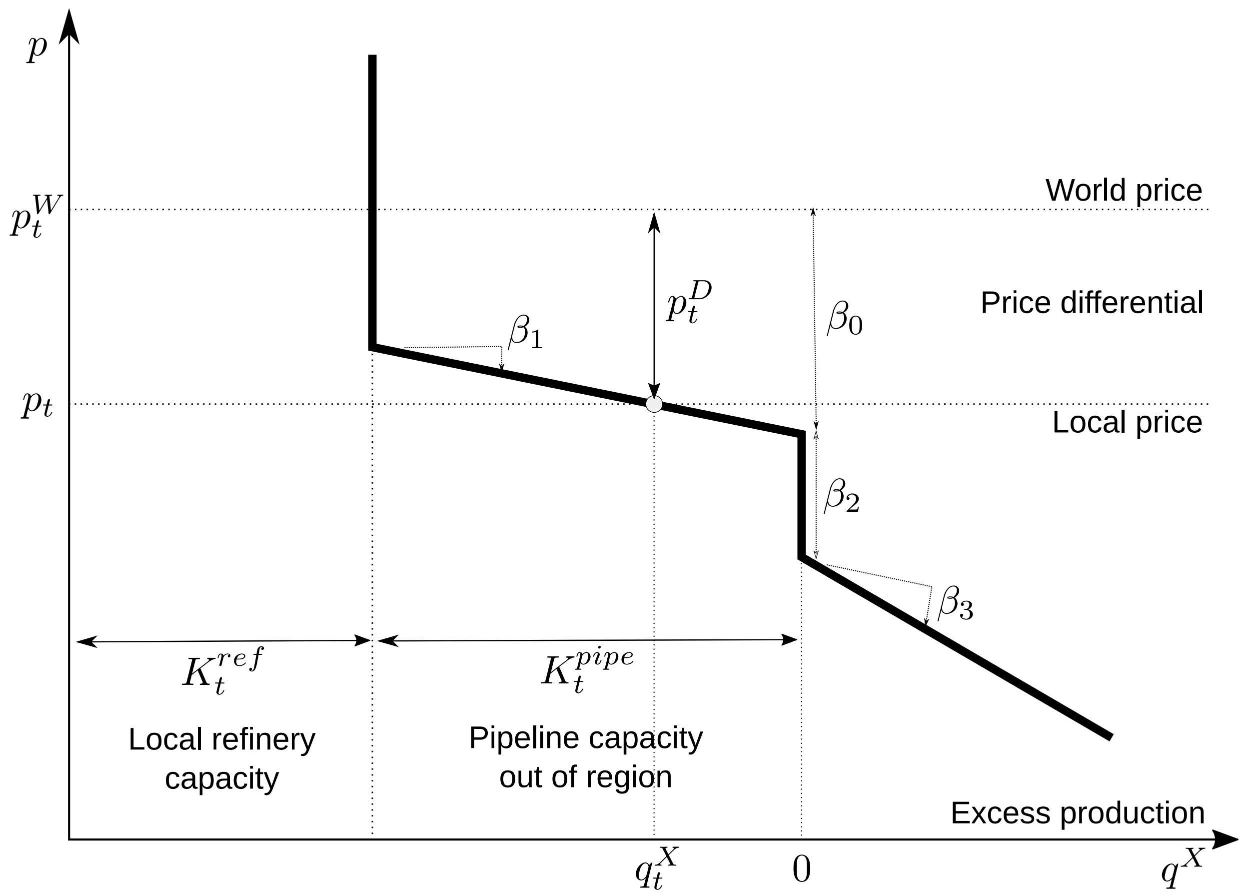
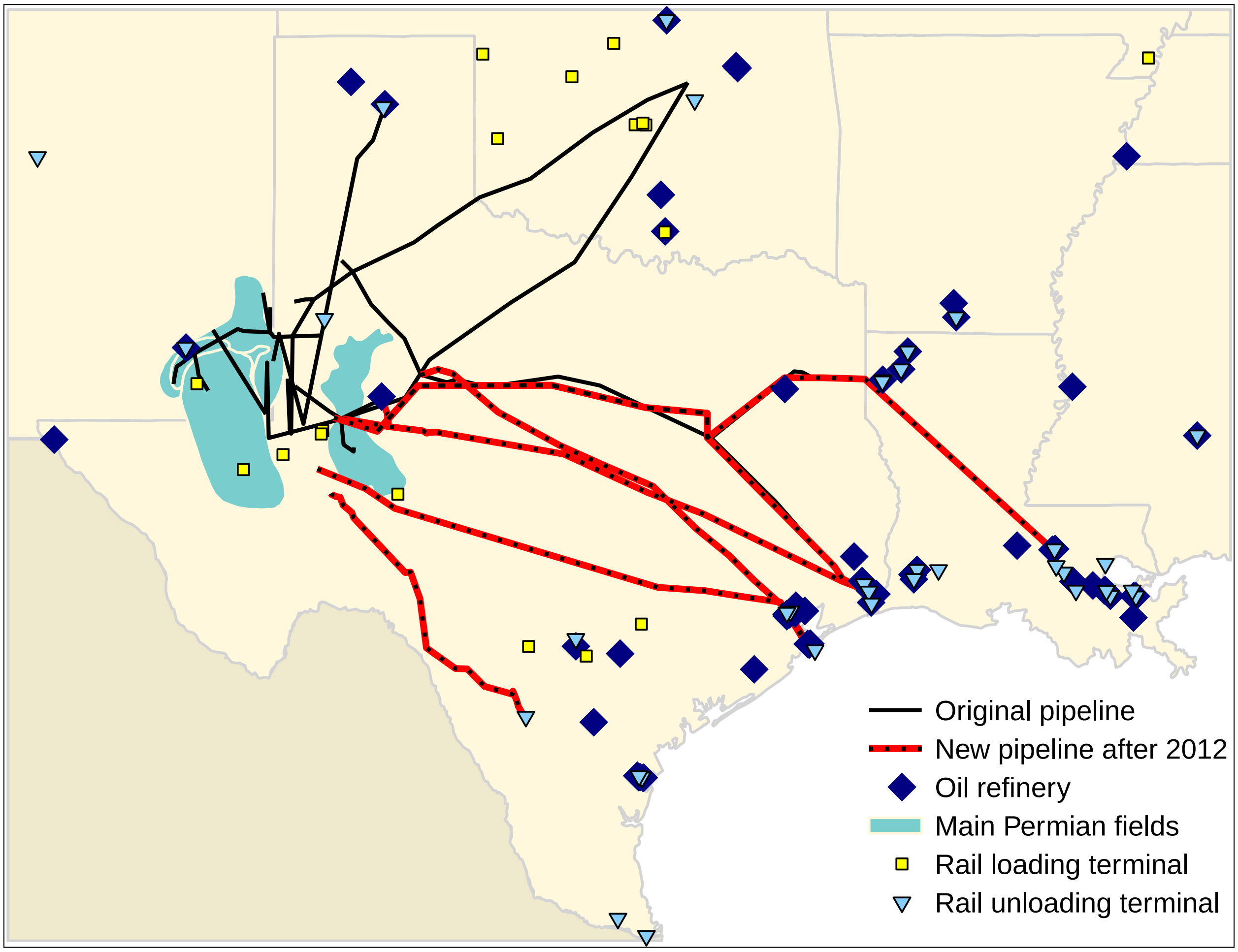


