ENVIRONMENTAL CONSEQUENCES OF
HYDROCARBON INFRASTRUCTURE POLICY

Thomas R. Covert
Ryan Kellogg

ABSTRACT

We study policies that aim to "keep carbon in the ground" by blocking fossil fuel infrastructure investment. Our analysis relies on a model of hydrocarbon production and transportation, incorporating substitution between pipeline infrastructure and flexible alternatives, like crude-by-rail. We apply the model to the Dakota Access Pipeline (DAPL), which moves oil from North Dakota to Texas and was controversially completed in 2017. Had DAPL's construction been enjoined, we estimate that 81% of the blocked pipeline flows would move by rail instead. This substitution induces both private costs and local environmental damage, since rail transport imposes greater local externalities than pipelines.

Thomas R. Covert
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
thomas.covert@chicagobooth.edu

Ryan Kellogg
University of Chicago
Harris School of Public Policy
1307 East 60th Street
Chicago, IL 60637
and NBER
kelloggr@uchicago.edu

1 Introduction

For more than a decade, environmental advocates and climate-focused regulators have argued for restrictions on fossil fuel development. Some "keep it in the ground" proposals would prevent the use of, or raise the price of, public lands in new fossil fuel development (Prest and Stock, 2023). Others would have governments or NGO's directly acquire and sequester fossil resources *in-situ*, as explored in Harstad (2012). In parallel, a growing number of investment funds have tried to raise the cost of capital for fossil fuel extraction through ESG efforts (Giglio, Maggiori, Stroebel, Tan, Utkus and Xu, 2023). These strategies all have limitations, however: U.S. public lands contain only a small fraction of undeveloped fossil fuels, ESG policies may reduce investment funds' returns, and it is hard for buyers of fossil fuel resources who plan to never develop them to outbid those who would develop.

An alternative "keep it in the ground" strategy is to block the construction of specialized fossil fuel transportation infrastructure, so that the fuels that would have been transported might instead never be developed. Infrastructure foreclosure policies have appeal because they can arguably stop fossil fuel production, even on private land, without the need to pay off investors or mineral rights holders. Activism for, and in some cases enactment of, such policies has become commonplace. Four major U.S. oil and gas pipeline projects—the Dakota Access Pipeline, Mountain Valley Pipeline, Atlantic Coast Pipeline, and Keystone XL—have been the subject of vociferous opposition, and it is unlikely that the latter two will ever be built (Tabuchi and Plumer, 2020). The U.S. Federal Energy Regulatory Commission (FERC) has been debating whether its natural gas pipeline permitting procedures should account for CO_2 emissions from the new gas production that each pipeline might induce (Wilson, 2022). Similarly, new coal export terminals on the U.S. West Coast have faced local opposition (McClure, 2021). These debates are not confined to the United States. Internationally, the East African Crude Oil Pipeline, Eastern Mediterranean Pipeline, and Canadian Trans Mountain Pipeline projects have all been opposed over their potential climate impacts (Dahir, 2023; Dalton, 2020; CBC, 2019).

Our goal in this paper is to improve our understanding of the economics of, and the trade-offs induced by, the foreclosure of fossil fuel infrastructure. Our analysis can be seen as the flip side of recent work that has explored how investments in electric transmission can reduce carbon emissions by inducing the production of renewable power (Fell, Kaffine and Novan, 2021; Davis, Hausman and Rose, 2023; Gonzales, Ito and Reguant, 2023). Here, we are instead interested in how *preventing* construction of infrastructure designed to transport fossil fuels might reduce carbon emissions by inhibiting production of those fuels.

The core trade-off in our analysis arises from the availability of alternative transportation

technologies. In the absence of a pipeline, crude oil can move over land via railroad or over water via ship, natural gas can be transported over water by liquefied natural gas carriers and potentially even over land by rail (Funk, 2023), and propane can move both by ship and by land on specialized vehicles. Policies that block oil and gas pipelines potentially induce increased usage of these alternatives, all of which have their own environmental externalities. Inspired by this dilemma, we develop and quantify a model of fossil fuel production and transportation mode choice, informing the policy question of whether blocking pipelines keeps resources in the ground or merely shifts their transport to other modes, and quantifying the net change in externalities.

We apply this model to the case of the Dakota Access Pipeline (DAPL), which was completed in June, 2017 and moves more than 500,000 barrels per day (500 mbbl/d) of oil from the Bakken Shale of North Dakota to the U.S. Gulf Coast. DAPL's construction faced substantial opposition from activists and policy-makers, and its operation remains under litigation today.[1] At the same time, railroads have played an important role in transporting Bakken oil production, with volumes peaking above 700 mbbl/d in 2015. Crude-by-rail has well-known safety externalities in the form of train derailments, such as the Lac-Megantic disaster in 2013 that killed 47 people, and it also generates local air pollution (Clay, Jha, Muller and Walsh, 2019). In this paper, we therefore use our model to ask: (1) had DAPL not been constructed, how much of the pipeline's oil flows would have stayed in the ground versus continued to flow, but by railroad instead of pipe; and (2) what would have been the environmental and economic surplus consequences of these counterfactual oil flows?

We begin in section 2 with a stylized model of crude oil production and transportation that develops intuition. Oil can flow through either fixed infrastructure (i.e., a pipeline) that has high up-front sunk costs and no ongoing marginal costs, or through an alternative that is more flexible but involves substantial ongoing marginal costs (i.e., railroads). In this environment, a continuum of oil shippers decides whether to use pipelines or rail to physically arbitrage price differences between a price-sensitive upstream supply source (North Dakota) and downstream markets (the Gulf Coast).[2] Given a fixed pipeline capacity, in equilibrium shippers will only use crude-by-rail when downstream oil prices are sufficiently high that the pipeline is operating at full capacity. At such times, the demand for rail shipping will drive a wedge between upstream and downstream oil prices that is equal to rail's marginal cost. This wedge depresses upstream production relative to a case in which more pipeline capacity were available and crude-by-rail shipping was not needed; it is by this mechanism

---

[1]In an analysis related to ours, ICF (2020) evaluates the potential oil production, price, and employment impacts of an abrupt, court-ordered 16-month DAPL shutdown.

[2]Throughout this paper, we follow transportation industry terminology by referring to pipeline and rail customers as "shippers". The pipelines and railroads themselves are known as *carriers*, not shippers.

that blocking pipeline construction can keep some oil in the ground.

We also use our model to endogenize pipeline capacity investment, capturing the institutional fact that pipeline shippers must make long-term capacity commitments to the pipeline before construction and before future oil prices are realized. We show that the equilibrium pipeline capacity balances the pipeline's tariff against the margin between the upstream and downstream oil prices that shippers expect to realize during the commitment period.

In section 3, we quantitatively apply our model to Bakken oil transportation. First, we estimate crude-by-rail costs using the history of crude-by-rail flows and price differentials from the Bakken to downstream markets. Here, we enhance our model by allowing for adjustment costs that dampen the response of rail flows to changes in price differentials, which better matches the data and aligns with institutional features of rail transport. Second, we estimate a model of Bakken upstream oil supply using data on drilling and production. Following Anderson, Kellogg and Salant (2018) and Newell and Prest (2019), our upstream model incorporates dynamics in which the drilling of new wells is price-responsive, but wells' subsequent production stream is not. Finally, we assume that shippers' beliefs about the long-run distribution of future oil prices as of June, 2014, when they made firm capacity commitments to DAPL, are consistent with historic oil price volatility. We then validate our estimated model in section 4 by assessing how well its predicted flows match actual flows given realized downstream oil prices, and by comparing the model's expected returns to (committed) pipeline shipment to DAPL's tariff.

We present our main counterfactual analysis in section 5, characterizing what would have happened if DAPL had not been built. We primarily evaluate our counterfactuals from the perspective of June, 2014, integrating over the distribution of possible future downstream oil prices during the 2017–2027 period during which DAPL's shippers are committed to pay for pipeline capacity. We find that if DAPL's construction had been enjoined, expected crude-by-rail flows would increase by 81% of the decrease in expected pipeline flows. Blocking DAPL would therefore have kept some crude oil in the ground, but by an amount considerably less than DAPL's capacity.

We then quantify the normative implications of these changes in oil production and transportation flows by evaluating firms' lost producer surplus and the changes in local pollution emissions associated with both pipeline and rail transport. We use per-barrel damage estimates from Clay et al. (2019) that account for spill risks, air pollution from diesel railroad locomotives, and air pollution from electricity generators that power pipeline pumping stations. These estimates highlight that NO_x emissions from locomotives are the dominant factor, and accordingly we find that foreclosing DAPL would increase local pollution damages on net. Dividing these impacts by the quantity of carbon that the policy would "leave in

the ground," we find costs per tonne of CO_2 abated of \$28 from decreased producer surplus and \$17 from increased local emissions, so that a significant portion of the abatement cost of blocking DAPL is an increase in local pollution damages.

We extend and discuss our results in section 6. First, we evaluate an alternative policy of directly regulating upstream production by considering a production tax that would reduce CO_2 emissions by the same amount as foreclosing DAPL. Such a policy is analogous to the idea of "royalty adders" that have been considered for oil production on federally-owned oil and gas estates (Prest, 2022; Prest and Stock, 2023). Unlike blocking pipeline construction, this policy leads to a reduction in local pollution and overall imposes a small cost per tonne of CO_2 abated of between \$1.01 and \$2.68 (combining the loss of producer surplus with gains from reduced local pollution). We also discuss the possibility that the counterfactual reductions in Bakken oil production might be offset by increases in production from other basins, per Prest (2022) and Prest, Fell, Gordon and Conway (2023). This production "leakage" would increase the cost per tonne of CO_2 estimates of both our DAPL foreclosure and upstream production tax counterfactuals. Finally, we discuss the possibility that the Bakken production reductions are merely production delays rather than long-term reductions.

We conclude in section 7 by discussing how our analyses inform the trade-offs of policies that enjoin the construction of fossil fuel infrastructure. We find that blocking DAPL would indeed have reduced Bakken production, but at a cost per tonne of CO_2 abated that is an order of magnitude greater than what could have been achieved by directly targeting upstream supply. Of course, such an upstream policy may not be feasible, in which case the relevant question is whether incurring the costs of blocking infrastructure is acceptable relative to doing nothing at all. Our paper highlights that these costs include not just reductions in producer surplus but also increases in local pollution damages from crude-by-rail flows. This trade-off between global and local pollution is not typical of environmental policies—local pollution reductions are often a co-benefit of CO_2 abatement—but it is not unprecedented. Other situations in which CO_2 emissions trade off against local pollution exposure include urban densification (Carozzi and Roth, 2023) and the operation of post-combustion emissions controls in fossil fuel boilers (Electric Power Research Institute, 2009).

Finally, while this paper focuses on crude oil transportation, the intuition that we capture is applicable to other settings in which a low-cost but inflexible investment competes with alternatives that involve high marginal costs but little commitment. For instance, our modeling framework can be used to evaluate and understand the trade-offs between urban light rail, which requires investments in dedicated tracks and passenger loading stations, and passenger buses that can be flexibly re-routed (Glaeser, 2020). It can also apply to investments in baseload and renewable power sources, which involve high up-front sunk costs

and low marginal generation costs, versus investments in more flexible gas-fired peaker units (Borenstein, 2005). These trade-offs between cost and flexibility also have a parallel in the finance literature that examines the returns of investments in illiquid versus liquid assets (Amihud and Mendelson, 1986; Pástor and Stambaugh, 2003).

2   Conceptual model of crude-by-rail flows, pipeline flows, and pipeline investment

We begin with a conceptual model that captures the most important economic forces that we believe govern how policies that limit pipeline capacity would affect crude oil flows, and how policies that target upstream supply directly would affect pipeline investment and subsequent oil flows. This simplified model delivers intuition; when we apply it to the case of DAPL in section 3 we will enrich it to better match the data and additional institutional features.

Consider a setting with one upstream oil-producing location and one downstream oil-consuming location. Upstream, consumption is zero, and firms produce a quantity Q of oil according to a strictly increasing supply function $Q = S(P_u)$, where P_u is the upstream price. The downstream market is "large" in the sense that changes in deliveries from the upstream location do not alter the downstream price P_d.[3]

Oil moves from the upstream to downstream location by pipeline or railroad. "Shippers" are the agents who own the oil that is moved and pay for transportation services. We assume that shippers are atomistic and able to freely enter and exit the industry. This assumption is motivated by the large number of potential parties who may act as shippers: upstream producers, downstream refiners, and speculative traders.[4] Shipping decisions take place in two periods. In period 1, shippers decide whether to commit to pipeline capacity. In period 2, committed shippers use the pipeline up to the committed capacity K, and the remainder use the railroad to transport oil. We solve the model backwards, starting with period 2.[5]

2.1   Oil flows with an exogenously given pipeline capacity

Shippers who committed to the pipeline in period 1 can use it in period 2 to move oil at zero marginal cost. Let Q_p denote the volume of pipeline shipments, where $0 \leq Q_p \leq K$.

---

[3]We take this perfect elasticity of downstream demand assumption as a reasonable representation because DAPL's terminus is on the Gulf of Mexico, with access to the global waterborne crude oil market.

[4]Covert (2015) documents that Bakken upstream production is not a concentrated industry.

[5]We ignore discounting here for expositional clarity. The enriched model that we take to the data in section 3 incorporates discounting and allows the pipeline to be used for many periods.

Figure 1: Simple model for production, pipeline and rail flows, and the price differential, given a downstream crude price P_d and pipeline capacity K
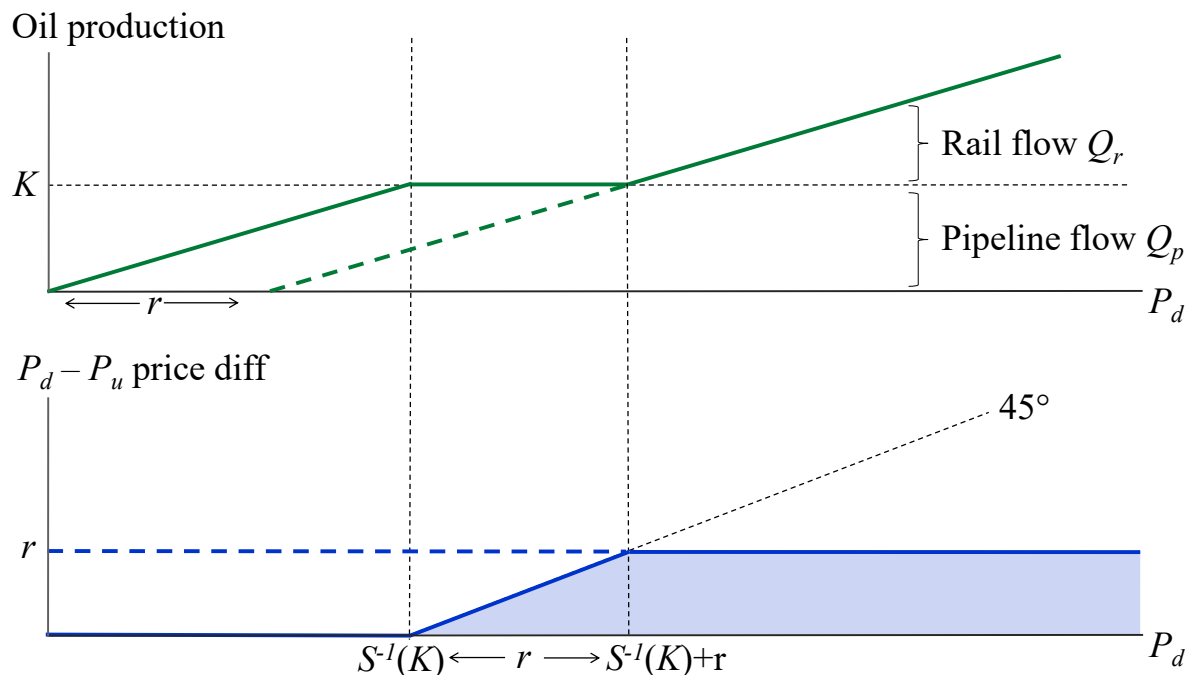
Top panel: solid line shows oil production as a function of the downstream oil price P_d when there is a pipeline with capacity K. Dashed line shows production when pipeline capacity is zero. Bottom panel: solid line shows the downstream minus upstream price differential as a function of the downstream price P_d when there is a pipeline with capacity K. Dashed line shows the price differential when pipeline capacity is zero.

Let Q_r denote the volume of rail shipments. The sum $Q_p + Q_r$ is total upstream production Q. There are no limits to crude-by-rail "capacity," but there are marginal costs $r > 0$.[6]

The pipeline capacity K and downstream price P_d determine Q_p and Q_r. For very low values of P_d, little crude oil is supplied by upstream producers, and the pipeline is not filled to capacity ($Q = Q_p < K$). Arbitrage then implies that $P_u = P_d$, and no crude flows by railroad. Q_p is increasing in P_d because the upstream supply curve is strictly upward-sloping, eventually filling the pipeline to capacity once $P_d = S^{-1}(K)$. Beyond this point, higher values of P_d cannot increase Q_p. Pipeline flows as a function of P_d and K are then given by $Q_p(P_d, K) = \min\{S(P_d), K\}$ and illustrated in the top panel of figure 1.

For downstream prices above $S^{-1}(K)$, Q_p is fixed, but rail may be used. Crude oil will move by rail only to the extent that the difference between P_d and P_u covers the marginal

---

[6]Railroads potentially exert market power, such that there may be a wedge between the true economic marginal cost of shipping crude by rail and the marginal cost paid by shippers (i.e., the freight rate) that is represented in our model by r. We do not take a stand on the extent to which r is true marginal costs versus market power rents.

cost r of railroad transport. When P_d lies in the interval $[S^{-1}(K), S^{-1}(K) + r]$, pipeline flow will be fixed at capacity K, rail flow Q_r will be zero, and the upstream price will be fixed at $S^{-1}(K)$. But for $P_d > S^{-1}(K) + r$, railroad volumes will be strictly positive and determined by the arbitrage condition $S^{-1}(K + Q_r) = P_u = P_d - r$. The function $Q_r(P_d, K)$ that governs rail flows as a function of P_d and K is $Q_r(P_d, K) = \max\{0, S(P_d - r) - K\}$ and is depicted in the top panel of figure 1.

The rail arbitrage condition ensures that rail shippers cannot earn economic profits in equilibrium. Pipeline shippers can, however, earn profits on the upstream versus down-stream price differential when the downstream price P_d is large enough that the pipeline is constrained. Their per-barrel profits $\pi_p(P_d, K)$ are given by equation 1 and depicted in the bottom panel of figure 1, capturing the feature that pipeline shippers' profits are capped, for high P_d, by the cost of railroad shipping.

$$\pi_p(P_d, K) = \begin{cases} 0 & \text{if } P_d \le S^{-1}(K) \\ P_d - S^{-1}(K) & \text{if } P_d \in (S^{-1}(K), S^{-1}(K) + r) \\ r & \text{if } P_d > S^{-1}(K) + r \end{cases} \tag{1}$$

Figure 1 also shows outcomes in the case that pipeline capacity is set to zero. For $P_d \ge S^{-1}(K) + r$, prices, production, and flows are unchanged from the case with a pipeline of capacity K because rail transport is on the margin in both cases, and therefore $P_u = P_d - r$ in both cases. For $P_d < S^{-1}(K) + r$, however, the upstream price is lower without the pipeline than with it, since rail is the only way to ship, and the cost of rail shipment induces a wedge r between P_u and P_d. The depressed upstream price causes oil production to fall relative to the case in which the pipeline was available; this decrease is depicted in the top panel of figure 1 by the gap between the solid and dashed lines to the left of $P_d = S^{-1}(K) + r$. The magnitude of this fall in oil production is an empirical question which depends on the cost of crude-by-rail transportation r and the elasticity of the upstream supply function $S(P_u)$.

2.2 Equilibrium pipeline capacity investment

We now analyze the equilibrium pipeline capacity commitments that occur in the first pe-riod. Two institutional details drive the assumptions we make in this analysis. First, because pipelines are irreversible investments that are subject to ex-post holdup, construction financ-ing requires firm, up-front commitments from shippers that they will pay for the pipeline's capacity whether they ultimately use that capacity or not (these commitments are therefore known as "ship-or-pay" agreements in the industry). These contracts imply that the most important risk associated with the project—the uncertainty about future price differentials

between the upstream and downstream locations—is borne primarily by shippers, not the pipeline owner. Second, pipelines could potentially exert market power over shippers, so their maximum tariffs are subject to cost-of-service rate regulation by FERC.

In period 1, shippers make K mbbl/d of firm pipeline capacity commitments before knowing P_d. They make these commitments on the basis of rational beliefs about the distribution of downstream prices $F(P_d)$. Committed shippers pay the regulated tariff τ for each barrel per day of capacity reserved. These committed shippers can then use their capacity in period 2 at zero marginal cost.

The equilibrium level of pipeline investment is then governed by an indifference condition in which shippers' expected per-barrel return from having pipeline access equals the tariff τ, which in turn is regulated to equal the pipeline's average per-barrel cost.[7] Taking expectations over the distribution $F(P_d)$, the equilibrium capacity investment K^* must satisfy equation 2:

$$\int \pi_p(P_d, K^*)dF(P_d) = \tau \tag{2}$$

The left-hand-side of equation 2 is strictly decreasing in K^*, since a larger capacity pipeline is less likely to be fully utilized, and it goes to zero as $K^* \to \infty$. It is also depicted in the bottom panel of figure 1 as the shaded area under the line representing pipeline shippers' returns $\pi_p(P_d, K)$, weighted by the probability distribution $F(P_d)$. Assuming that the cost of pipeline construction is not so large that not building the pipeline at all is optimal, equation 2 will then identify an interior solution for K^*.

3 Empirical model of Bakken oil production, pipeline transport, and rail transport

We now apply the intuition introduced in section 2 to DAPL and the production and movement of crude oil out of the Bakken region of North Dakota. Sections 3.1 through 3.4 describe the components of the model needed to quantitatively simulate Bakken production, transportation choices, and upstream prices, given a time series of downstream prices. Then,

---

[7]Note that this equilibrium condition for K^* potentially differs from that which would maximize total surplus to the extent that there are economies of scale in pipeline capacity. Total surplus is maximized when the expected per-barrel return to pipeline shippers equals the marginal cost of capacity, not its average cost. The two conditions coincide only when pipeline costs are constant returns to scale. The divergence between market equilibrium and surplus-maximizing investment in the presence of increasing returns to scale is driven by average-cost regulation of pipeline tariffs and is emblematic of rate regulation in natural monopoly settings. We assume constant returns to scale here; allowing for increasing returns would involve replacing the fixed tariff τ with a decreasing function $\tau(K)$. See footnote 33 for a discussion of how allowing for increasing returns would affect our counterfactual simulations involving an upstream production tax.

section 3.5 discusses how we model firms' beliefs—as of June, 2014 when shippers made firm commitments to DAPL—about the distribution of future downstream prices.

3.1 Price differentials, and crude-by-rail flows and costs

We begin with an empirical model for crude-by-rail flows and costs. Unlike the simpler model we developed in section 2, here we recognize that rail can move crude oil from the Bakken to not just the U.S. Gulf Coast (where DAPL terminates), but also the East and West Coasts. To infer crude-by-rail's cost structure, we use data on crude-by-rail flows and oil price differentials for each downstream location. We summarize these data here and provide additional detail in appendix B.1.

We obtain monthly oil prices at the three coastal destinations from Bloomberg (2023). We use the price of Light Louisiana Sweet (LLS) crude as the Gulf Coast price, Brent crude as the East Coast price, and a grade-adjusted version of Alaska North Slope (ANS) crude as the West Coast price. For the price of crude in the Bakken, we use data from S&P Global (2021), which publishes a Bakken "local" price that producers receive when they sell locally. We convert the raw nominal price data to real dollars as of June, 2014 using a CPI from Bureau of Labor Statistics (2023).

We obtain monthly data on Bakken oil production from U.S. Energy Information Administration (2021b) (hereafter "EIA"). To infer the volumes of oil that leave the Bakken by pipeline and rail, we multiply monthly production by monthly transportation mode share estimates from North Dakota Pipeline Authority (2023a) (hereafter "NDPA"). Then, to infer the share of rail shipments flowing to each coastal destination each month, we use the EIA's "Movements of Crude Oil and Selected Products by Rail" report, which provides estimates of rail flows that the EIA infers from the Surface Transportation Board's (STB's) waybill sample.

We display our price and crude-by-rail volume data together in figure 2, spanning May, 2010 (the earliest date for which we have upstream price data) through 2019. The top panel shows the time series for the crude oil price at the U.S. Gulf Coast. The most prominent feature of this series is the substantial price decrease that occurred during the second half of 2014. Baumeister and Kilian (2016) attributes this price decrease primarily to a decrease in global demand for crude oil, with a smaller role played by global production shocks.

The middle panel of figure 2 shows the difference, for each coastal destination, between the downstream oil price and the upstream price in the Bakken. This panel reveals three facts about these prices. First, the the three coastal prices are tightly correlated, typically differing by no more than a few $/bbl. Overall, the West Coast price increased, and the

Figure 2: Oil prices, price differentials, and crude-by-rail flows
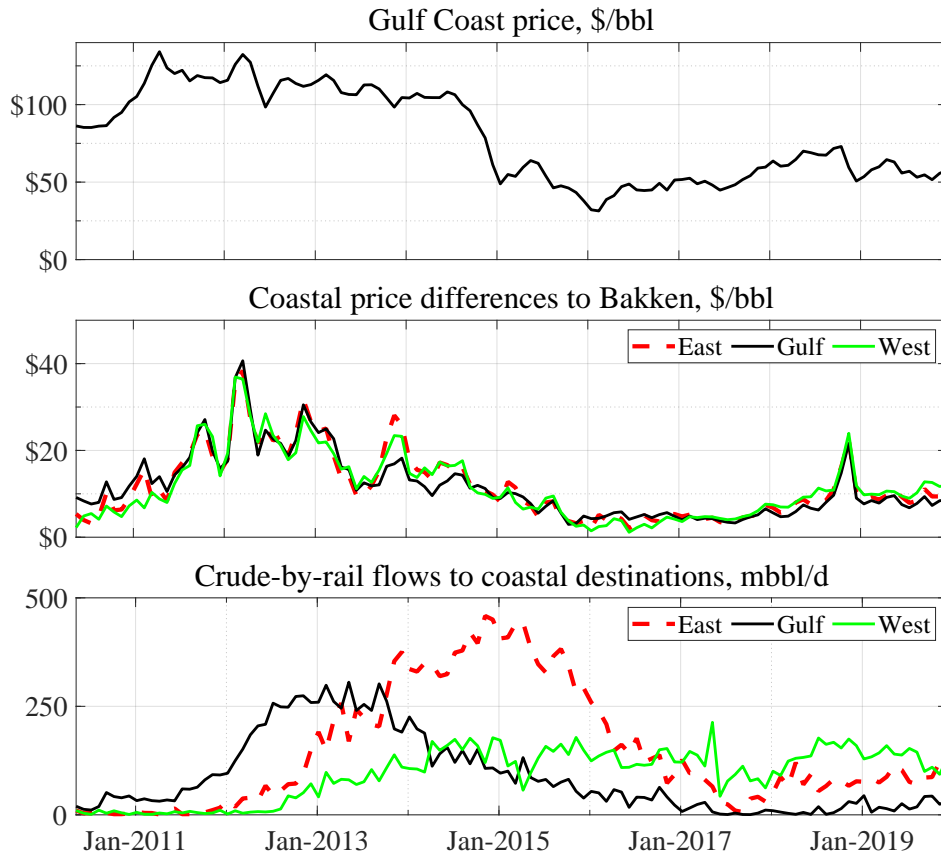
Note: The top panel shows the price for crude oil delivered to the Gulf Coast, measured by the Light Louisiana Sweet (LLS) price. The middle panel shows the difference between prices at coastal markets and in the Bakken, where we use Brent for the East Coast price and grade-adjusted Alaska North Slope (ANS) for the West Coast Price. The bottom panel shows crude-by-rail flows to each coastal destination in thousands of barrels of oil per day (mbbl/d). All prices are real June, 2014 dollars and aggregated to the monthly level.

Gulf Coast price decreased, relative to the East Coast price during 2010–2019, but overall no destination held a persistent price advantage over another. Second, the upstream Bakken price is substantially discounted relative to the coastal destinations. Third, the upstream price discount contracted following the decrease in the coastal prices in late 2014.

Finally, the bottom panel of figure 2 shows crude-by-rail flow volumes from the Bakken to the East, Gulf, and West Coasts. Total shipments rise substantially beginning in 2012, peak in late 2014, and generally decrease thereafter. Comparing the middle and bottom panels, the increase and subsequent decrease in crude-by-rail volumes follows the increase and decrease in price differentials that were realized by rail shippers, though it is clear that rail volumes do not respond to price differentials immediately but rather with a lag. Changes in the share of rail flows going to each destination also evolve gradually rather than respond

10

instantly to changes in relative prices.

The sluggish response of crude-by-rail flows to price changes rejects the simple model posed in section 2, in which rail transport just involves a constant marginal cost $r > 0$. Moreover, a multi-destination version of this simple model would imply that all rail volumes should flow each month to the destination that offers the highest downstream price net of shipping costs, and that total rail flows should be zero if, for all locations, the price differential less the shipping cost is strictly negative. The rail flow data shown in figure 2 defy both of these predictions: every destination has strictly positive rail flows in every month from 2012 onward, including 2015–2018 when the price differential to all locations hovered around \$5/bbl, considerably lower than industry reports of crude-by-rail costs of \$8 to \$15/bbl (Frittelli, Parfomak, Ramseur, Andrews, Pirog and Ratner, 2014; ICF, 2020).

To reconcile our model with these facts, marginal rail shipping costs must vary over time and differ across destinations. We assume that period-to-period adjustment costs are the key time-varying force in shipping costs. In practice, these adjustment costs take the form of investments (and dis-investments) in capital such as rail cars, loading facilities in North Dakota, and unloading facilities in downstream locations.[8] Thus, while rail shipping is not as frictionless in practice as assumed by our simple model from section 2, it is still considerably more flexible than shipping by pipeline.

We model crude-by-rail frictions with a quadratic adjustment cost specification in the spirit of Hall (2004), which we interpret as a reduced form for the process of making multiple small discrete investments or dis-investments in crude-by-rail capacity. Our model's marginal cost of shipping oil to destination i in month t therefore includes a dynamic term $\gamma(Q_{rit} - Q_{ri,t-1})$ that is linear in the difference between the current and previous months' rail volumes to destination i, with $\gamma \geq 0$.

The total marginal cost of shipping to destination i in month t is then given by the sum $r_i + \gamma(Q_{rit} - Q_{ri,t-1})$, where r_i is a destination-specific static marginal cost. We allow this cost to be destination-specific because each of the three coastal destinations is a different distance from the Bakken area, and because there may be other route-specific factors that impact costs. For instance, rail traffic through the U.S. to the East Coast will typically pass through the congested Chicago area.

Our assumption that shippers are atomistic and can freely enter and exit then implies

---

[8]These investments are reported to be substantially cheaper, per unit of capacity, than pipeline investments. For example, at the low end, Fielden (2018) documents that the Plains All American loading facility in Manitou, ND cost \$40 million for 65,000 bbl/d of capacity, or roughly \$600 per bbl/d of capacity. At the higher end, Area Development News Desk (2018) documents that Enbridge spent \$145 million on an 80,000 bbl/d facility, or \$1,813 per bbl/d of capacity. Both of these figures lie substantially below the \$9,200 per bbl/d of capacity invested for DAPL (Energy Transfer Partners LP, 2017).

that in equilibrium, the price differential to location i must equal $r_i + \gamma(Q_{rit} - Q_{ri,t-1})$ whenever rail flows are strictly positive.[9] Letting P_{it} and P_{ut} denote the destination i price and upstream price in month t, this arbitrage condition is given by equations 3 and 4:

$$P_{it} - P_{ut} - r_i - \gamma(Q_{rit} - Q_{ri,t-1}) \leq 0 \tag{3}$$

$$Q_{rit}(P_{it} - P_{ut} - r_i - \gamma(Q_{rit} - Q_{ri,t-1})) = 0. \tag{4}$$

To estimate γ and the r_i cost parameters, we rearrange equation 4 and add a disturbance term ε_{it} to obtain equation 5:

$$Q_{rit} - Q_{ri,t-1} = \frac{-r_i}{\gamma} + \frac{1}{\gamma}(P_{it} - P_{ut}) + \varepsilon_{it}. \tag{5}$$

The disturbance term accounts for two types of unobservables. First, rail flows Q_{rit} are measured with error because the EIA estimates these flows from the STB's sample of waybills rather than from the universe of actual rail flows. Second, there may be unobserved shocks to the cost of rail shipping. These shocks would affect both current period rail flows Q_{rit} and the upstream price P_{ut}, which would lead an OLS estimate of equation 5 to be inconsistent. We therefore estimate equation 5 via 2SLS, instrumenting for the $(P_{it} - P_{ut})$ term using the first three lags of the East Coast (Brent) oil price. These instruments have strong first stage predictive power, and they will also satisfy an exclusion restriction if changes in crude oil flows out of the Bakken do not affect the global Brent benchmark price of oil. We estimate equation 7 using data from August 2012 (when rail shipments to the West Coast first exceeded 10 mbbl/d) through December 2019 (the end of our sample).

We present the 2SLS estimates of equation 5 in table 1. The estimated marginal cost intercepts for shipping to the East, Gulf, and West Coasts are \$9.49/bbl, \$12.64/bbl, and \$8.69/bbl, respectively. These estimates are in line with industry reports (Frittelli et al., 2014; ICF, 2020) and reflect the fact that the West Coast is the shortest distance to travel from the Bakken (Clay et al., 2019). Our estimate of γ is \$1.28/bbl per mbbl/d, which implies that increasing rail flows to a destination by 10 mbbl/d from one month to the next increases the marginal shipping cost by \$12.76/bbl. Thus, adjustment costs are a substantial portion of total rail shipping costs.