Figure 5: Crude oil refining and transportation infrastructure out of the Permian Basin



Notes: Oil field locations are from the EIA "Tight Oil and Shale Gas Plays" shape file. Refinery locations are from the EIA "Petroleum Refineries" shape file. Most pipeline paths are from the EIA "Crude Oil Pipelines" shape file. Crude oil rail loading and unloading terminals, as at November 2014, are from the EIA "Crude Oil Rail Terminals" shape file. These map files are available at `https://www.eia.gov/maps/layer_info-m.php`. The routes for the Permian Express II and PELA pipelines are not included in the EIA data. Approximate routes were created using OpenStreetMap (`http://www.openstreetmap.org`), based on information from `http://www.sunocologistics.com/Customers/Business-Lines/Asset-Map/241/`.

Figure 6: Crude oil production and combined refinery and pipeline capacity in the Permian Basin

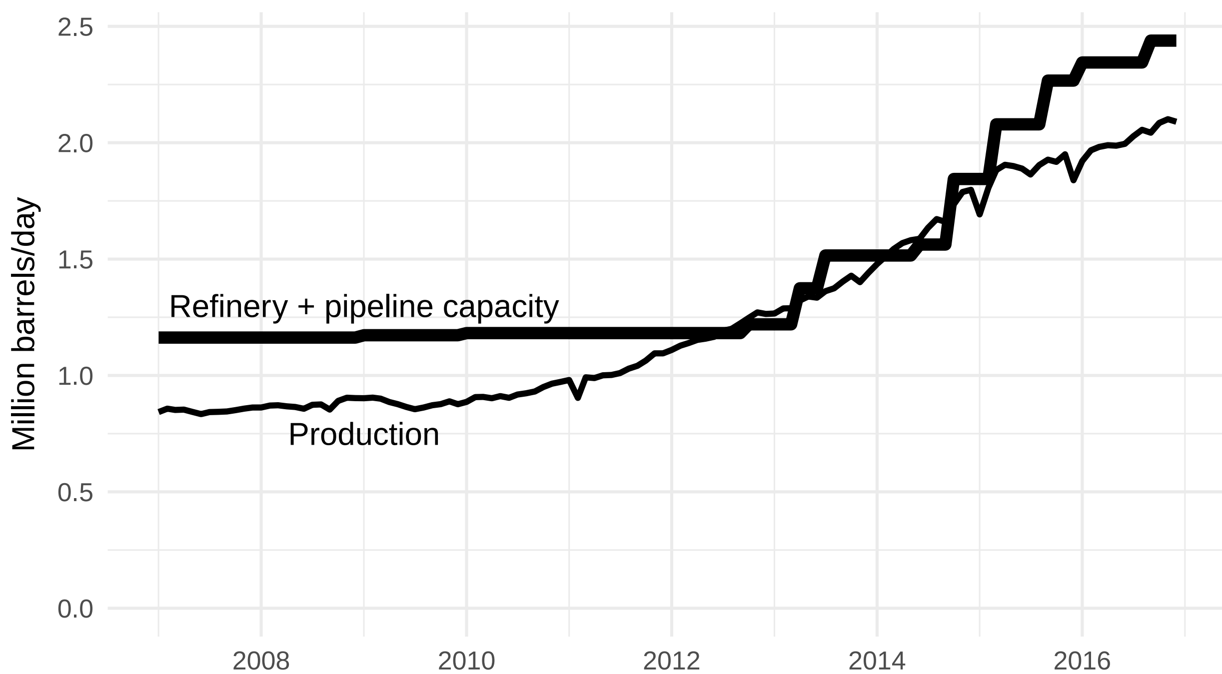


Figure 7: Excess supply, price differentials, and rail volumes for the Permian Basin