---

[9]The assumption that shippers are atomistic implies that they will be price takers both upstream and downstream. The free entry and exit assumption implies that the arbitrage condition given by equations 3 and 4 holds, rather than a classic Euler equation that would include a forward-looking term. That is, in our model free entry and exit compete away current-period railroad shipping rents, rather than the sum of current and expected future rents.

Table 1: Rail cost function estimates

| | Rail cost intercepts r_i | | | |
	East Coast	Gulf Coast	West Coast	Friction parameter γ
Point estimate	9.49	12.64	8.69	1.28
Standard error	(3.32)	(2.05)	(2.87)	(0.29)
Units	\$/bbl	\$/bbl	\$/bbl	\$/bbl per mbbl/d

Table shows estimates of equation 5 via 2SLS. Standard errors on the structural parameters are computed using the delta method and Driscoll and Kraay (1998), which allows for both spatial and temporal correlation in the residuals, using the plug-in bandwidth per Newey and West (1994). The first-stage F-statistic is 23.08 ($p < 0.001$). "mbbl/d" denotes units of thousands of barrels per day. All costs are in real June, 2014 dollars.

3.2   Model of upstream oil production

We next estimate a model of upstream supply in the spirit of Anderson et al. (2018). This framework makes the physically realistic assumption that "new" production, from newly drilled wells, is sensitive to crude oil prices, while "old" production, from pre-existing wells, is not.[10] Thus, in the short run, total upstream crude oil production is highly inelastic with respect to the oil price, since "old" production is a large share of total production. In the long run, however, production can respond significantly to persistent price shocks as changes in the rate of drilling affect production rates. Thus, the model we implement allows for considerably richer production dynamics than the simple static model from section 2.

Let Q_{ot}, Q_{nt}, and Q_t denote old, new, and total Bakken oil production each month. Per section 3.1, data on Q_t come from U.S. Energy Information Administration (2021b). To compute Q_{ot} and Q_{nt}, we use the EIA's "Drilling Productivity Reports" (DPRs), which estimate the contribution of new drilling to oil production each month (U.S. Energy Information Administration, 2021a). We derive Q_{nt} from the DPR data (details in appendix B.3) and then compute Q_{ot} as $Q_t - Q_{nt}$.

Following Anderson et al. (2018), we model the evolution of production from old wells as following equation 6, which specifies an exponential decline with a decay parameter $\beta \in (0, 1)$:

$$Q_{ot} = \beta Q_{t-1}. \tag{6}$$

---

[10]Anderson et al. (2018) explains that this behavior can be rationalized by the combination of low per-bbl marginal extraction costs once a well has been completed with the existence of a geologic production capacity constraint that is a function of the remaining reserves. Anderson et al. (2018) only studies conventional wells, but Newell and Prest (2019) shows that the price non-responsiveness of production from existing wells is also a feature of shale oil production.

We estimate equation 6 by projecting Q_{ot} onto Q_{t-1}, with no constant. We estimate $\beta = 0.955$, with a standard error of 0.002.[11]

Unlike production of old wells, the drilling of new wells is price-responsive. We follow Newell and Prest (2019)—which estimates the price-responsiveness of drilling across all major U.S. shale oil plays—by modeling new production Q_{nt} as a constant elasticity supply function of current and lagged upstream prices per equation 7. As discussed in Newell and Prest (2019), lagged prices are important because planning and executing the drilling of a new well can take several months. We also allow the productivity of Bakken drilling to evolve over time, which we capture with time-varying intercepts θ_t.

$$\log Q_{nt} = \theta_t + \sum_{\ell=0}^{L_n} \theta_{P\ell} \log P_{u,t-\ell}, \tag{7}$$

To estimate the elasticity parameters in equation 7, we follow Newell and Prest (2019) by taking first differences so that we have stationary series on both the left and right-hand sides of our estimating equation. We set the maximum lag L_n equal to nine months, since we find that longer lags do not significantly impact drilling. We also pool the monthly price coefficients to the quarterly level in order to avoid over-fitting the data, resulting in the following estimating equation:[12]

$$\Delta \log Q_{nt} = \alpha_1 + \alpha_{P0} \sum_{\ell=0}^{2} \Delta \log P_{u,t-\ell} + \alpha_{P1} \sum_{\ell=3}^{5} \Delta \log P_{u,t-\ell} + \alpha_{P2} \sum_{\ell=6}^{8} \Delta \log P_{u,t-\ell} + \varepsilon_{nt}. \tag{8}$$

The disturbances ε_{nt} correspond to shocks to the productivities θ_t in the supply function given by equation 7. These shocks may be simultaneously determined with the contemporaneous upstream price P_{ut}, which would bias downward our estimate of α_{P0}. While the inclusion of the α_1 parameter controls for linear technical progress over time, our main strategy for addressing this simultaneity is to instrument for $\sum_{\ell=0}^{2} \Delta \log P_{u,t-\ell}$ with the sum of the first three differences of the logged Brent oil price. As was the case when we estimated our crude-by-rail cost model in section 3.1, this instrument will be valid if shocks to Bakken production do not materially influence the global oil price.

We present the estimates of equation 8 in table 2. The total elasticity of new Bakken oil production to a permanent price change, given by three times the sum of α_{P0}, α_{P1}, and α_{P2},

---

[11]The estimated standard error is robust to heteroskedasticity and autocorrelation in the residuals per the Andrews (1991) HAC estimator.

[12]That is, we set $\theta_{P0} = \theta_{P1} = \theta_{P2} = \alpha_{P0}$, with similar expressions holding for the longer lags. If we instead estimate each monthly coefficient $\theta_{P\ell}$, the overall long-run elasticity we estimate is similar to what we obtain from estimating equation 8, but a few of the coefficients have negative point estimates.

Table 2: Upstream supply function estimates

| | Trend parameter α_1 | Upstream price elasticities α_P | | |
		Current quarter α_{P0}	Lagged quarter α_{P1}	Second lagged quarter α_{P2}
Point estimate	0.017	0.13	0.20	0.11
Standard error	(0.005)	(0.04)	(0.03)	(0.04)

Table shows estimates of equation 8 via 2SLS. Standard error estimates are robust to heteroskedasticity and autocorrelation in the residuals per the Andrews (1991) HAC estimator. The first-stage F-statistic is 210.17 ($p < 0.001$).

is 1.32. This value is comparable to the supply elasticity of 1.1 to 1.2 estimated in Newell and Prest (2019) using data from all of the major U.S. shale plays. Thus, even though the inelasticity of production from old wells limits the price-responsiveness of Bakken production in the short-run, the long-run price elasticity of production is substantial.

Finally, we solve for the monthly productivity intercepts θ_t by inverting equation 7 for each period t. Because the imputed θ_t are noisy (owing to noise in the raw production data), when we simulate the model we use smoothed values obtained from fitting a sixth-degree polynomial fit to the imputed θ_t.[13] Both the imputed and smoothed θ_t series are shown in appendix figure A.2.

3.3 Bakken pipeline capacity and the Dakota Access Pipeline

The Dakota Access Pipeline (DAPL) was put into service in June, 2017 with a capacity to move 520 mbbl/d from the Bakken to the Gulf Coast.[14] Our empirical model also accounts for pipelines other than DAPL that export crude oil from the Bakken. North Dakota Pipeline Authority (2023a) reports 763 mbbl/d of non-DAPL pipeline capacity, including the Double H Pipeline that was completed in February, 2015 with a capacity of 84 mbbl/d and expanded in January, 2016 to 108 mbbl/d. However, some of this capacity was not actually able to move oil all the way to the U.S. Gulf Coast due to downstream capacity constraints (ICF, 2020).[15] We therefore take total non-DAPL capacity to be equal to the average rate of

---

[13]The production data Q_t are sufficiently noisy that there are 4 months during 2011–2019 in which we impute $\exp(\theta_t) = 0$. This outcome occurs when βQ_{t-1} exceeds Q_t in the data.

[14]DAPL itself moves oil to Patoka, IL. Its completion was coincident with the completion of the Energy Transfer Crude Oil Pipeline (ETCO) from Patoka, IL to Nederland, TX on the Gulf Coast. For brevity, in the rest of this paper we refer to DAPL and ETCO jointly as just DAPL.

[15]Pre-existing pipelines out of the Bakken connect to other pipeline systems to the east and south (in particular the Enbridge Mainline, Platte, and Pony Express pipelines) to move oil onward to demand centers, but these pipelines have been constrained because they also carry oil from western Canada. In contrast,

Bakken pipeline exports from January, 2016 through February, 2017. We compute this value to be 586 mbbl/d, based on multiplying total Bakken production (per the EIA) by the share of production exported by pipeline (per the NDPA) each month.[16]

We also account for the fact that, unlike our analytic model from section 2, there is non-zero, albeit small, local demand for oil in the Bakken area. Per North Dakota Pipeline Authority (2023a), existing local refining capacity in June, 2014 was 88 mbbl/d. The NDPA also reports small volumes of oil that are trucked north from the Bakken and injected into spare capacity on the Canadian pipeline network. We model local Bakken oil demand as inframarginal—and therefore perfectly inelastic—at a quantity equal to the average combined local refining and trucking volumes during June, 2017 (when DAPL went into service) through December, 2019 (the end of our sample period): 139 mbbl/d.

3.4   Forward simulation of the full model

This section characterizes our model's equilibrium Bakken oil production, pipeline flows, and rail flows in each month t, given contemporaneous downstream prices and available pipeline capacity (DAPL and non-DAPL) K_t. Let P_{Et}, P_{Gt}, and P_{Wt} denote the East, Gulf, and West coast prices, respectively. Pipeline shipments of oil can only access the Gulf Coast market, but rail shipments can access any of the three markets. Let P_u^t denote the history of upstream prices at t (not including P_{ut}), and let Q_L denote local Bakken area oil consumption.

Equilibrium in our model in each period t is then defined by the following conditions:

1. Upstream production Q_t equals decayed old supply plus new supply: $Q_t = \beta Q_{t-1} + Q_{nt}(P_{ut}, P_u^t, \theta_t)$, where the function $Q_{nt}(P_{ut}, P_u^t, \theta_t)$ is given by equation 7.

2. Pipeline flows Q_{pt} are given by:

$$
Q_{pt}(P_{Gt}, P_{ut}, K_t) \begin{cases} = 0 & \text{if } P_{Gt} < P_{ut} \\ \in [0, K_t] & \text{if } P_{Gt} = P_{ut} \\ = K_t & \text{if } P_{Gt} > P_{ut} \end{cases}
$$

3. Rail flows Q_{rit} to each destination i are given by $Q_{rit} = \max\{0, Q_{ri,t-1} - \frac{r_i}{\gamma} + \frac{1}{\gamma}(P_{it} - P_{ut})\}$

DAPL, together with the ETCO pipeline, carries oil directly to the U.S. Gulf Coast.

[16]January, 2016 was the month the Double H pipeline was expanded, and February, 2017 was the last month before oil began to flow into DAPL to fill and commission the pipeline before it came fully into service in June, 2017. The difference of 177 mbbl/d between nameplate and effective non-DAPL capacity that we find is similar to the value of 243 mbbl/d reported in ICF (2020).

4. Upstream production equals the sum of local consumption and all pipeline and rail flows: $Q_t = Q_L + Q_{pt} + Q_{rEt} + Q_{rGt} + Q_{rWt}$

Given K_t, P_u^t, the downstream prices P_{it}, the lagged rail flows $Q_{ri,t-1}$, and the supply intercept θ_t, these four equilibrium conditions determine Q_{pt}, the Q_{rit}, Q_t, and P_{ut} each period. We prove the existence and uniqueness of this equilibrium in appendix C. Thus, starting from the initial conditions and the upstream price history in June, 2014 (when shippers committed to DAPL), we can use our estimated model to forward simulate Bakken production, pipeline flows, and rail flows given time series for P_{it} and K_t. Formally, letting $\mathbf{P_t}$ denote the vector of downstream prices $[P_{Et}, P_{Gt}, P_{Wt}]$ and letting $\mathbf{Q_{r,t-1}}$ denote the vector of lagged rail flows $[Q_{rE,t-1}, Q_{rG,t-1}, Q_{rW,t-1}]$, we can forward simulate:

$$P_{ut} = P_u(\mathbf{P_t}, P_u^t, \mathbf{Q_{r,t-1}}, K_t, \theta_t) \tag{9}$$

$$Q_{pt} = Q_p(\mathbf{P_t}, P_u^t, \mathbf{Q_{r,t-1}}, K_t, \theta_t) \tag{10}$$

$$Q_{rit} = Q_{ri}(\mathbf{P_t}, P_u^t, \mathbf{Q_{r,t-1}}, K_t, \theta_t) \text{ for } i \in \{E, G, W\} \tag{11}$$

$$Q_t = Q_t(\mathbf{P_t}, P_u^t, \mathbf{Q_{r,t-1}}, K_t, \theta_t) \tag{12}$$

Relative to the equilibrium in the simple model from section 2, the dynamics from the upstream supply function and crude-by-rail adjustment costs in this model cause crude oil flows to respond to downstream price shocks gradually rather than immediately. Thus, if downstream prices are rising quickly, the price differential can exceed the baseline rail shipping cost r_i. Alternatively, following a decrease in downstream prices rail flow can be strictly positive even if the price differential is strictly less than r_i.

3.5 Shippers' beliefs and equilibrium pipeline capacity

Thus far, we have described and estimated a model that allows us to forward simulate oil production and transportation flows given time series inputs of pipeline capacity and downstream prices. We next describe a model of pipeline capacity commitment and investment that builds on the framework introduced in section 2.2, wherein shippers' capacity commitment equates the shipping margin they expect to realize from having pipeline access to the pipeline's tariff τ. Applying this framework to DAPL requires us to specify shippers' beliefs about the distribution of future downstream oil prices as of June, 2014, when they executed binding ten-year "ship-or-pay" contracts with DAPL (Energy Transfer Partners LP, 2014). Quantifying this part of the model has two payoffs. First, doing so allows us to simulate equilibrium pipeline investment in the upstream production tax counterfactual discussed in section 6.1. Second, by specifying shippers' future price beliefs we gain the ability to evaluate

all of our policy counterfactuals from the perspective of June, 2014, when the future path of oil prices was not yet known.

Let $t = 0$ denote June, 2014, let $t_c = 37$ denote June, 2017 (when DAPL went into service following its construction), and let $T = 156$ denote May, 2027 (the last month of shippers' ten-year commitment period). Let $\delta = 0.9919$ denote the monthly discount factor.[17] Let K_{nt} denote the time series of non-DAPL capacity, and let K_d denote DAPL's capacity. Finally, let $\mathcal{P}$ denote a time series of downstream prices, $\mathcal{P} = \{\mathbf{P_1}, \mathbf{P_1}, ..., \mathbf{P_T}\}$, and let $F(\mathcal{P} \mid \mathbf{P_0})$ denote shippers' beliefs about the distribution of $\mathcal{P}$ at $t = 0$. The equilibrium condition for K_d is then given by equation 13, which is the analog to equation 2 from the simpler model in section 2.2, now enriched to accommodate multiple periods, multiple rail destinations, and dynamics in rail transport and upstream production:

$$\tau = \frac{1}{\sum_{t=t_c}^{T} \delta^t} \int \left(\sum_{t=t_c}^{T} \delta^t \left(P_{Gt} - P_u(\mathbf{P_t}, P_u^t, \mathbf{Q_{r,t-1}}, K_{nt} + K_d, \mathbb{E}[\theta_t]) \right) \right) dF(\mathcal{P} \mid \mathbf{P_0}). \quad (13)$$

To compute the right-hand side of equation 13, we need to specify the distribution $F(\mathcal{P} \mid \mathbf{P_0})$, the expected intercepts $\mathbb{E}[\theta_t]$ of the Bakken upstream supply function, and the volume of capacity K_d that shippers committed to. Given these ingredients, we compute the right-hand side of equation 13 by Monte Carlo simulation, where for each draw of $\mathcal{P}$ from $F(\mathcal{P} \mid \mathbf{P_0})$, $P_u(\mathbf{P_t}, P_u^t, \mathbf{Q_{r,t-1}}, K_{nt} + K_d, \mathbb{E}[\theta_t])$ is solved for constructively for periods $t = 1$ through $t = T$ using the per-period equilibrium conditions from section 3.4, starting from the initial conditions $\mathbf{P_0}$, P_u^1, and $\mathbf{Q_{r0}}$.

To specify and estimate $F(\mathcal{P} \mid \mathbf{P_0})$, we first focus on the process for the Gulf Coast price P_G, since this price is most relevant to pipeline shippers. We adopt a parsimonious approach and assume that $\log P_{Gt}$ follows the AR(1) process in equation 14, where the innovations ϵ_{Gt} are iid normal with mean zero and variance σ_G^2:

$$\log P_{Gt} = \phi_0 + \phi_1 \log P_{G,t-1} + \epsilon_{Gt}. \quad (14)$$

When estimating the parameters ϕ_0, ϕ_1, and σ_G^2, we target two objects: the long-run expected value of P_G, denoted $\mathbb{E}[P_G]$, and the evolution of its variance from the short-run (one month ahead) through the long-run (156 months ahead).[18] We summarize our

---

[17]The monthly discount factor δ is based on an annual nominal discount rate of 12.5% (Kellogg, 2014) and a June, 2014 annual inflation forecast of 2.0% (Federal Reserve Bank of Atlanta, 2014).