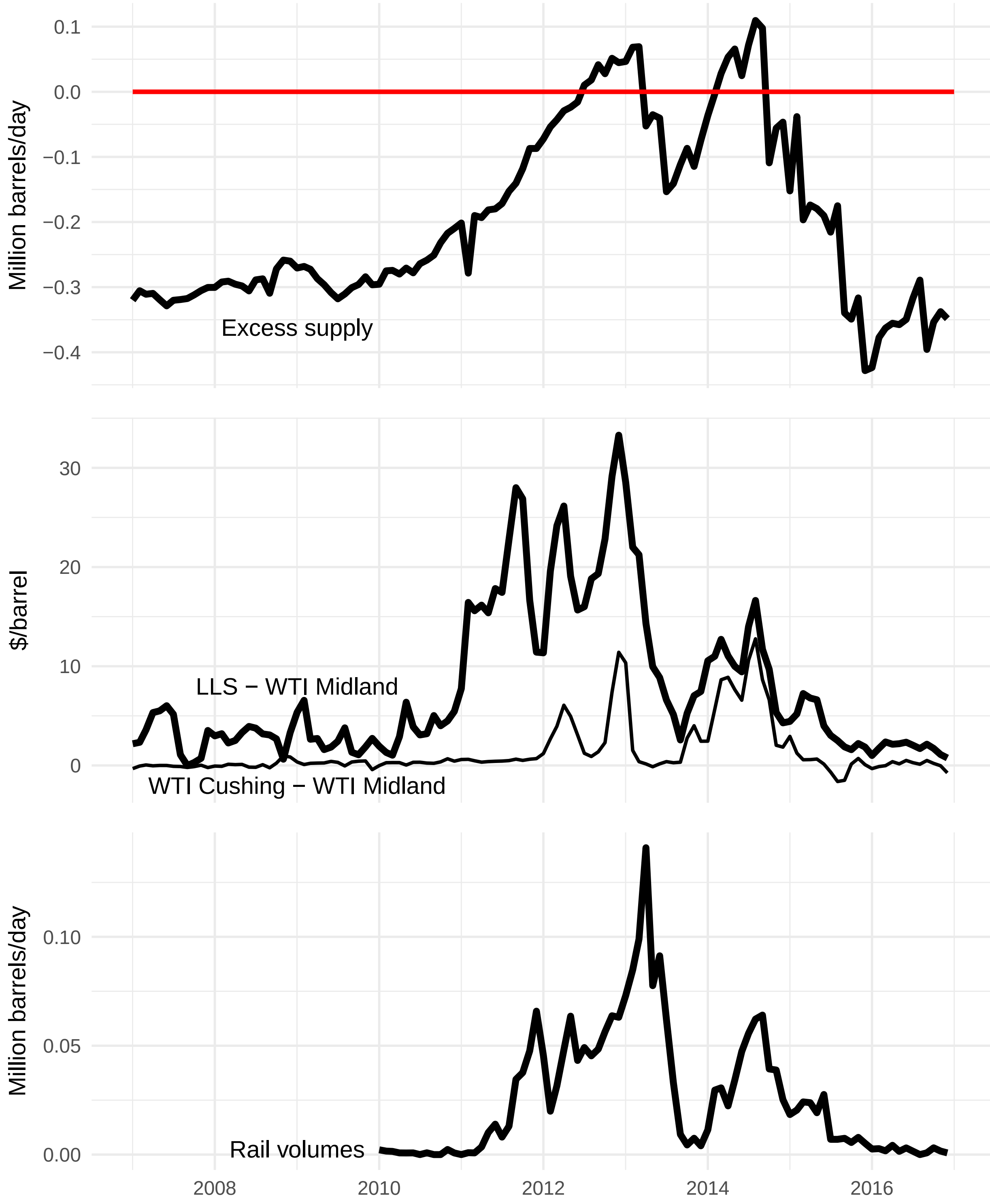
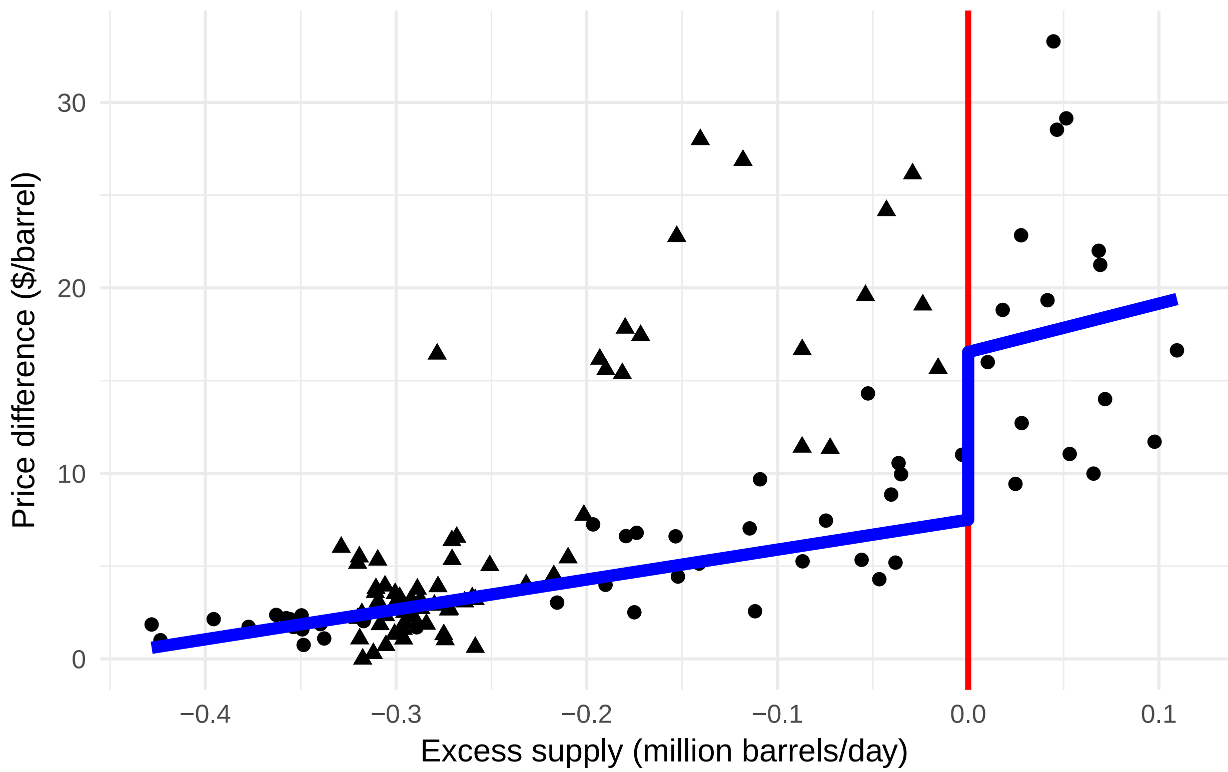