[18]An alternative strategy would be to estimate equation 14 directly via OLS, using the monthly LLS price series. We eschew this strategy both to make use of the known June, 2014 futures price and because we want to target long-run oil price volatility directly, and modeled volatility is very sensitive to ϕ_1 for $\phi_1 \approx 1$.

estimation procedure here and provide more detail in appendix B.4. For $\mathbb{E}[P_G]$, we use the June, 2014 three-year Brent crude future price of \$93.47/bbl (Quandl, 2017).[19] Note that this expected price is higher than prices that were realized through the rest of the 2010s, since oil prices fell in late 2014 (see figure 2). For the variance, at each horizon $t \in [1, 156]$ we target the historic variance of log oil prices, computed as the variance of t-month differences in logged Brent prices.[20]

We fit our AR(1) model to the historic variances using a minimum distance estimator and obtain estimates of ϕ_1 and σ_G equal to 0.9925 and 0.098, respectively. Intuitively, the short-run variance identifies σ_G^2, and the long-run variance identifies ϕ_1 (the long-run variance is larger the closer is ϕ_1 to 1). These estimates imply that future oil price volatility increases from 10.3% at a one month horizon, to 69.0% at 37 months, and to 114% at 13 years, closely matching historic volatilities (see appendix figure A.3). Finally, the expected price $\mathbb{E}[P_G] =$ \$93.47/bbl pins down the estimate of $\phi_0 = 0.0293$.

We next specify and estimate the processes for the East and West Coast prices P_E and P_W. We assume that differences between these prices and P_G follow the joint AR(1) process given by equations 15 and 16.

$$P_{Et} - P_{Gt} = \phi_E + \phi_{EE}(P_{E,t-1} - P_{G,t-1}) + \phi_{EW}(P_{W,t-1} - P_{G,t-1}) + \epsilon_{Et} \tag{15}$$

$$P_{Wt} - P_{Gt} = \phi_W + \phi_{WE}(P_{E,t-1} - P_{G,t-1}) + \phi_{WW}(P_{W,t-1} - P_{G,t-1}) + \epsilon_{Wt} \tag{16}$$

We assume that ϵ_{Et} and ϵ_{Wt} are bivariate normal with mean zero. We estimate all the parameters of equations 15 and 16 by OLS, using price data from the same August, 2012 to December, 2019 sample period that we used when estimating the rail cost function in section 3.1. The estimates are consistent with strong mean reversion of spatial price differences: the eigenvalues of the matrix formed by the ϕ_{EE}, ϕ_{EW}, ϕ_{WE}, and ϕ_{WW} parameters are 0.72 and 0.79.[21] The estimates overall imply that the long-run coastal price differences are positive but small: $\mathbb{E}[P_E - P_G] =$ \$0.91/bbl and $\mathbb{E}[P_W - P_G] =$ \$0.84/bbl. These positive estimates increase the value of crude-by-rail and diminish DAPL's value to pipeline shippers.

Next, we estimate a time series of expected Bakken upstream supply function intercepts

---

When we estimate 14 directly, we obtain an estimate of $\phi_1 = 0.9852$, similar to the value of 0.9925 that we estimate from long-differenced oil price volatility.

[19] We use the three-year future because this is the longest horizon at which contracts are liquidly traded, and we use Brent rather than Louisiana Light Sweet (LLS) because there is no LLS futures market and because the Brent and LLS prices have historically been quite close (see figure 2). We equate the futures price with the price expectation because Anderson et al. (2018) finds a CAPM beta for oil of nearly zero using data through April, 2015, implying little risk premium or discount in oil futures.

[20] We use Brent rather than LLS to be consistent with our use of Brent to measure $\mathbb{E}[P_G]$.

[21] The point estimates are $\phi_{EE} = 0.60$, $\phi_{EW} = 0.23$, $\phi_{WE} = $ -0.10, and $\phi_{WW} = 0.91$. The jointly normal distribution of ϵ_{Et} and ϵ_{Wt} has $\sigma_E =$ \$1.36/bbl, $\sigma_W =$ \$1.25/bbl, and $\sqrt{\sigma_{EW}} =$ \$0.93/bbl.

$\mathbb{E}[\theta_t]$ for $t \in [1, 156]$. We do not use our intercepts estimated in section 3.2 from realized production data because we are interested in specifying firms' expectations of production as of June, 2014, and because we need to estimate these expectations through 2027. We estimate the $\mathbb{E}[\theta_t]$ using a June, 2014 production forecast from North Dakota Pipeline Authority (2014). This forecast is discussed in more detail in appendix B.4 and plotted in appendix figure A.4. Average expected production during DAPL's ten-year contract period is 1620 mbbl/d. This production forecast, combined with $\mathbb{E}[P_G]$ and the estimated supply parameters from section 3.2, pins down the expected supply intercepts $\mathbb{E}[\theta_t]$.[22]

Finally, we specify shippers' capacity commitment K_d. DAPL's capacity when it opened in June, 2017 was 520 mbbl/d, and we use this value for K_d when we compute equation 13. However, as discussed in more detail in appendix B.4, it is difficult to be certain of the total capacity to which Bakken pipeline shippers committed in June, 2014. We therefore test the sensitivity of our results to alternative values of 320, 450, and 570 mbbl/d.

4 Validation

We conduct two validation exercises to assess how well our model "fits the data" and can match some data features that we haven't used in estimation. First, starting from initial conditions in June, 2014, we forward simulate the model through 2019 given the time series of realized downstream (but not upstream) oil prices, pipeline capacity, and estimated upstream supply intercepts. We then compare the model's simulated oil flows to actual flows observed in our data. This comparison tests how well the full set of equilibrium conditions in our model, given in section 3.4, can rationalize observed flows, given limited input information.

The simulated pipeline flows and aggregate rail flows from this simulation are shown in figure 3.[23] The model overall captures several salient features of actual pipeline and rail flows over 2014–2019. First, both simulated and actual rail flows decline gradually from June, 2014 through June, 2017, driven by the decline in oil prices that occurred in late 2014 (recall figure 2). Pipeline flows, in contrast, do not decrease during this time. Second, the model captures the gradual increase in pipeline flows following DAPL's completion in June, 2017. Actual pipeline flows initially increase more quickly than our simulated flows, potentially reflecting anticipatory effects that are not present in our model.[24] Finally, our

---

[22] In an alternative specification, we allow for uncertainty in upstream supply by letting the supply intercept be stochastic, using a conservative production forecast from the NDPA. See appendix B.4 for details, and see appendix tables A.2 and A.3 for simulation results that allow for supply uncertainty.

[23] Appendix figure A.5 presents alternative simulations that use simplified versions of our model that include only a single rail destination and shut down the dynamics of upstream production and crude-by-rail flows. These models fit the data less well.

[24] Figure 3 shows that pipeline flows actually start to increase a few months prior to June, 2017. These flows

Figure 3: Simulated and actual pipeline and rail flows, June, 2014 through 2019
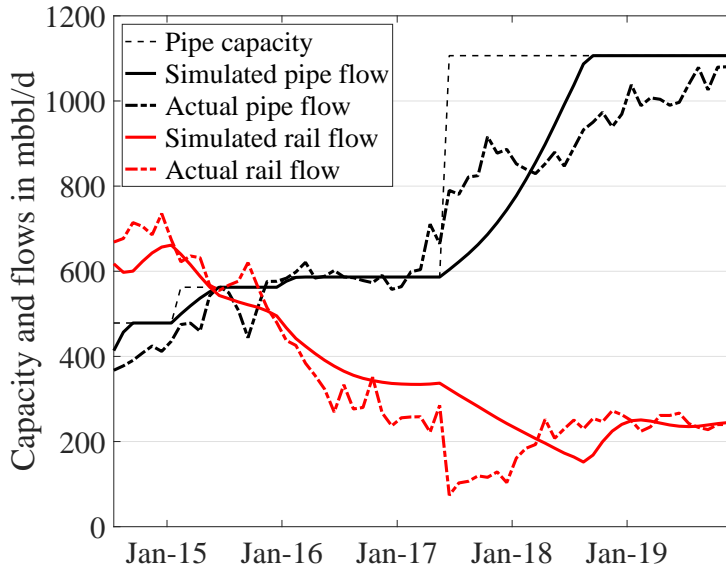

Figure shows actual and forward simulated pipeline flows and aggregate rail flows from the Bakken, using realized downstream prices through 2019. The simulation uses the model discussed in sections 3.1 through 3.4, starting from initial conditions as of June, 2014.

model captures the modest increase in crude-by-rail volumes that started in 2018 following a rebound in oil prices.[25]

In our second validation exercise, we test the equilibrium condition that the expected return for DAPL's committed shippers as of June, 2014 should equal DAPL's tariff, per equation 13. This exercise serves two purposes. First, it informs how well our model would predict pipeline investment under counterfactual policies that directly regulate upstream production. Second, it informs how well our estimated distribution of future oil prices $F(\mathcal{P} \mid \mathbf{P_0})$ matches shippers' beliefs, which in turn speaks to our policy evaluations that take expectations as of June, 2014, integrating over $F(\mathcal{P} \mid \mathbf{P_0})$.

The actual DAPL tariff for committed shippers is \$5.50–\$6.25/bbl (Gordon, 2017), with

---

may reflect DAPL's "line fill": the filling of the pipeline with oil before it was formally placed into service. The total line fill for DAPL is nearly 2 million barrels, given the pipeline's 1,872 mile length (including ETCO) and its 30 inch diameter (Dutta and Huchzermeyer, 2017).

[25]In appendix figure A.6, we break out the simulated and actual crude-by-rail flows by destination. Although our model does a good job of matching the decline of crude-by-rail volumes to the Gulf Coast, it does less well at capturing the divide between flows to the East versus West Coasts, potentially reflecting a delayed build-out of West Coast crude-by-rail terminal capacity (Fielden, 2013). In appendix B.2, we show an alternative specification of our model that better fits destination-specific rail flows by changing the r_E and r_W cost parameters. Our counterfactual simulations are qualitatively unchanged when we use this alternative specification.

shippers committing higher volumes paying a tariff at the lower end of this range. When we compute the right-hand-side of equation 13 using our model and our estimate of $F(\mathcal{P} \mid \mathbf{P_0})$, we obtain an expected return of \$6.17/bbl.[26] We view this result as a close match to the actual DAPL tariff.

Appendix table A.1 shows that the expected shipper returns that we would calculate under alternative specifications match the DAPL tariff less well. These specifications use a random walk model for the evolution of the Gulf Coast oil price P_G, or they simplify the model by assuming only a single rail destination and shutting down dynamics.[27]

5 Counterfactual simulations: what if DAPL had not been constructed?

5.1 Bakken oil production and transportation without DAPL

We now use our estimated model to evaluate Bakken oil production and transportation in a counterfactual in which DAPL's construction had been foreclosed. Figure 4 shows simulated counterfactual flows given realized downstream oil prices through 2019. In this counterfactual, pipeline volumes stay fixed at the non-DAPL pipeline capacity of 586 mbbl/d from mid-2016 onward. Unlike our baseline simulation that includes DAPL's capacity starting in June, 2017, rail flows in our counterfactual increase rather than decrease after this date. By December, 2019, simulated crude-by-rail flows without DAPL are 595 mbbl/d, compared to 247 mbbl/d with DAPL. Thus, by December, 2019, 348 mbbl/d (67%) of the 520 mbbl/d of pipeline oil flows that would have been eliminated by foreclosing DAPL would still be produced, but moved out of the Bakken by railroad rather than by pipeline.

Actual downstream oil prices after June, 2014 were unknown when DAPL shippers signed firm transportation agreements that month. Recognizing this uncertainty, we focus in the remainder of the paper on evaluating counterfactuals from the perspective of shippers and policy-makers in June, 2014, taking expectations over the distribution of downstream prices $F(\mathcal{P} \mid \mathbf{P_0})$ that we estimated in section 3.5.

Table 3 presents simulated expected pipeline, rail, and production volumes of Bakken oil, both with and without DAPL. In either scenario, expected Bakken production is greater than

---

[26]The simulation error from our 10000 draw Monte Carlo simulation of the integral in equation 13 implies a 95% confidence interval of $\pm\$0.15$/bbl around our expected return estimate of \$6.17/bbl.

[27]Appendix table A.2 then presents expected returns using the baseline model with alternative values for DAPL's committed capacity. For commitments of 450 mbbl/d or 570 mbbl/d, the expected return is similar to our value of \$6.17/bbl that uses the actual installed DAPL capacity of 520 mbbl/d. The expected return is substantially higher, however, if one assumes that only 320 mbbl/d of new Bakken pipeline capacity were committed to in June, 2014.

Figure 4: Simulated pipeline and crude-by-rail flows, with and without DAPL
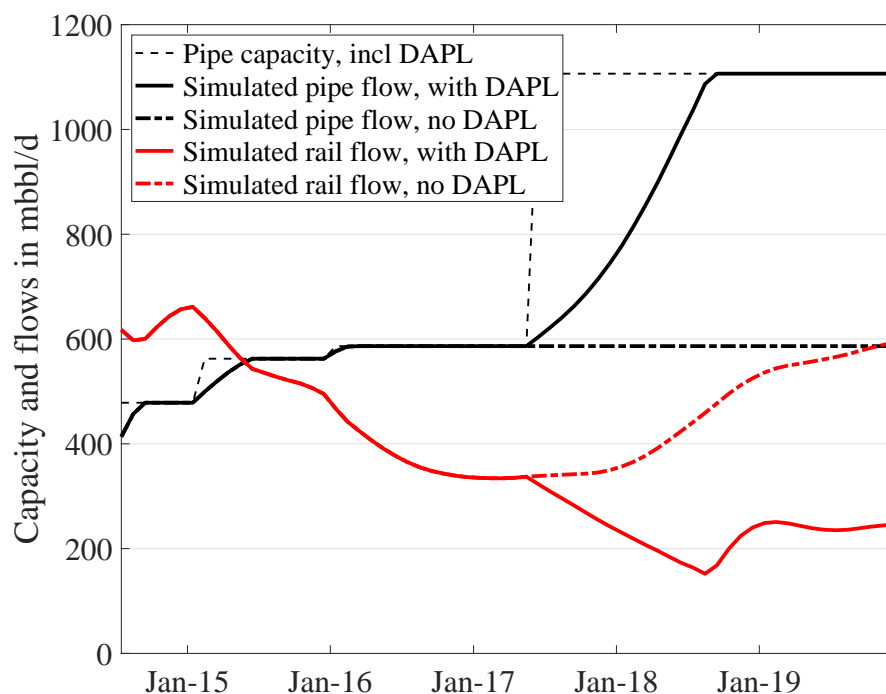

Figure shows forward simulations using the full model discussed in section 3, starting from initial conditions as of June, 2014 and using realized downstream prices through 2019, both with and without the addition of 520 mbbl/d of DAPL capacity in June, 2017. Simulated flows "with DAPL" are identical to those shown in figure 3.

what was realized over 2014–2019 because expected downstream prices, as of June, 2014, were higher than realized prices. In expectation, removing DAPL reduces pipeline flows by 306 mbbl/d. This magnitude is smaller than DAPL's 520 mbbl/d capacity because for very low downstream price realizations, DAPL is not fully utilized. Expected rail flows increase by 248 mbbl/d, offsetting 81% of the decrease in pipeline flows.[28] Bakken oil production decreases by 58 mbbl/d (4%). Thus, railroads' ability to effectively, even if incompletely, substitute for Bakken pipeline transportation implies that blocking pipeline construction would cause most of the precluded pipeline oil flow to divert to the railroads rather than "stay in the ground."

---

[28]Appendix table A.3 shows that the share of reduced pipeline flows that is offset by the increase in crude-by-rail flows is nearly invariant across alternative specifications that use different values for DAPL's committed capacity, allow for a random walk belief for future oil prices, or allow for supply uncertainty.

Table 3: Simulated flows with and without DAPL, in expectation

	Volume with DAPL mbbl/d	Volume without DAPL mbbl/d	Change in volume mbbl/d	Percent change in volume
Pipeline flows	827	521	-306	-37%
Rail flows	702	950	248	35%
Local Bakken consumption	139	139	0	0%
Bakken production	1529	1471	-58	-4%

All expectations are taken as of June, 2014 and are averages over 10000 Monte Carlo draws of possible downstream price paths. For each simulated price path, we compute average discounted volumes (pipeline, rail, local, total) during DAPL shippers' ten-year commitment period (June, 2017 - May, 2027).