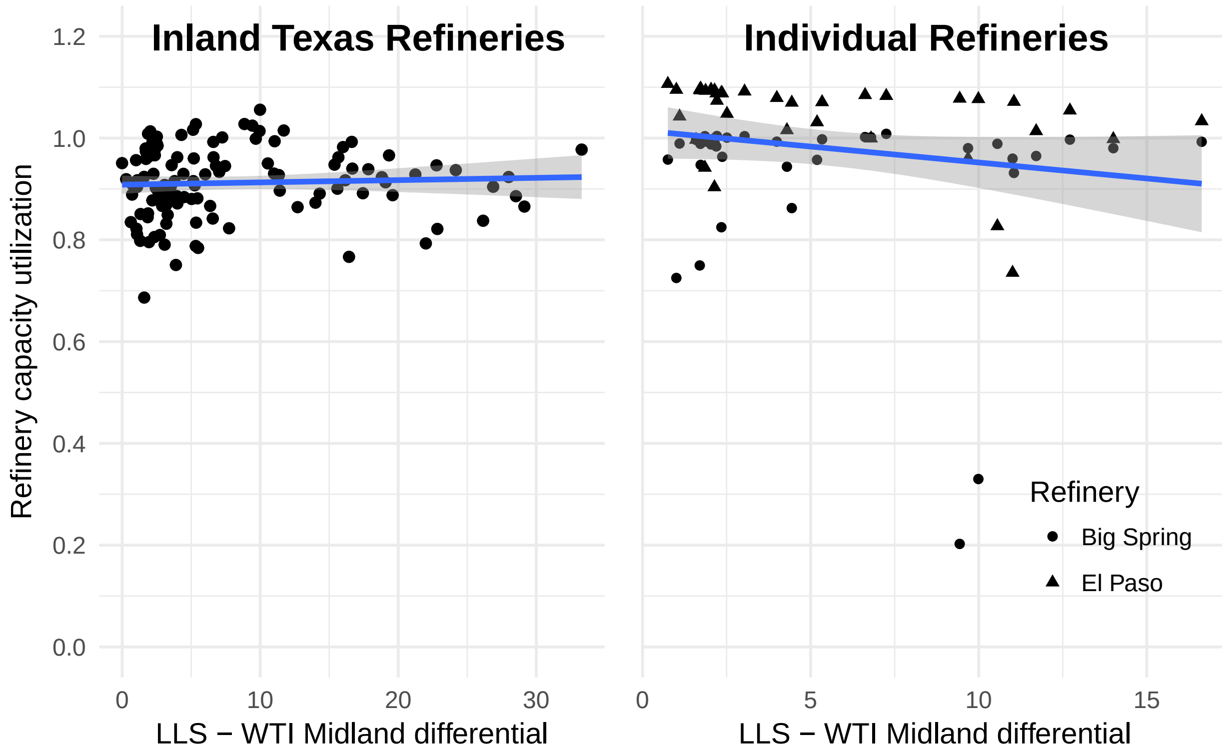