5.2 Environmental and economic surplus impacts

5.2.1 Pollution emission factors and damages

To assess the environmental consequences of transporting Bakken crude, we rely primarily on emissions factors and damages estimates from Clay et al. (2019). This paper estimates: (1) emissions of CO_2, NO_x, and SO_x associated with pipeline transportation from the Bakken to the USGC, based on pipeline pumping stations' electricity consumption and on marginal generators' emissions factors; (2) emissions of CO_2, NO_x, VOCs, and particulate matter associated with railroad transportation from the Bakken to each of the three coastal destinations, using locomotive emissions factors; (3) pollution damage valuations from the AP3 integrated assessment model (Muller, 2014), which uses an EPA 2014 VSL of $8.5 million; and (4) expected damages from spills and accidents for both pipeline and rail transport. Because Clay et al. (2019) provides pipeline emissions factors for 2011 and rail emissions factors for 2014, we adjust its reported values to account for changes in the electric generation and locomotive fleets over time using data from U.S. Environmental Protection Agency (2023) and U.S. Department of Transportation (2018). Details are provided in appendix B.5.

We present the $/bbl damage estimates for 2019 in table 4; appendix figure A.7 shows the evolution of these damages over time. As in Clay et al. (2019), the greatest environmental damage from Bakken oil transport comes from railroad NO_x emissions, owing both to locomotives' high NO_x emissions factors and to the fact that these emissions often occur in densely populated areas. At a social cost of carbon (SCC) of $100 per metric ton (tonne) of CO_2, monetized local pollution damages exceed CO_2 damages for rail transport to all three destinations, but for pipeline transportation CO_2 damages exceed costs from local pollution.

Our analysis also considers the CO_2 emissions associated with the oil itself. Each pro-

Table 4: Estimated damages from pipeline and rail transit of Bakken crude in 2019, \$/bbl

	Local air pollution	Spills	CO$_2$ (at \$100/tonne)
Pipeline	\$0.35	\$0.11	\$0.83
Rail to East Coast	\$3.00	\$0.73	\$0.79
Rail to Gulf Coast	\$1.66	\$0.73	\$0.79
Rail to West Coast	\$0.76	\$0.52	\$0.57

Estimates computed from Clay et al. (2019), U.S. Environmental Protection Agency (2023), and U.S. Department of Transportation (2018). Values are in real 2014 dollars. See text for details.

duced barrel of oil emits 0.432 tonnes of CO$_2$ when it is consumed (U.S. Environmental Protection Agency, 2022a). Thus, at a SCC of \$100/tonne, the avoided climate damages from preventing the production (and consumption) of oil are \$43.22/bbl, significantly greater than the per-barrel damages from oil transportation shown in table 4.

5.2.2 Producer surplus calculations

Private surplus losses in our model are borne entirely by oil producers. With exogenous downstream prices (an assumption that changes in Bakken oil exports are too small to affect the global oil market), consumer surplus impacts of any quantity change are zero. Similarly, because we assume rail shippers arbitrage away rail shipping profits, they too earn zero surplus. Finally, our pipeline commitment equilibrium implies that committed pipeline shippers earn zero rents in expectation.

To compute producer surplus, we take advantage of the fact that there are no economic distortions other than environmental externalities in our model. Thus, we can compute the private surplus loss from foreclosing DAPL by evaluating a hypothetical pipeline tax that would have been sufficient to cause no shippers to commit to DAPL. More precisely, let λ be a \$ per bbl/d tax on pipeline capacity, and let $K_d(\lambda)$ denote committed DAPL capacity as a function of the tax. The $K_d(\lambda)$ function is implicitly defined by a modification of the original equilibrium condition 13 that now includes the tax λ:

$$\tau + \lambda = \frac{\int \left(\sum_{t=t_c}^{T} \delta^t \left(P_{Gt} - P_u(\mathbf{P_t}, P_u^t, \mathbf{Q_{r,t-1}}, K_{nt} + K_d(\lambda), \mathbb{E}[\theta_t]) \right) \right) dF(\mathcal{P} \mid \mathbf{P_0})}{\sum_{t=t_c}^{T} \delta^t}. \quad (17)$$

Let $\lambda_0 > 0$ be the value of the pipeline capacity tax such that $K_d(\lambda_0) = 0$; in our

Table 5: Changes in environmental and private surplus from foreclosing DAPL, in expectation

	Expectations over 10-year DAPL contract
Δ Pipe flows (mbbl/d)	-306
Δ Pipe local pollution damage ($1000/d)	-$144
Δ Pipe CO_2 emissions (mtonnes/d)	-2.5
Δ Rail flows (mbbl/d)	248
Δ Rail local pollution damage ($1000/d)	$588
Δ Rail CO_2 emissions (mtonnes/d)	1.8
Decrease in producer surplus ($1000/d)	$716
Δ CO_2 from combustion (mtonnes/d)	-25.2
Local damages per tonne CO_2 abated ($/tonne)	$17
Lost PS per tonne CO_2 abated ($/tonne)	$28
Damages + lost PS per tonne CO_2 abated ($/tonne)	$45

All expectations are taken as of June, 2014 and are averages over 10000 Monte Carlo draws of possible downstream price paths. For each simulated price path, we compute average discounted outcomes during DAPL shippers' ten-year commitment period (June, 2017 - May, 2027). All monetary values are in real June, 2014 dollars. "mtonnes/d" denotes thousands of metric tons per day.

estimated model we find $\lambda_0 = \$3.03$/bbl.[29] We then evaluate the change in producer surplus $\Delta(PS)$ as the following Harberger triangle:[30]

$$\Delta(PS) = \int_0^{\lambda_0} s(K_d(s) - K_d(0))ds. \tag{18}$$

5.2.3 Welfare impacts of foreclosing DAPL

Table 5 presents our estimates of the expected environmental and producer surplus effects of foreclosing DAPL in June, 2014. The decrease in pipeline flows of 306 mbbl/d causes a reduction in local environmental harm of $144,000 per day and a reduction in CO_2 emissions

---

[29]The sum $\tau + \lambda_0$ is $9.20/bbl and is the shadow value of pipeline capacity at $K_d = 0$. This shadow value increases as total pipeline capacity decreases but will be bounded above by the cost of crude-by-rail.

[30]We evaluate equation 18 by Gauss-Legendre quadrature, solving equation 17 at each quadrature node. We set $\tau = \$6.17$/bbl, the value of the right-hand-side of equation (17) when $\lambda = 0$, with $K_d(0) = 520$ mbbl/d. We use the Harberger triangle approach in our producer surplus calculations because we cannot directly evaluate the level of producer surplus at $K_d = 520$ mbbl/d and $K_d = 0$. In particular, our upstream production model is designed to capture the dynamics with which Bakken production responds to price shocks, but it is not founded on a model of dynamically optimizing behavior that would allow us to directly evaluate producers' profit functions.

of 2.5 thousand metric tonnes (2.5 mtonnes) per day. The 248 mbbl/d increase in crude-by-rail flows, however, increases local pollution damage by \$588,000 per day and increases CO_2 emissions by 1.8 mtonne/d. Thus, even though the magnitude of the decrease in pipeline oil flow exceeds the increase in rail flow, overall local environmental damage increases because the per-barrel harm from railroad transport exceeds that from pipeline transport.

We estimate a decrease in producer surplus of \$716,000 per day, greater than the change in monetized local environmental harm but of a comparable magnitude. The decrease in oil production of 58 mbbl/d is associated with a 25.2 mtonne/d decrease in CO_2 emissions from oil consumption. This emissions reduction is much greater than the changes in emissions associated with changes in pipeline and rail flows, even though the change in oil production is much smaller in magnitude than the swing in transportation from pipeline to rail.

The bottom section of table 5 computes the the cost of foreclosing DAPL in terms of dollars per tonne of avoided CO_2 emissions. The increase in local pollution damages imposes a cost of \$17/tonne, while the decrease in producer surplus amounts to \$28/tonne. The total cost, with equal weight on environmental damages and producers' surplus losses, is \$45/tonne. This cost is similar to the U.S. government's contemporaneous value for the SCC of \$42/tonne (Interagency Working Group on the Social Cost of Carbon, 2013) but lower than the value of \$190 per tonne proposed in U.S. Environmental Protection Agency (2022b).[31] Section 6 below further characterizes these costs by evaluating alternative policies and discussing factors outside our model that may be important.

6 Discussion and extensions

6.1 Policy alternative: upstream production tax

In this section, we compare outcomes from blocking pipeline construction to outcomes from policies that directly tax oil production upstream in the Bakken. An example of such a policy is a "royalty adder," like that under consideration for federally-owned oil and gas resources (Prest, 2022; Prest and Stock, 2023). The vast majority of Bakken oil lies on private rather than public mineral estates, so in the Bakken the tax would instead have to take the form of a state-imposed severance tax.[32]

---

[31]The average SCC in Interagency Working Group on the Social Cost of Carbon (2013) for 2014, using a 3% discount rate to evaluate future climate damages, is \$37/tonne in 2007 dollars. Accounting for inflation per Bureau of Labor Statistics (2023) yields \$42/tonne. The \$190 per tonne figure is for the year 2020 in 2020 dollars and was computed using a 2% discount rate to evaluate future climate damages.

[32]The State of North Dakota already imposes a 5% production tax and a 5% extraction tax, so this counterfactual calculation would not represent something totally unprecedented. See `https://www.tax.nd.gov/business/oil-and-gas-severance-tax`.

We model a Bakken production tax as a wedge between the price received by upstream producers and the price paid by pipeline and rail shippers for oil in the Bakken. To normalize our comparisons between policies, we set the value of the tax so that the induced reduction in CO_2 emissions is the same as that achieved by blocking the pipeline (26.0 mtonnes/d). We set the effective date of the tax as June, 2017 (DAPL's in-service date).

We model the tax in two ways. First, we treat the tax as being announced after DAPL shippers made their firm commitments, so that we hold DAPL's capacity fixed at 520 mbbl/d. In this case, a production tax of \$3.68/bbl (equivalent to \$8.52/tonne CO_2) reduces emissions by the same amount as blocking DAPL's construction. Second, we model the tax as being announced before the firm contracts were executed, and we compute the new equilibrium pipeline capacity that accounts for the tax. In this case, the emissions-equivalent production tax is \$3.24/bbl (\$7.51/tonne CO_2), and the new pipeline capacity is 443 mbbl/d.[33]

Table 6 compares outcomes from the DAPL foreclosure policy and the two tax policies. Because the policies are normalized on CO_2 emissions, they all induce roughly the same reduction in Bakken oil production. But the production taxes do not induce the large shift in flows from pipeline to rail that characterizes the ban on DAPL's construction. Holding DAPL's capacity fixed, the production tax reduces both pipeline and rail volumes, while if DAPL's capacity responds to the tax, the entire volume reduction comes from pipeline flows.

The fourth and fifth rows of table 6 show that the production tax induces a large transfer from producers to the government of five to six million dollars per day, depending on how the tax's timing is modeled.[34] This transfer is nearly an order of magnitude greater than the \$716,000 per day decrease in producers' surplus when DAPL is banned. It arises from the fact that the tax affects inframarginal barrels of oil, not just the transported oil that is on the margin. An alternative policy of a production quota with tradeable credits would avoid the large transfer and leave all other impacts unchanged from those shown in table 6.

The tax reduces combined producer surplus and government revenue by \$109,000 per day if DAPL's capacity is held fixed, and by \$96,000 per day if DAPL's capacity responds to the tax. These surplus reductions—which are the deadweight loss of the tax ignoring environmental effects—are considerably smaller than the \$716,000 per day reduction in producers' surplus when DAPL is banned. Intuitively, banning DAPL induces a large, distortionary shift in oil transport mode that increases the industry's costs, whereas upstream production taxes do not.

Turning to environmental impacts, we find that while banning DAPL increases local

---

[33]The counterfactual capacity of 443 mbbl/d assumes no economies of scale in pipeline construction. To the extent that economies of scale are important, the upstream production tax would cause a larger decrease in DAPL's capacity than what what we simulate here.

[34]These transfers may induce a positive fiscal externality that we do not account for in our analysis.

Table 6: Expected impacts from foreclosing DAPL and from taxing
Bakken oil production

	Blocking DAPL	Production tax, DAPL capacity fixed	Production tax, DAPL capacity adjusts
Impacts per day			
Δ Pipe flows (mbbl/d)	-306	-29	-59
Δ Rail flows (mbbl/d)	248	-30	0
Δ Total flows (mbbl/d)	-58	-59	-59
Δ Producer surplus ($1000/d)	-$716	-$6035	-$5316
Δ Tax revenue ($1000/d)	-	$5926	$5221
Δ PS + Δ tax revenue ($1000/d)	-$716	-$109	-$96
Δ Local pollution damage ($1000/d)	$444	-$82	-$26
Δ CO_2 emissions (mtonnes/d)	-26.0	-26.0	-26.0
Costs per tonne of CO_2 abated			
Lost producer surplus	$27.56	$232.26	$204.59
Tax revenue	-	$228.07	$200.90
Lost PS - tax revenue	$27.56	$4.19	$3.68
Increase in local pollution damages	$17.08	-$3.17	-$1.01
Lost PS - tax revenue + local pollution	$44.63	$1.01	$2.68

All expectations are taken as of June, 2014 and are averages over 10000 Monte Carlo draws of possible downstream price paths. For each simulated price path, we compute average discounted outcomes during DAPL shippers' ten-year commitment period (June, 2017 - May, 2027). The production tax in column (2) is $3.68/bbl ($8.52/tonne CO_2), and that in column (3) is $3.24/bbl ($7.51/tonne CO_2). The simulation used to generate column (3) assumes that DAPL's equilibrium capacity investment anticipates the tax and is therefore 443 mbbl/d rather than 520 mbbl/d. All monetary values are in real June, 2014 dollars. "mtonnes/d" denotes thousands of metric tons per day.

environmental damages, taxing upstream production modestly reduces these damages. The increase in local damages from banning DAPL stems from the large increase in crude-by-rail traffic. Taxing upstream production instead weakly decreases both pipeline and rail flows.

The bottom half of table 6 summarizes these results in terms of costs per tonne of abated CO_2 emissions. Here again the tax's transfer from oil producers to the government is large: at least $201 per tonne. Its deadweight loss is small, however: just $3.68 per tonne in the case in which DAPL's capacity responds to the tax. Putting equal weight on all impacts, the total cost per tonne of CO_2 from the production tax is $1.01 holding DAPL's capacity fixed, and $2.68 if DAPL's capacity responds to the tax. This overall impact is considerably smaller than the total cost of blocking DAPL: $45 per tonne. This difference reflects both the tax's relatively small deadweight loss and its induced decrease (rather than increase) in

local environmental pollution damage.

6.2   Potential impacts on oil production from other basins

Thus far, our welfare analysis has assumed that every barrel of avoided Bakken oil production translates to a barrel of avoided oil consumption. However, this assumption will be violated if decreased Bakken production leads to increases in oil production elsewhere via the re-equilibration of the global market. The mechanism for such "leakage" of production would be an increase in world oil prices that induced production increases outside of the Bakken. The strength of this mechanism depends on supply and demand elasticities: leakage will be greatest when global oil demand is inelastic and non-Bakken supply is elastic (Prest et al., 2023; Weisbach, Kortum, Wang and Yao, 2023).[35]

Rather than develop and quantify a model of the global oil market—a task we view as outside the scope of this paper—we rely on recent research to assess how leakage might affect the results from our counterfactual analyses. Prest (2022) estimates a leakage rate of 52–72% for policies that reduce production on U.S. federal lands, using data from International Energy Agency (2019); these estimates were used to evaluate federal oil royalty adders in Prest and Stock (2023). Prest et al. (2023) calculates a leakage rate of 55% using a meta-analysis of global oil demand and supply elasticity estimates from the literature. Based on this work, we re-evaluate our counterfactual analyses using leakage rates of 52% and 72%.

When we account for production leakage, the policies' effects on CO_2 emissions associated with combustion of the produced oil are all attenuated by a factor of $(1 - \ell)$, where ℓ is the leakage rate. This attenuation then inflates the policies' costs per tonne of CO_2 abated, as shown in table 7. At a leakage estimate of 72%, the combined private and local environmental cost from foreclosing DAPL increases from \$45 to \$159 per tonne of CO_2 abated. The combined cost of taxing upstream production increases from \$1.01 to \$3.62 per tonne (or from \$2.68 to \$9.56 per tonne if DAPL's capacity responds to the tax).

6.3   Limitations

While our model of upstream supply is dynamic in the sense that it allows oil production each period to be a function of lagged prices, it does not incorporate a Hotelling-style model of resource exhaustion. One potential concern with our approach is then that policies that reduce oil production today are really just postponing production rather than truly "keeping it in the ground."

---

[35]Note that our infinite Bakken downstream oil demand elasticity assumption in the analysis thus far is consistent with either highly elastic global oil demand or highly elastic non-Bakken supply.

Table 7: Expected costs per tonne of CO_2 abated from foreclosing DAPL and from taxing Bakken oil production, accounting for production leakage to other basins

	Blocking DAPL	Production tax, DAPL capacity fixed	Production tax, DAPL capacity adjusts
No leakage			
Lost PS - tax revenue	$27.56	$4.19	$3.68
Increase in local pollution damages	$17.08	-$3.17	-$1.01
Lost PS - tax revenue + local pollution	$44.63	$1.01	$2.68
52% leakage			
Lost PS - tax revenue	$57.41	$8.72	$7.67
Increase in local pollution damages	$35.57	-$6.61	-$2.10
Lost PS - tax revenue + local pollution	$92.99	$2.11	$5.58
72% leakage			
Lost PS - tax revenue	$98.42	$14.95	$13.16
Increase in local pollution damages	$60.98	-$11.33	-$3.60
Lost PS - tax revenue + local pollution	$159.41	$3.62	$9.56

All expectations are taken as of June, 2014 and are averages over 10000 Monte Carlo draws of possible downstream price paths. For each simulated price path, we compute average discounted outcomes during DAPL shippers' ten-year commitment period (June, 2017 - May, 2027). The production tax in column (2) is $3.68/bbl ($8.52/tonne CO_2), and that in column (3) is $3.24/bbl ($7.51/tonne CO_2). The simulation used to generate column (3) assumes that DAPL's equilibrium capacity investment anticipates the tax and is therefore 443 mbbl/d rather than 520 mbbl/d. All monetary values are in real June, 2014 dollars.

Similar to Prest (2022) and Prest et al. (2023), we do not view dynamics related to resource exhaustion as a first-order threat to our analyses. The Bakken production data do not exhibit any evidence of a long-run decline: since the Covid-19 interruption in spring 2020, production has been roughly constant between 1,000 and 1,100 mbbl/d through at least July, 2023 (U.S. Energy Information Administration, 2021b). Thus, to the extent that reserve exhaustion has been important, productivity improvements like learning-by-doing (Kellogg, 2011; Covert, 2015; Agerton, 2020) have been sufficient to compensate. North Dakota Pipeline Authority (2023b), moreover, forecasts that under the 2023 EIA price forecast, Bakken production will actually increase in the coming decade and not fall below 1,000 mbbl/d until sometime after its last forecast year of 2047. We therefore conclude that policies that reduce Bakken oil production today will keep the otherwise-produced oil in the ground for a very long time, quite plausibly beyond the time at which substitute fuels and technologies (or strong carbon policies) substantially reduce oil demand.

Second, there are limits to the scope of environmental effects we quantify in this paper.

We do not quantify harms from local pollution associated with upstream extraction. There is a literature that studies pollution from shale oil production using large-scale data, focusing mainly on the Marcellus shale in Pennsylvania (Currie, Greenstone and Meckel, 2017; Bonetti, Leuz and Michelon, 2021; Zhang, Li, Khanna, Krupnick, Hill and Sullivan, 2023), but we are unaware of a comprehensive valuation study analogous to Clay et al.'s (2019) study of transportation-related emissions. The extant literature suggests that impacts are local to drilling and production sites, and given the Bakken area's low population density we see these damages as being small in magnitude relative to those we quantify.[36] Additionally, we do not incorporate local environmental damages from downstream oil refining or the consumption of refined products, nor do we evaluate emissions associated with oil's substitutes in consumption. These considerations may be important, though they would not affect our relative cost per tonne estimates across policies because they would proportionally re-scale the cost per tonne of each.

7   Conclusions

This paper demonstrates that initiatives to "keep carbon in the ground" by blocking fossil fuel transportation infrastructure can present difficult trade-offs. For oil pipelines, these trade-offs arise because the availability of crude-by-rail as a substitute can still allow oil to reach demand centers, and because this substitute induces larger environmental externalities than the blocked infrastructure itself. Despite the differences between pipelines' and railroads' technology and cost structure, we find that these two transport modes are strongly substitutable, such that if the construction of DAPL had been enjoined, 81% of the blocked pipeline flows would have moved by rail instead in expectation. This quantitative conclusion is of course specific to DAPL, but our modeling framework can be applied to other situations, across which the cost and availability of substitute transport may vary.

We find that the combined private plus local environmental cost of blocking DAPL, per tonne of CO_2 abated, is \$45. This value is on par with the contemporaneous (2014) U.S. social cost of carbon (SCC) of \$42/tonne (Interagency Working Group on the Social Cost of Carbon, 2013) but significantly lower than the value of \$190 per tonne recently proposed in U.S. Environmental Protection Agency (2022b). Thus, blocking DAPL may still pass a Kaldor-Hicks cost-benefit test under recent valuations of the SCC, though this conclusion can be overturned if "leakage" of oil production to other basins is sufficiently large.

---

[36]Even if local damages from extraction were of the same per-bbl magnitude as local damages from transportation, the large changes in transportation mode share (relative to the change in production) associated with blocking DAPL implies that our results on local pollution damages would be qualitatively unchanged.

Moving beyond a standard cost-benefit lens, blocking DAPL presents an environmental justice dilemma, since the policy reduces global climate damages while imposing local pollution damages onto communities near railroad corridors. This trade-off between local and global pollution is typically absent from other carbon abatement policies that directly target fossil fuel production or consumption itself. In these cases, substitution is often to other energy technologies that have lower, not higher, local pollution externalities.

Finally, we find that alternative policies that directly target upstream production can avoid the local environmental externalities imposed by blocking DAPL while also imposing less of a burden on productive efficiency. However, upstream policies face their own challenges. Upstream carbon taxes, which could be implemented as "royalty adders" or severance taxes, generate large transfers from industry to the government that can potentially exceed monetized climate benefits and lead to political resistance. And upstream policies in general require authority over upstream producers and resource owners, presenting a challenge in the U.S. because the vast majority of onshore oil and gas resources are privately-owned. Pipeline infrastructure projects, in contrast, present multiple veto points because approvals are needed in every state that the proposed line would pass through. In situations in which upstream policies are not on the table because local authorities are unable or unwilling to enact them, blocking pipeline infrastructure may therefore still present itself as an attractive option to advocates and policy-makers who strongly value carbon reductions.

References

Agerton, Mark, "Learning Where to Drill: Drilling Decisions and Geological Quality in the Haynesville Shale," 2020. working paper.

Amihud, Yakov and Haim Mendelson, "Asset Pricing and the Bid-Ask Spread," *Journal of Financial Economics*, 1986, *17*, 223–249.

Anderson, Soren T., Ryan Kellogg, and Stephen W. Salant, "Hotelling Under Pressure," *Journal of Political Economy*, 2018, *126*, 984–1026.

Andrews, Donald W. K., "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 1991, *59* (3), 817–858.

Area Development News Desk, "Enbridge Pipelines Completes Berthold, North Dakota Rail Terminal, Continues Pipeline Expansion Project," 2018. [Online; `http://www.areadevelopment.com/newsItems/8-19-2013/`

`enbridge-pipelines-sandpaper-pipeline-project-berthold-north-dakota289123.`
`shtml`; accessed 6 September, 2018].

Baumeister, Christiane and Lutz Kilian, "Understanding the Decline in the Price of Oil since June 2014," *Journal of the Association of Environmental and Resource Economists*, 2016, *3*, 131–158.

Bloomberg, "U.S. Crude oil futures prices," accessed 31 July, 2023.

Bonetti, Pietro, Christian Leuz, and Giovanna Michelon, "Large-sample evidence on the impact of unconventional oil and gas development on surface waters," *Science*, 2021, *373*, 896–902.

Borenstein, Severin, "The long-run efficiency of real-time electricity pricing," *Energy Journal*, 2005, pp. 93–116.

Bureau of Labor Statistics, "Consumer Price Index, all urban, all items less energy, not seasonally adjusted (CUUR0000SA0LE)," https://www.bls.gov/cpi/data.htm; accessed 9 August, 2023.

Carozzi, Felipe and Sefi Roth, "Dirty density: Air quality and the density of American cities," *Journal of Environmental Economics and Management*, 2023, *118*, 102767.

CBC, "'Climate Leaders Don't Build Pipelines': Indigenous Advocate Tells Trudeau," *CBC News*, May 2019.

Clay, Karen, Akshaya Jha, Nicholas Muller, and Randall Walsh, "The External Costs Of Transporting Petroleum Products: Evidence From Shipments Of Crude Oil From North Dakota by Pipelines and Rail," *Energy Journal*, 2019, *40* (1), 55–72.

Covert, Thomas R., "Experiential and Social Learning in Firms: the Case of Hydraulic Fracturing in the Bakken Shale," 2015. working paper.

Currie, Janet, Michael Greenstone, and Katherine Meckel, "Hydraulic fracturing and infant health: New evidence from Pennsylvania," *Science Advances*, 2017, *3*, e1603021.

Dahir, Abdi Latif, "An Oil Rush Threatens Natural Splendors Across East Africa," *The New York Times*, March 2023.

Dalton, Jane, "Greece and Israel Agree to Deal to Build World's Longest Underwater Gas Pipeline Despite Pledge to Cut Fossil Fuels," *The Independent*, January 2020.

Davis, Lucas W, Catherine Hausman, and Nancy L Rose, "Transmission Impossible? Prospects for Decarbonizing the US Grid," June 2023. NBER working paper 31377.

Driscoll, John C. and Aart C. Kraay, "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *Review of Economics and Statistics*, 1998, *80* (4), 549–560.

Dutta, Ashok and Laura Huchzermeyer, "Feature: Dakota Access Pipeline Started Up With 50,000-b/d More Crude Oil Capacity," *Platts News*, 2017.

Electric Power Research Institute, "Survey of Impacts of Environmental Controls on Plant Heat Rate," Palo Alto, CA, report 1019003 2009.

Energy Transfer Partners LP, "Energy Transfer Announces Crude Oil Pipeline Project Connecting Bakken Supplies to Patoka, Illinois and to Gulf Coast Markets," 2014. [Online; `http://ir.energytransfer.com/phoenix.zhtml?c=106094&p=irol-newsArticle&ID=1942689`; accessed 13 September, 2017].

_ , "Energy Transfer Announces the Bakken Pipeline is in Service Transporting Domestic Crude Oil from the Bakken/Three Forks Production Areas," 2017. [Online; `http://ir.energytransfer.com/phoenix.zhtml?c=106094&p=irol-newsArticle&ID=2278014`; accessed 13 March, 2018].

Federal Reserve Bank of Atlanta, "Atlanta Fed Survey of Business Inflation Expectations," June 2014. https://www.atlantafed.org/-/media/Documents/news/pressreleases/BIEsurvey/2014/1406.pdf.

Fell, Harrison, Daniel T Kaffine, and Kevin Novan, "Emissions, transmission, and the environmental value of renewable energy," *American Economic Journal: Economic Policy*, 2021, *13* (2), 241–272.

Fielden, Sandy, "Coast Bound Train — The Future of Crude By Rail to the West Coast Part 2," *RBN Energy Blog*, 2013. [Online; `https://rbnenergy.com/coast-bound-train-the-future-of-crude-by-rail-to-the-west-coast-part-2`; accessed 21 February, 2017].

_ , "Crude Loves Rocking Rail – Plains, Enbridge And Global Terminals In The Bakken," 2018. [Online; `https://rbnenergy.com/plains-enbridge-and-global-terminals-in-the-bakken`; accessed 21 February, 2017].

Frittelli, John, Paul W. Parfomak, Jonathan L. Ramseur, Anthony Andrews, Robert Pirog, and Michael Ratner, "U.S. Rail Transportation of Crude Oil: Background and Issues for Congress," Congressional Research Service report R43390 2014.

Funk, Josh, "Rule allowing rail shipments of LNG will be put on hold to allow more study of safety concerns," *Associated Press*, September 2023.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, Zhenhao Tan, Stephen Utkus, and Xiao Xu, "Four Facts About ESG Beliefs and Investor Portfolios," April 2023. NBER working paper 31114.

Glaeser, Edward L., "Infrastructure and Urban Form," December 2020. NBER working paper 28287.

Gonzales, Luis E., Koichiro Ito, and Mar Reguant, "The Investment Effects of Market Integration: Evidence from Renewable Energy Expansion in Chile," *Econometrica*, 2023, *91* (5), 1659–1693.

Gordon, Meghan, "Dakota Access, ETCO Oil Pipelines to Start Interstate Servce May 14," *Platts News*, April 2017.

Hall, Robert E., "Measuring Factor Adjustment Costs," *Quarterly Journal of Economics*, 2004, *119* (3), 899–927.

Harstad, Bård, "Buy Coal! A Case for Supply-Side Environmental Policy," *Journal of Political Economy*, 2012, *120* (1), 77–115.

ICF, "Economic Impacts of a Dakota Access Pipeline Shutdown," Technical Report September 2020.

Interagency Working Group on the Social Cost of Carbon, "Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866," May 2013.

International Energy Agency, "World Energy Outlook 2019," 2019.

Kellogg, Ryan, "Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch," *Quarterly Journal of Economics*, 2011, *126* (4), 1961–2004.

_ , "The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling," *American Economic Review*, 2014, *104* (6), 1698–1734.

McClure, Robert McClure, "Activist action thwarts fossil fuel projects in Pacific Northwest," 2021. [Online; `https://www.registerguard.com/story/news/2021/01/25/activists-thwart-pacific-northwest-fossil-fuel-projects/4207658001/`; accessed 21 August, 2023].

Muller, Nicholas Z., "Boosting GDP Growth by Accounting for the Environment," *Science*, 2014, *345*, 873–874.

Newell, Richard G. and Brian C. Prest, "The Unconventional Supply Boom: Aggregate Price Response from Microdata," *Energy Journal*, 2019, *40* (3), 1–30.

Newey, Whitney K. and Kenneth D. West, "Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 1994, *61* (4), 631–653.

North Dakota Pipeline Authority, "Governor's Pipeline Summit," June 2014. Online; `https://ndpipelines.files.wordpress.com/2012/04/kringstad-pipeline-summit-6-24-2014.pdf`; accessed 13 June, 2018.

__ , "Monthly Update Slide Decks and US Williston Basin Crude Oil Export Options," Online; `https://northdakotapipelines.com`; accessed on 11 August, 2017, 24 August, 2017, and 28 January, 2023.

__ , "North Dakota Midstream Update," July 2023. Online; `https://ndpipelines.files.wordpress.com/2023/07/kringstad-nd-ogrp-7-21-23.pdf`; accessed 25 August, 2023.

Pástor, Ľuboš and Robert F. Stambaugh, "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy*, 2003, *111*, 642–685.

Prest, Brian C., "Supply-side Reforms to Oil and Gas Production on Federal Lands: Modeling the Implications for CO_2 Emissions, Federal Revenues, and Leakage," *Journal of the Association of Environmental and Resource Economists*, 2022, *9*, 681–720.

__ **and James H. Stock**, "Climate royalty surcharges," *Journal of Environmental Economics and Management*, 2023, *120*, 102844.

__ , **Harrison Fell, Deborah Gordon, and TJ Conway**, "Estimating the Emissions Reductions from Supply-side Fossil Fuel Interventions," Resources for the Future working paper 23-11, 2023.

Quandl, "Brent Futures Prices," Online; `https://www.quandl.com/c/futures`; accessed on 16 May, 2017.

S&P Global, "Platts Oilgram Price Report," Accessed via Factiva on 6 Feb, 2018 and 6 April, 2021.

Tabuchi, Hiroko and Brad Plumer, "Is This the End of New Pipelines?," *The New York Times*, July 2020.

U.S. Department of Transportation, "Railroad Energy Intensity and Criteria Air Pollutant Emissions," report DOT/FRA/ORD-18/34 2018.

U.S. Energy Information Administration, "Drilling Productivity Report," https://www.eia.gov/petroleum/drilling/xls/dpr-data.xlsx; accessed 20 May, 2021.

— , "Petroleum and Other Liquids Data," https://www.eia.gov/energyexplained/oil-and-petroleum-products/data/US-tight-oil-production.xlsx and https://www.eia.gov/dnav/pet/PET_MOVE_RAILNA_A_EPC0_RAIL_MBBL_M.htm; accessed 26 April, 2021 and 17 March, 2021.

U.S. Environmental Protection Agency, "Emissions Factors for Greenhouse Gas Inventories," 2022. https://www.epa.gov/system/files/documents/2022-04/ghg_emission_factors_hub.pdf; accessed 10 March, 2023.

— , "EPA External Review Draft of Report on the Social Cost of Greenhouse Gases: Estimates Incorporating Recent Scientific Advances," https://www.epa.gov/system/files/documents/2022-11/epa_scghg_report_draft_0.pdf September 2022.

— , "Emissions and Generation Resource Integrated Database," https://www.epa.gov/egrid/summary-data and https://www.epa.gov/egrid/historical-egrid-data; accessed 9 April, 2023.

Weisbach, David A., Samuel Kortum, Michael Wang, and Yujia Yao, "Trade, Leakage, and the Design of a Carbon Tax," *Environmental and Energy Policy and the Economy*, 2023, *4* (1), 43–90.

Wilson, Miranda, "FERC retreats on gas policies as chair pursues clarity," 2022. [Online; https://www.eenews.net/articles/ferc-retreats-on-gas-policies-as-chair-pursues-clarity/; accessed 21 August, 2023].

Zhang, Ruohao, Huan Li, Neha Khanna, Alan J. Krupnick, Elaine L. Hill, and Daniel M. Sullivan, "Air Quality Impacts of Shale Gas Development in Pennsylvania," *Journal of the Association of Environmental and Resource Economists*, 2023, *10*, 447–486.