Figure 8: Excess supply and price differentials for the Permian Basin



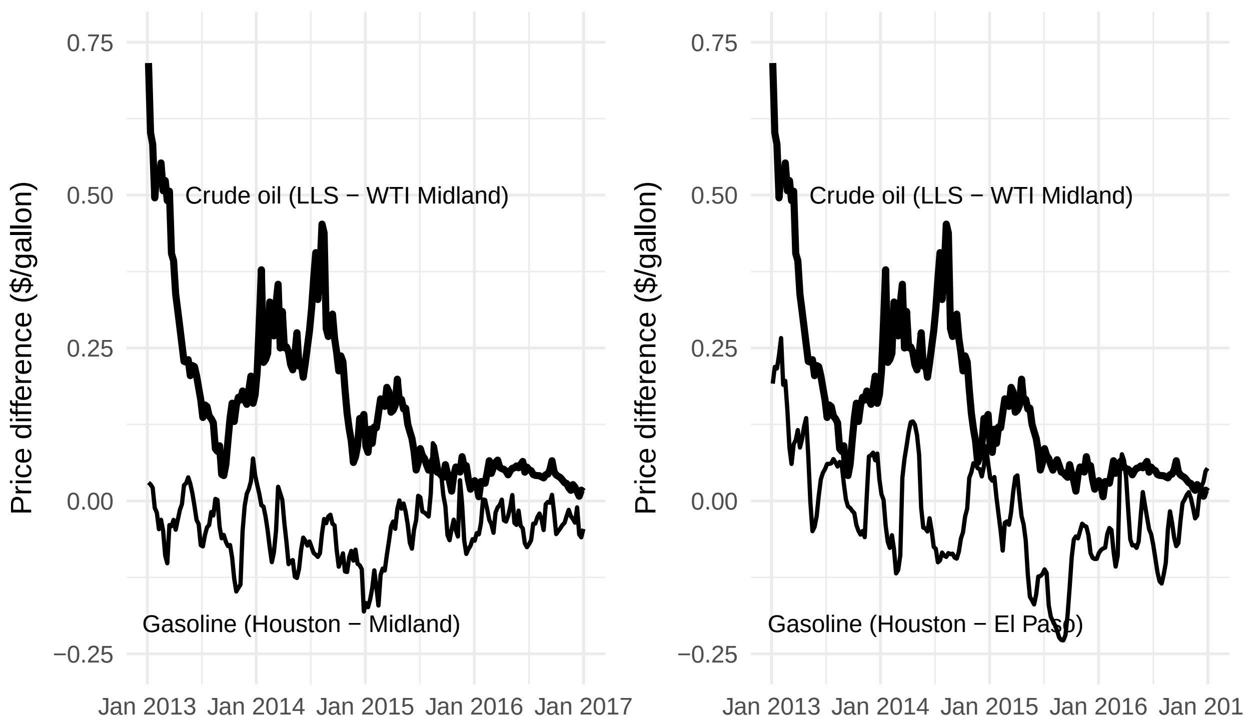
Notes: The scatter graph shows the 120 observations of q_t^X and p_t^D used for the regressions in Table 4. Observations with a triangle are those for which the marginal barrel of oil production is sent by pipeline to Cushing (and so the LLS–WTI Midland price differentials need to be adjusted by the WTI Cushing–LLS differential). The fitted line is based on Column 3 of 4, assuming the year is 2016 and the WTI Cushing–LLS differential is zero.

Figure 9: Refinery capacity utilization and price differentials



Notes: The left graph shows the monthly mean capacity utilization of the refineries in the Inland Texas refining district, plotted against the LLS to WTI Midland price differential. Refinery data is from the monthly EIA Refinery Utilization and Capacity report (`https://www.eia.gov/dnav/pet/pet_pnp_unc_dcu_nus_m.htm`). The right graph shows the monthly capacity utilization for the two Texas refineries in the Permian Basin (Big Spring and El Paso), for the period 2014 to 2016. Individual refinery data is from the refinery Monthly Report and Operations Statements, made available by the Railroad Commission of Texas (`http://www.rrc.texas.gov/oil-gas/research-and-statistics/refinery-statements/`).

Figure 10: Price differentials for crude oil and retail gasoline in Midland and El Paso



Notes: Both graphs show the weekly mean difference between the WTI Midland and LLS oil prices, in dollars per gallon. The left graph shows the weekly mean difference between the daily mean gasoline price in Midland and the daily mean gasoline price in Houston. The right graph shows this comparison for El Paso and Houston. Retail gasoline prices are from Gas Buddy.

Figure 11: Production, infrastructure, and price differentials for the Bakken region