Online appendix for "Environmental Consequences of Hydrocarbon Infrastructure Policy"

A   Additional figures and tables

Figure A.1: Map of EIA Petroleum Administration for Defense Districts (PADDs)
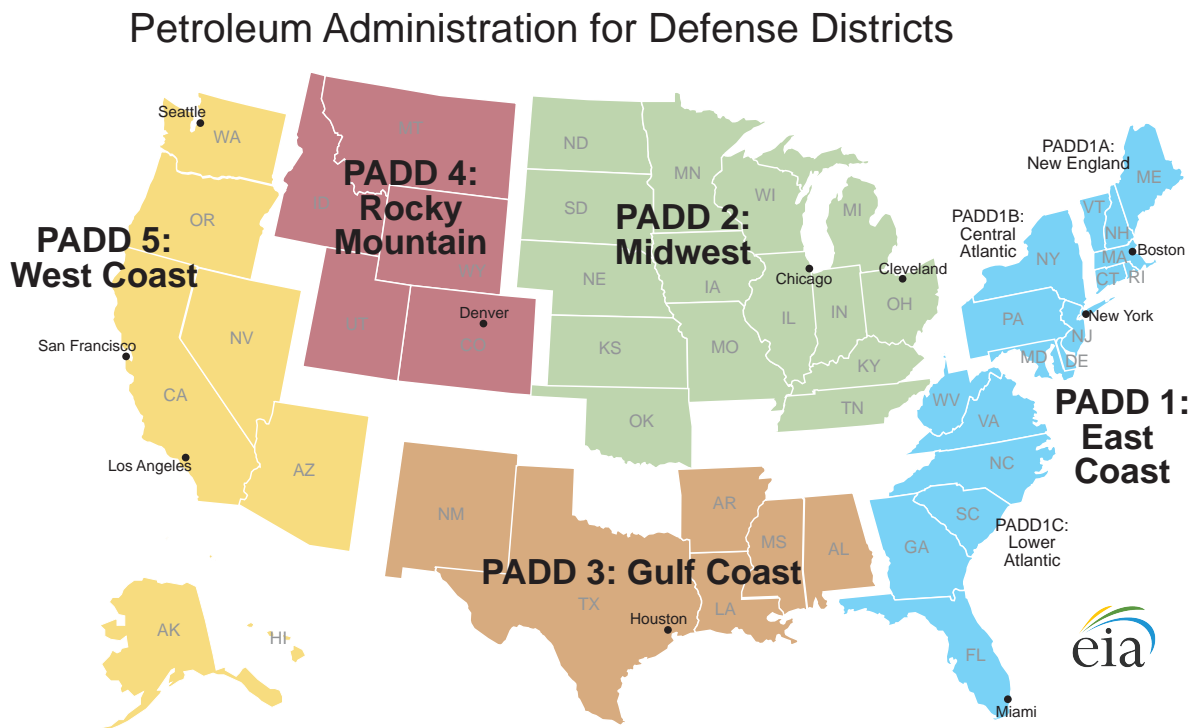
Source: U.S. Energy Information Administration (2023)

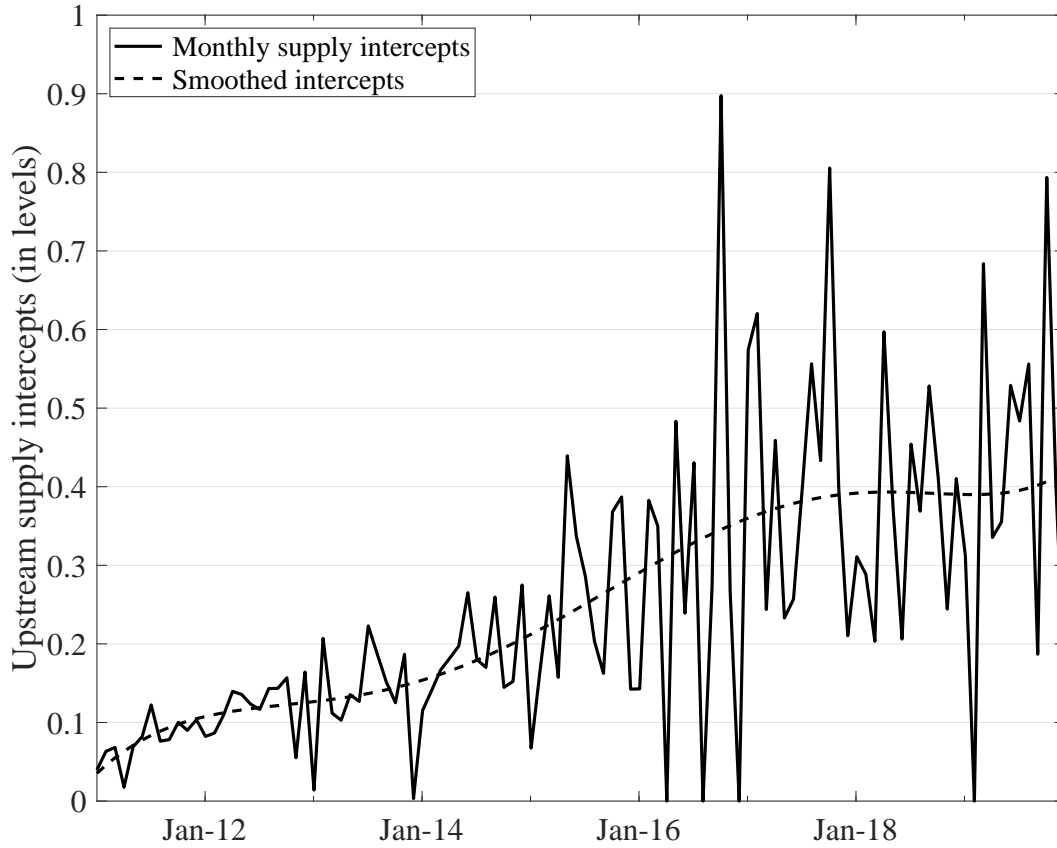**Figure A.2:** Upstream supply function productivity intercepts

Figure shows the imputed intercepts $\exp(\theta_t)$ of upstream supply equation 7. Units are mbbl/d, with price entering equation 7 in units of \$/bbl. The smoothed intercepts are the result of a sixth-degree polynomial fit to the imputed intercepts. There are 4 months in which the imputed intercepts have a value of zero. These zeros arise because in those months, the quantity of decayed "old" production exceeds the month's total production reported in the data.

Figure A.3: Brent price volatility estimates at horizons up to 13 years
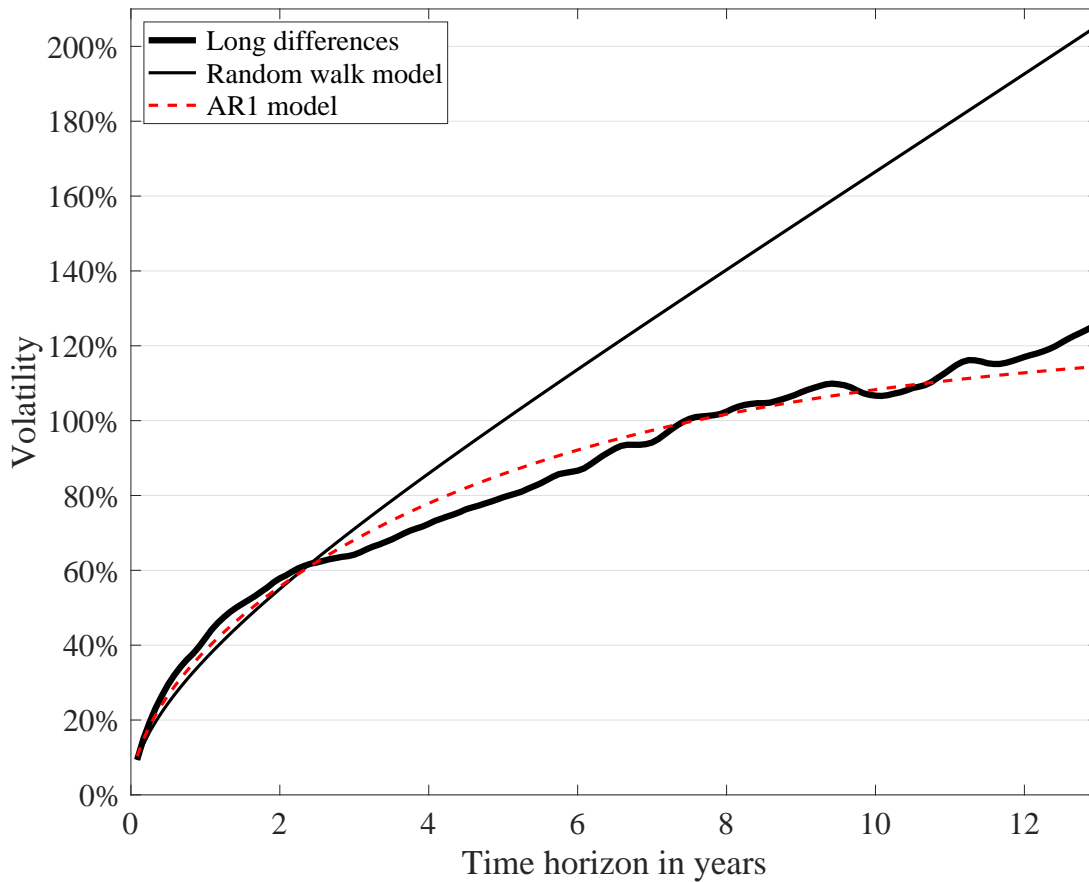
Note: The long differenced volatility at each horizon t is calculated as the standard deviation of t-month differences in historic logged (real June, 2014) Brent crude prices. The random walk model extrapolates the historic one-month Brent volatility to longer time horizons by multiplying the one-month volatility by $\sqrt{t}$. The AR(1) model is the best fit of equation 14 to the series of t-month historic volatilities. Volatilities in percent are calculated for each horizon by exponentiating the standard deviation, subtracting one, and multiplying by 100.

Figure A.4: NDPA production forecasts as of June, 2014
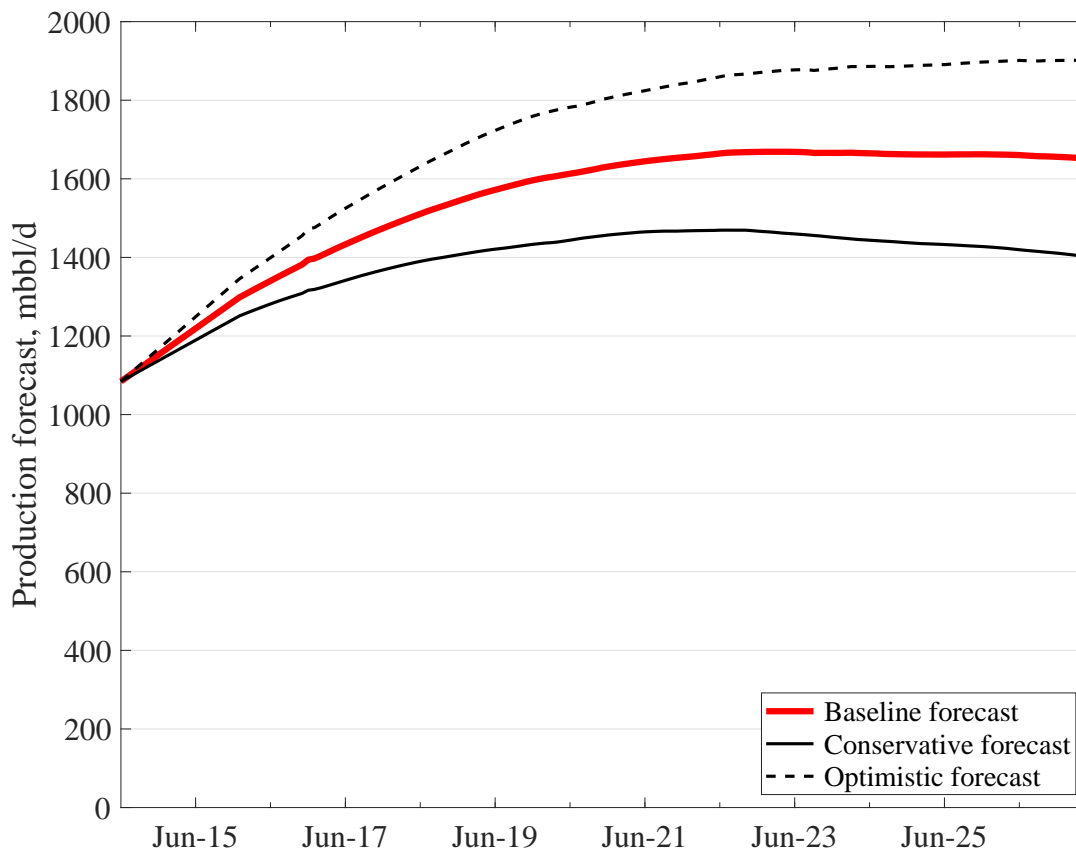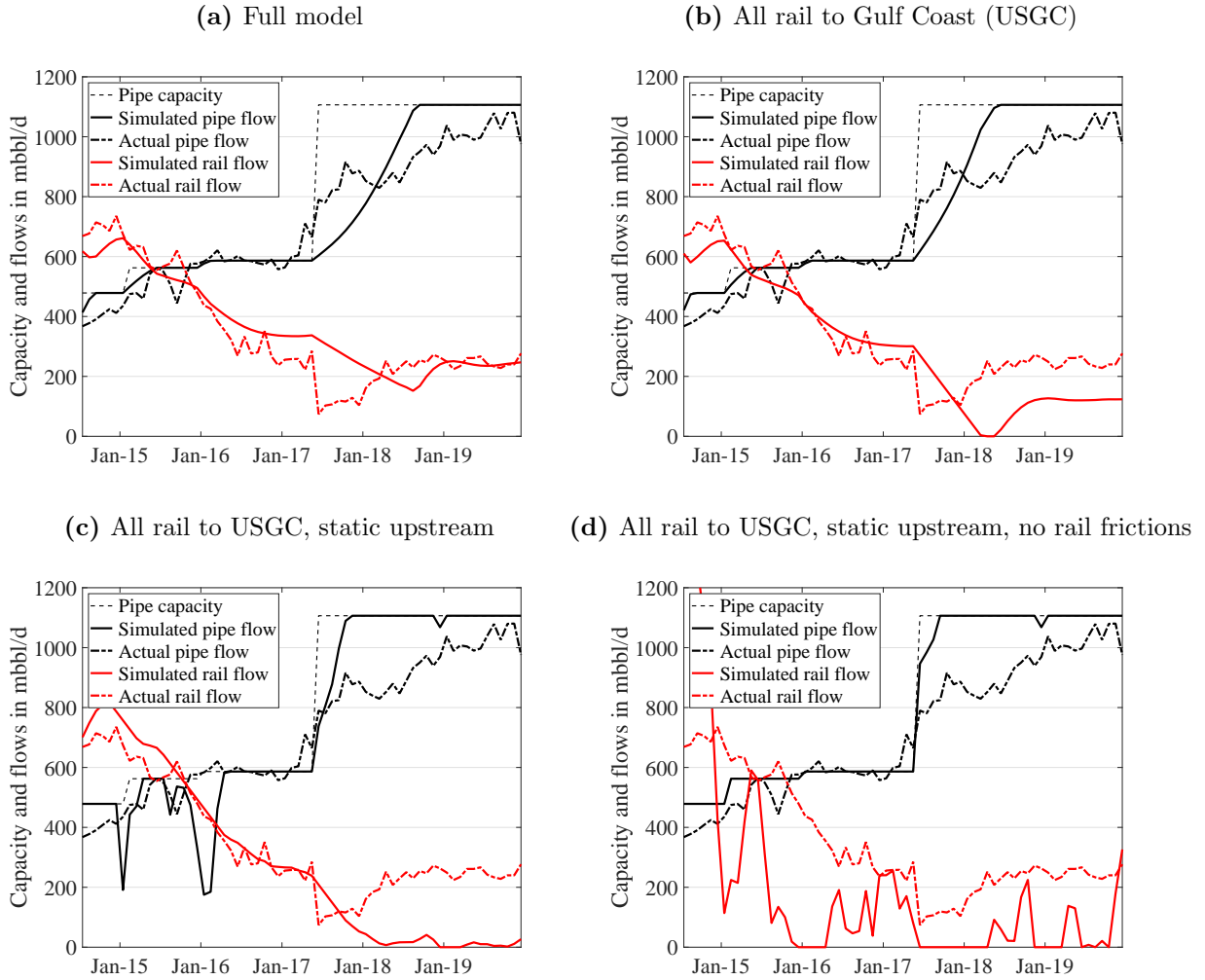

Figure shows the NDPA's production forecasts as of June, 2014 (North Dakota Pipeline Authority, 2014). The NDPA publishes a figure showing the production forecast time series but not the underlying data. We digitized the NDPA figure and then evaluated production for each month $t \in [1, 156]$ by applying a local linear smoother to the digitized points (Calonico, Cattaneo and Farrell, 2019). Our digitized forecast starts in January, 2016; we linearly interpolate this forecast back to July, 2014 using realized production from June, 2014. The conservative forecast is the NDPA's "case 2" forecast. We construct the optimistic forecast by adding, in each month, the difference between the baseline and conservative forecast to the baseline forecast.