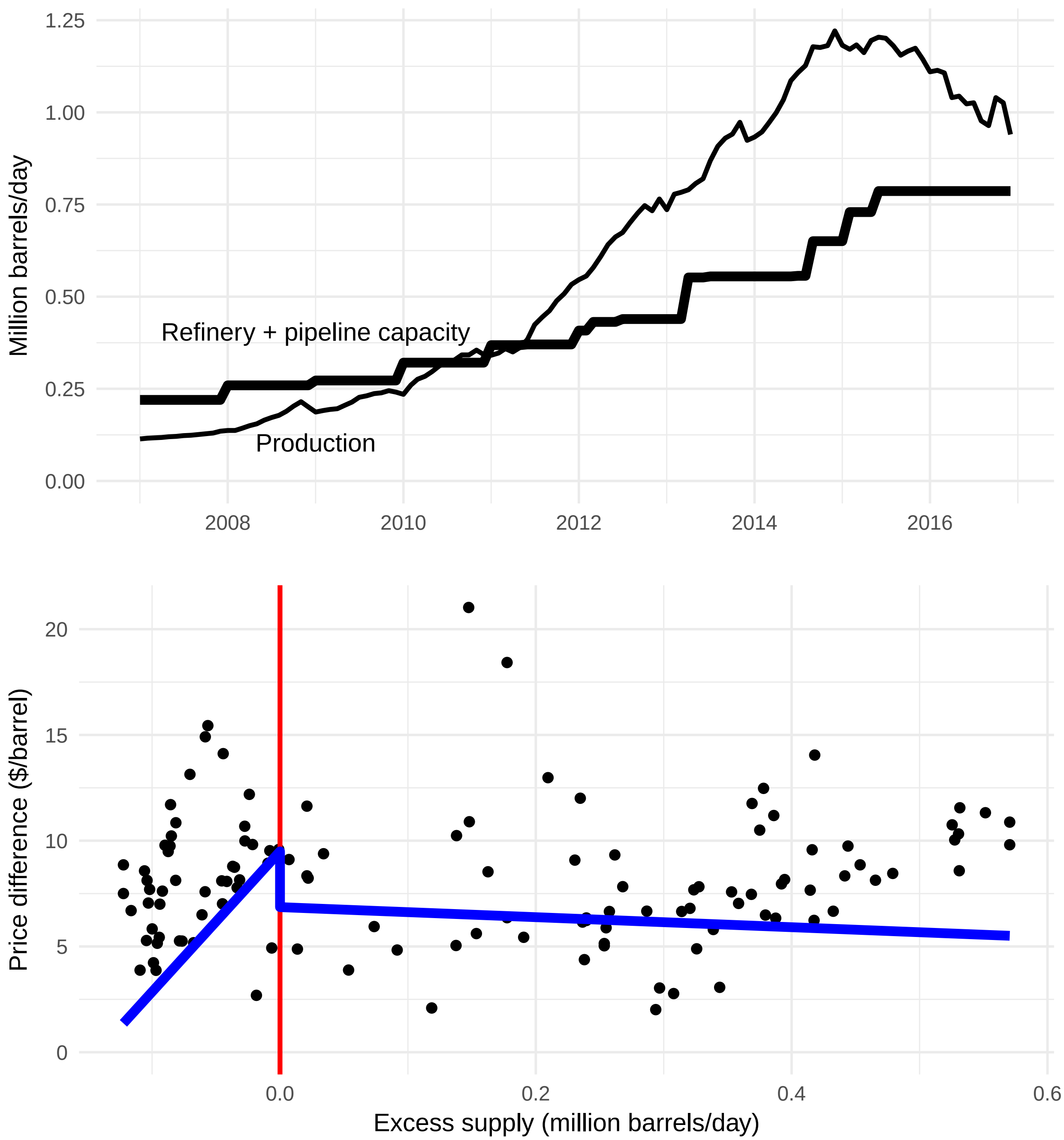


Figure 12: Production, infrastructure, and price differentials for Colorado

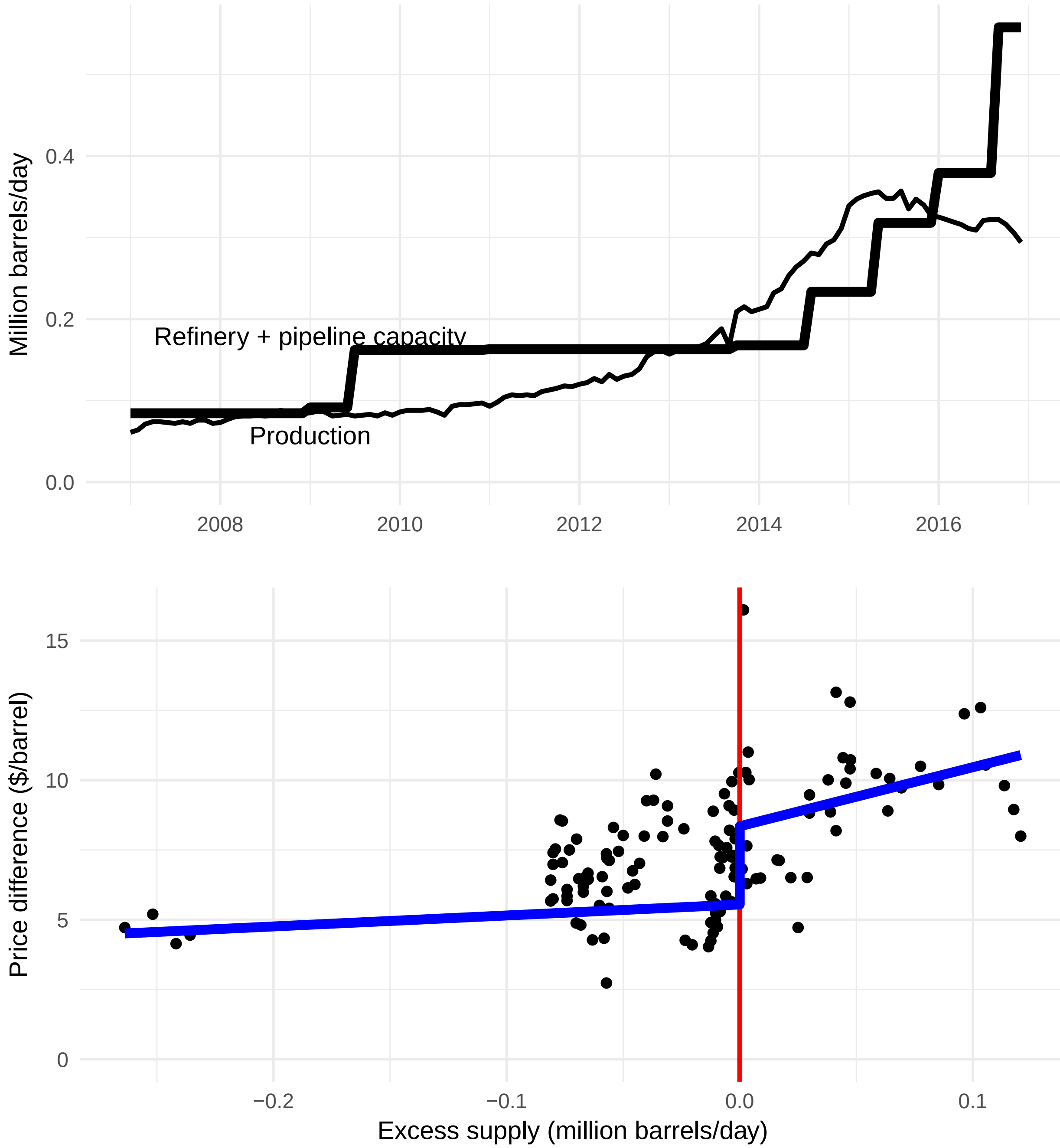


Table 1: Summary statistics about U.S. crude oil production, infrastructure, and prices, 2007–16

Variable	2008	2010	2012	2014	2016	Mean
U.S. crude oil production (mbd)	5.00	5.48	6.49	8.76	8.88	6.76
Permian basin (mbd)	0.87	0.92	1.19	1.63	2.02	1.26
Bakken (mbd)	0.18	0.32	0.69	1.12	1.06	0.63
Colorado (mbd)	0.08	0.09	0.14	0.26	0.32	0.17
Eagle Ford region (mbd)	0.06	0.08	0.63	1.45	1.25	0.65
Total pipeline length (000 miles)	50.96	54.63	57.46	66.81	75.74	59.82
Interstate pipeline length (000 miles)		37.94	39.48	47.00	51.93	43.92
Intrastate pipeline length (000 miles)		16.69	17.98	19.81	23.80	19.66
Refinery deliveries by pipeline (mbd)	7.34	7.71	8.44	9.43	10.17	8.46
Refinery deliveries by rail (mbd)	0.01	0.01	0.09	0.43	0.33	0.15
Refinery deliveries by truck (mbd)	0.19	0.20	0.36	0.42	0.45	0.31
Rail shipments within U.S. (mbd)		0.06	0.38	0.87	0.39	0.47
U.S. wellhead price ($/barrel)	94.22	74.64	94.63	87.71	38.37	74.84
WTI Cushing price ($/barrel)	100.33	79.35	94.21	93.67	43.01	78.52
LLS price ($/barrel)	103.09	82.62	111.62	97.34	44.70	85.02
Mean abs. price deviation ($/barrel)	3.53	2.16	6.02	2.50	1.21	3.13
Mean discount from LLS ($/barrel)	8.87	7.97	16.99	9.63	6.33	10.18

Notes: The table shows annual means for every second year between 2008 and 2016, as well as the overall mean for the 10-year period from 2007 to 2016. Crude oil production is from the EIA "Crude Oil Production" and Drilling Productivity reports. Crude oil pipeline length data is from the Pipeline and Hazardous Materials Safety Administration annual report data from hazardous liquids operators (available at `https://www.phmsa.dot.gov/pipeline/library/data-stats/distribution-transmission-and-gathering-lng-and-liquid-annual-data`). Refinery delivery data is from the EIA "Refinery Receipts of Crude Oil by Method of Transportation" report. Domestic rail shipments of crude oil are from the EIA "Movements of Crude Oil and Selected Products by Rail" report. U.S. wellhead prices are from the EIA "Domestic Crude Oil First Purchase Prices by Area" report. WTI Cushing and LLS prices are from Bloomberg.

mbd = millions of barrels per day.

Table 2: Crude oil pipelines out of the Permian Basin

Pipeline name	Start date	Destination	Length (miles)	Capacity (barrels/day)[1]
Original pipelines				
Basin		Cushing, OK	530	450,000
Centurion[2]		Cushing, OK	2,700	140,000
Phillips[3]		Borger, TX	289	132,000
West Texas Gulf[4]		Goodrich, TX	580	340,000
New pipelines since 2012				
Amdel[5]	Sep 2012	Nederland, TX	503	40,000
Longhorn[6]	Apr 2013	Houston, TX	450	275,000
Bridgetex	Sep 2014	Houston, TX	400	300,000
Cactus	Feb 2015	Gardendale, TX	310	330,000
Permian Express II	Aug 2015	Nederland, TX	334	200,000
PELA[7]	Aug 2016	Baton Rouge, LA		100,000

[1] Unless otherwise stated, length and capacity data are from RBN Energy (2017).

[2] Centurion pipeline length includes oil gathering pipelines in both the Texas and New Mexico parts of the Permian Basin

[3] Information from `http://www.phillips66partners.com/EN/Pages/borger-crude-assets.aspx`. Capacity is for the combined Line 80 and Line WA pipelines.

[4] The capacity of the West Texas Gulf pipeline was increased from 250,000 to 340,000 barrels per day in 2013.

[5] Amdel was an existing pipeline that was reversed in 2012.

[6] The Longhorn pipeline was built in the 1950s to transport oil from the Permian Basin to the Gulf Coast. It was reversed to 2005 and converted to carry refined products from Houston to El Paso. In 2013 it was reversed again and converted back to its original purpose. Its capacity was expanded to 275,000 barrels/day in mid-2014 (Leroux, 2012).

[7] The PELA pipeline project included a combination of new pipeline segments and the reversal of existing pipelines.

Table 3: Summary statistics about Permian Basin crude oil production, infrastructure, and prices, 2007–16

Variable	2008	2010	2012	2014	2016	Mean
Oil production (mbd)	0.87	0.92	1.19	1.63	2.02	1.26
Local refining capacity (mbd)	0.27	0.29	0.29	0.29	0.30	0.29
Adjusted	0.25	0.27	0.27	0.27	0.27	0.26
Pipeline export capacity (mbd)	0.97	0.97	1.01	1.68	2.31	1.33
Adjusted	0.91	0.91	0.95	1.58	2.17	1.25
Intra-PADD 3 rail shipments (mbd)		0.00	0.05	0.04	0.00	0.03
LLS ($/barrel)	103.09	82.62	111.62	97.34	44.70	85.02
WTI Cushing ($/barrel)	100.33	79.35	94.21	93.67	43.01	78.52
WTI Midland ($/barrel)	100.19	79.06	90.33	86.81	42.93	77.11
LLS - WTI Midland ($/barrel)	2.90	3.56	21.29	10.53	1.77	7.91
q_t^X (mbd)	-0.29	-0.26	-0.00	0.02	-0.36	-0.19
$I(q_t^X > 0)$	0.00	0.00	0.50	0.58	0.00	0.13
V_{marg}	1.00	1.00	0.50	0.00	0.00	0.55

Notes: Values for refining and pipeline capacities are as at December. All other variables are the means over the corresponding year, with the final column showing the means over the 10-year period from 2007 to 2016. mbd = millions of barrels per day.

Table 4: Estimation of price differentials for Permian Basin (WTI Midland)

	No β_4	Base	Year FE	$\beta_4 = -1$	Unadj K_t^{ref}
	(1)	(2)	(3)	(4)	(5)
q_t^X	39.49***	23.34***	16.12***	13.43***	15.21**
	(12.95)	(4.00)	(5.42)	(4.83)	(6.02)
$I(q_t^X > 0)$	4.59	11.56**	9.04**	12.61***	13.09**
	(6.18)	(5.69)	(3.91)	(4.42)	(6.06)
$I(q_t^X > 0)\, q_t^X$	−68.19**	−52.04*	9.99	−17.43	−73.73
	(30.92)	(27.06)	(25.38)	(22.88)	(56.04)
$I(q_t^X < 0\,\&\, V_{marg})(p_t^W - p_t^V)$		−0.76***	−0.68***		−0.65***
		(0.04)	(0.15)		(0.20)
Constant	15.44***	8.47***	5.96***	4.13***	6.10**
	(4.31)	(1.02)	(2.04)	(1.52)	(2.50)
Year fixed effects	*N*	*N*	*Y*	*Y*	*Y*
Observations	120	120	120	120	120
Adj. R^2	0.543	0.841	0.883	0.829	0.889

Notes: Each observation is one month, with data covering the period January 2007 to December 2016. The dependent variable in all equations is the difference between the LLS and WTI Midland prices, in $ per barrel. The variable q_t^X is the difference between oil production and refinery and pipeline capacities, in millions of barrels per day. V_{marg} is equal to 1 when marginal pipeline flows are delivered to Cushing, assumed to be before March 2013. In equation 4, the coefficient on the fourth line is set equal to -1 (implying a complete passthrough of the WTI Cushing to LLS differential when Cushing pipelines are marginal). In equation 5, refinery capacities are unadjusted barrels per calendar day. Standard errors are Newey-West with 12 lags.
*p<0.1; **p<0.05; ***p<0.01

Table 5: Welfare decomposition for hypothetical new oil pipeline out of the Permian Basin, as of September 2014

$ million/day	(2)	(3)	(4)	(5)
Local oil price increase ($/barrel)	6.36	10.04	10.87	5.71
Increase in oil producer revenue (A+B+C+D)	10.95	17.65	19.19	9.81
Reduction in oil refiner profits (A)	1.71	2.70	2.92	1.54
Reduction in oil shipper profits (B)	7.94	13.01	13.94	7.01
Reduction in oil transportation costs (C)	0.90	0.95	1.17	0.94
Profit from higher oil production (D)	0.40	0.99	1.15	0.32
Transfer % of higher producer profit	88.21	89.01	87.87	87.16

Notes: The results in the table show a welfare decomposition for the effect of a hypothetical new pipeline out of the Permian Basin, with a capacity of 100,000 barrels per day, placed in service in September 2014. Each column shows the decomposition based on the estimates for the corresponding columns in Table 4. The first row shows the predicted increase in the WTI Midland price due to the pipeline. The next rows show the effects on producer, refiner, and shipper profits, with the letters corresponding to the areas in Figure 3. Profits from higher oil production are based on an assumed supply elasticity of 1.0. The final row shows the share of the higher oil producer profits that correspond to transfers from oil refiners and shipper.

Table 6: Estimation of retail gasoline price changes on benchmark oil prices

	ΔHouston	ΔEl Paso	ΔMidland	ΔHouston	ΔEl Paso	ΔMidland
	(1)	(2)	(3)	(4)	(5)	(6)
ΔLLS_t	0.26***	0.31***	0.24***	0.34***	0.43***	0.32***
	(0.07)	(0.07)	(0.07)	(0.13)	(0.14)	(0.10)
ΔLLS_{t-1}	0.31***	0.37***	0.34***	0.31**	0.46***	0.32**
	(0.06)	(0.06)	(0.06)	(0.12)	(0.13)	(0.13)
ΔLLS_{t-2}	0.22***	0.26***	0.20**	0.26	0.34**	0.21
	(0.07)	(0.06)	(0.08)	(0.19)	(0.14)	(0.18)
ΔLLS_{t-3}	0.05	0.07	0.07	0.04	0.15	0.09
	(0.11)	(0.10)	(0.11)	(0.09)	(0.11)	(0.09)
ΔWTI Midland_t	−0.12*	−0.18***	−0.06	−0.05	−0.07	−0.01
	(0.06)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)
ΔWTI Midland_{t-1}	0.03	−0.03	0.04	0.0004	0.03	0.01
	(0.06)	(0.04)	(0.06)	(0.10)	(0.08)	(0.11)
ΔWTI Midland_{t-2}	0.06	0.04	0.03	0.09	0.09	0.05
	(0.06)	(0.06)	(0.08)	(0.10)	(0.10)	(0.15)
ΔWTI Midland_{t-3}	0.13	0.09	0.11	0.10	0.13	0.12
	(0.11)	(0.12)	(0.12)	(0.13)	(0.18)	(0.15)
ΔWTI Cushing_t				−0.17	−0.25	−0.14
				(0.17)	(0.18)	(0.13)
ΔWTI Cushing_{t-1}				0.05	−0.15	0.07
				(0.18)	(0.19)	(0.21)
ΔWTI Cushing_{t-2}				−0.07	−0.13	−0.02
				(0.25)	(0.22)	(0.28)
ΔWTI Cushing_{t-3}				0.04	−0.12	−0.03
				(0.14)	(0.24)	(0.18)
Observations	209	209	209	209	209	209

Notes: Each observation is one week, with data covering the period January 2013 to December 2016. The dependent variable in each equation is the change in the weekly mean gasoline price in each city. Covariates are the change in the weekly mean benchmark crude oil price (in $/gallon) and up to three lags. Standard errors are Newey-West with 12 lags.

*p<0.1; **p<0.05; ***p<0.01