Figure A.5: Simulated and actual pipeline and rail flows, using the full model and alternatives

(a) Full model

(b) All rail to Gulf Coast (USGC)

(c) All rail to USGC, static upstream

(d) All rail to USGC, static upstream, no rail frictions

"Full model" in panel (a) shows forward simulations using the full model discussed in section 3, starting from initial conditions as of June, 2014. This panel is identical to figure 3 in the main text. Panels (b), (c), and (d) present simulations from simplifications of the full model. "All rail to USGC" forces all crude-by-rail to flow to the USGC; this model divides the friction parameter γ by 3 so that it is comparable to the multiple-destination full model. "All rail to USGC, static upstream" additionally simplifies the upstream model so that production Q_t each period is a constant elasticity function of the upstream price P_{ut}, with the elasticity equal to the sum of the coefficients estimated in equation 8; i.e., 1.32. Finally, "All rail to USGC, static upstream, no rail frictions" additionally sets $\gamma = 0$; simulated rail flows in this model in 2014 exceed 1400 mbbl/d. Actual flows and pipeline capacity are identical across panels. All simulations use realized downstream prices through 2019.

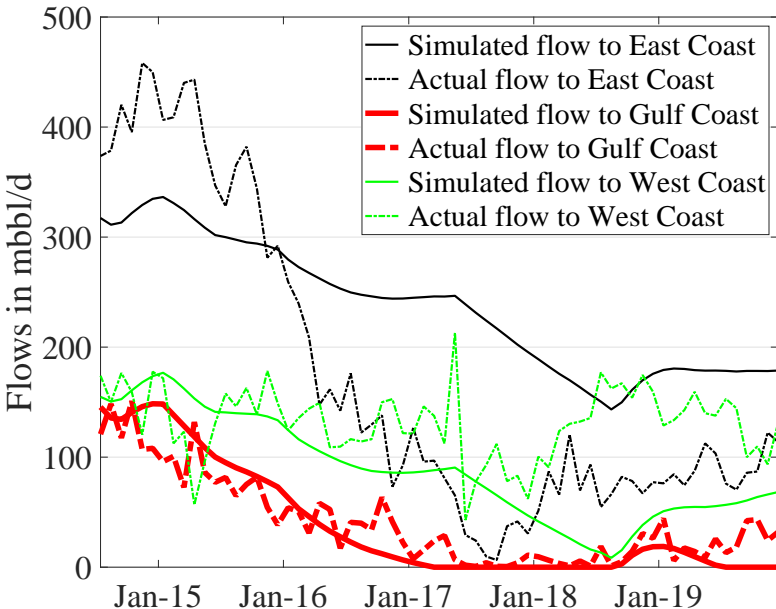**Figure A.6:** Simulated and actual crude-by-rail flows to each destination
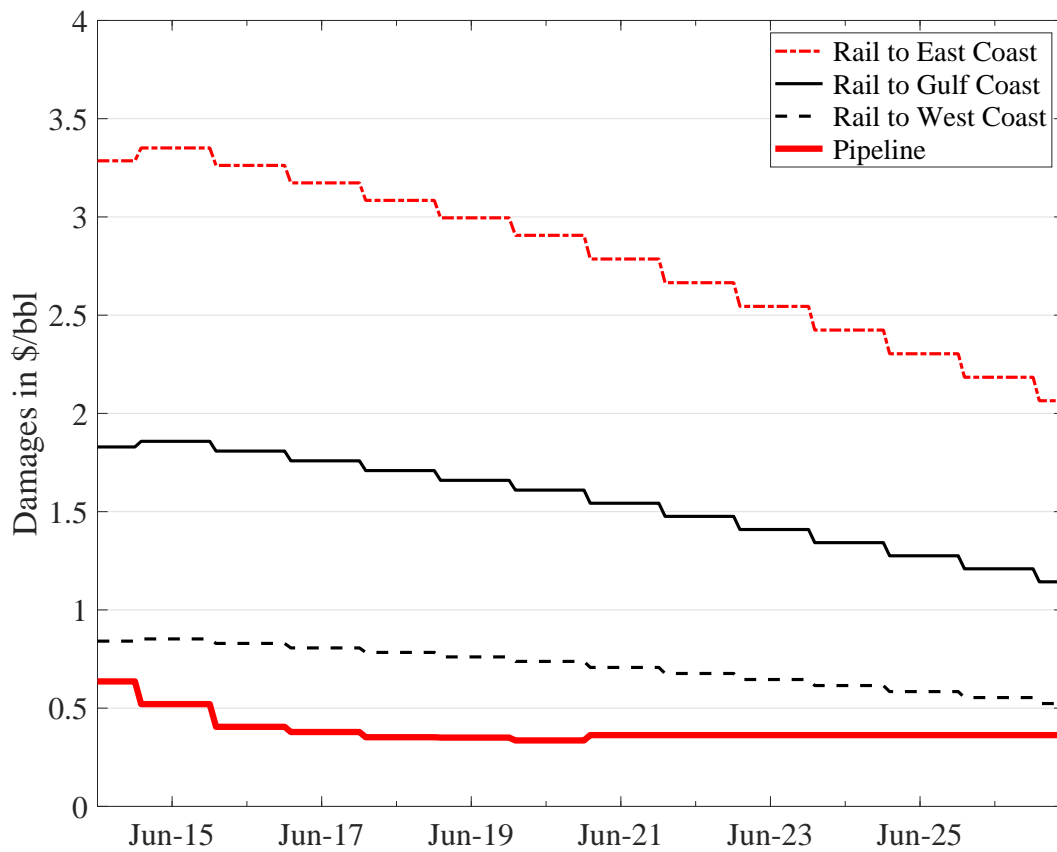

Figure shows forward simulations using the full model discussed in section 3, starting from initial conditions as of June, 2014, and using realized downstream prices through 2019.

Figure A.7: Estimated local air pollution damages from pipeline and rail transportation of Bakken crude

Estimates computed from Clay, Jha, Muller and Walsh (2019), U.S. Environmental Protection Agency (2023), and U.S. Department of Transportation (2018). See section 5.2.1 and appendix B.5 for details.

Figure A.8: Simulated pipeline and crude-by-rail flows, with and without an upstream production tax, DAPL capacity held fixed

Figure shows forward simulations using the full model discussed in section 3, starting from initial conditions as of June, 2014 and using realized downstream prices through 2019, both with and without the imposition of a $3.68/bbl upstream production tax ($8.52/tonne carbon tax) starting in June, 2017. Simulated flows without the tax are are identical to those shown in figure 3. See section 6.1 for discussion.

Figure A.9: Simulated pipeline and crude-by-rail flows, with and without an upstream production tax, DAPL capacity responds to the tax
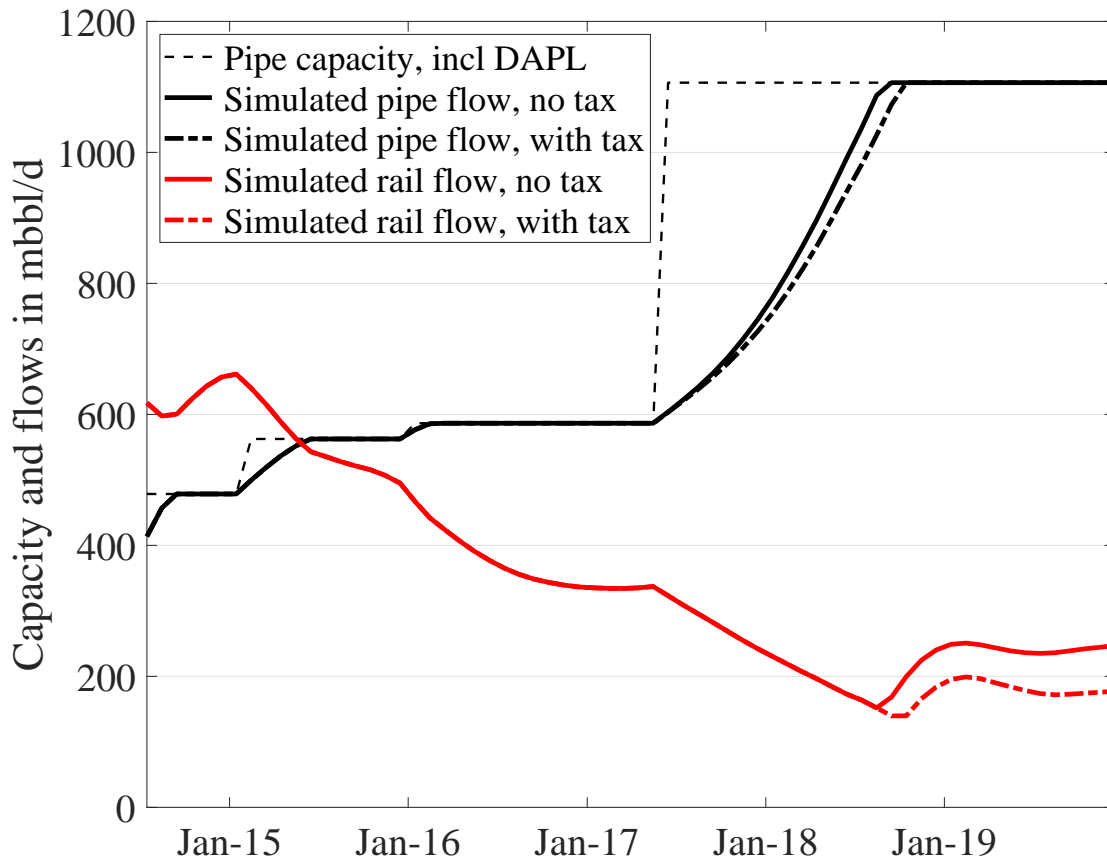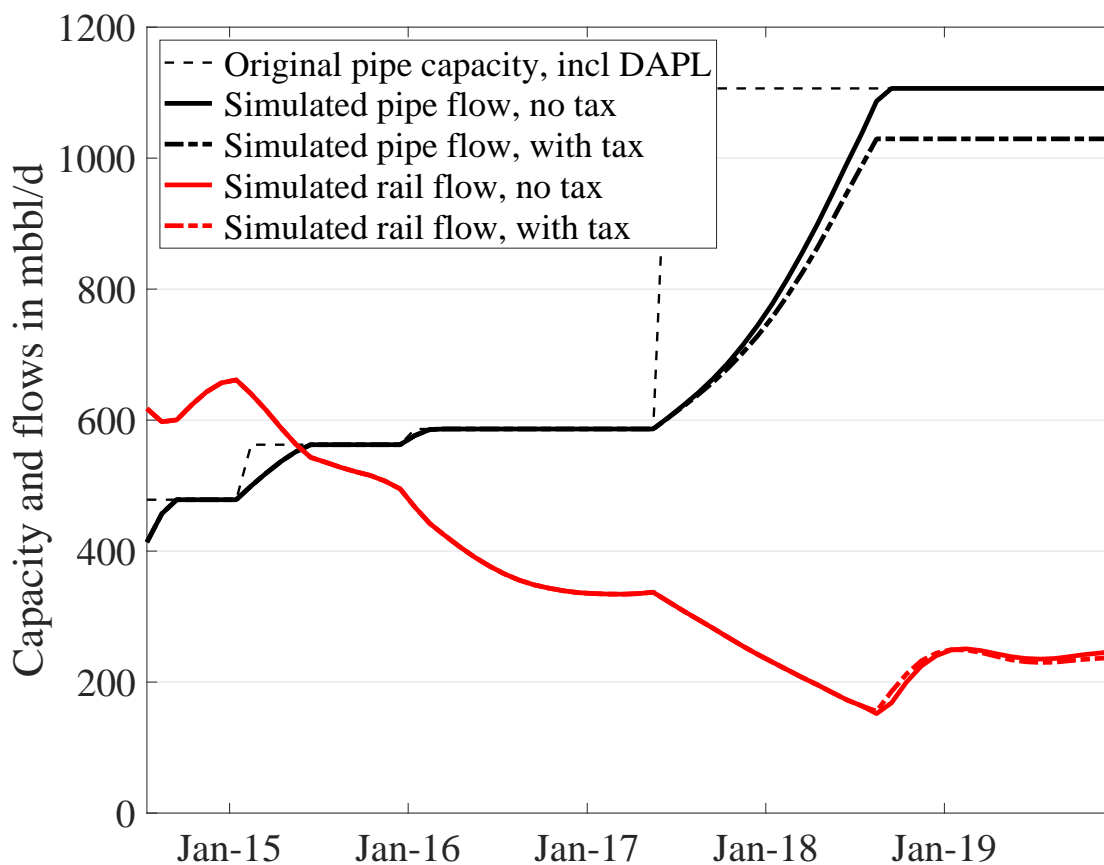

Figure shows forward simulations using the full model discussed in section 3, starting from initial conditions as of June, 2014 and using realized downstream prices through 2019, both with and without the imposition of a $3.24/bbl upstream production tax ($7.51/tonne carbon tax) starting in June, 2017. DAPL's equilibrium capacity investment anticipates this tax. Simulated flows without the tax are are identical to those shown in figure 3. See section 6.1 for discussion.

Table A.1: Expected return to DAPL shippers, using the full model and alternatives

	Expected return
Full model	$6.17/bbl
Random walk belief for P_G	$7.59/bbl
All rail to Gulf Coast (USGC)	$7.73/bbl
All rail to USGC, static upstream	$8.93/bbl
All rail to USGC, static upstream, no rail frictions	$6.09/bbl

Expected returns are calculated as the right-hand-side of equation 13, starting from initial conditions as of June, 2014. "Full model" uses the full model discussed in section 3. "Random walk belief for P_G" assumes shippers have a random walk belief regarding the future Gulf Coast oil price P_g rather than the AR(1) belief estimated in section 3.5. "All rail to USGC" returns to the AR(1) price belief but forces all crude-by-rail to flow to the USGC; this model divides the friction parameter γ by 3 so that it is comparable to the multiple-destination full model. "All rail to USGC, static upstream" additionally simplifies the upstream model so that production Q_t each period is a constant elasticity function of the upstream price P_{ut}, with the elasticity equal to the sum of the coefficients estimated in equation 8; i.e., 1.32. Finally, "All rail to USGC, static upstream, no rail frictions" additionally sets $\gamma = 0$. The actual DAPL committed tariffs were $5.50–$6.25/bbl (Gordon, 2017). All returns are in real June, 2014 dollars.

Table A.2: Expected return to DAPL shippers; alternative specifications

	Expected returns, $/bbl			
	$K_d = 320$	$K_d = 450$	$K_d = 520$	$K_d = 570$
Full model	$7.15	$6.50	$6.17	$5.96
Full model, plus supply uncertainty	$7.14	$6.48	$6.16	$5.95

Expected returns are calculated as the right-hand-side of equation (13), starting from initial conditions as of June, 2014. "Full model" uses the full model discussed in section 3. "Full model, plus supply uncertainty" endows prospective shippers with uncertainty over future Bakken oil supply, per the discussion in appendix B.4. Each column provides expected returns with an assumed DAPL capacity commitment of 320, 450, 520, or 570 mbbl/d. The actual DAPL committed tariffs were $5.50–$6.25/bbl (Gordon, 2017). All returns are in real June, 2014 dollars.

Table A.3: Simulated oil flow changes from foreclosing DAPL, in expectation, with alternative specifications

	Δ pipe flow mbbl/d	Δ rail flow mbbl/d	Δ production volume mbbl/d	Δ rail as % of Δ pipe
Baseline model	-306	248	-58	-81%
$K_d = 320$mbbl/d	-208	169	-39	-81%
$K_d = 450$mbbl/d	-274	222	-52	-81%
$K_d = 570$mbbl/d	-328	265	-63	-81%
Random walk belief	-289	236	-53	-82%
Supply uncertainty	-305	247	-58	-81%

All expectations are taken as of June, 2014 and are averages over 10000 Monte Carlo draws of possible downstream price paths. For each simulated price path, we compute average discounted volumes (pipeline, rail, local, total) during DAPL shippers' ten-year commitment period (June, 2017 - May, 2027). The first row ("Baseline model") corresponds to table 3 in the main text. Rows 2 through 4 use alternative values for DAPL's committed capacity (the baseline model uses $K_d = 520$ mbbl/d). Row 5 assumes shippers have a random walk belief regarding the future Gulf Coast oil price P_g rather than the AR(1) belief estimated in section 3.5. Row 6 endows prospective shippers with uncertainty over future Bakken oil supply, per the discussion in appendix B.4.

B Detail on data sources and estimation

B.1 Oil price and crude-by-rail flow data

The LLS, Brent, and ANS prices that we use are given by Bloomberg keys USCRLLSS, EUCRBRDT, and USCRANSW, respectively. ANS crude is 32 API and 0.96% sulfur (S&P Global, 2017), making it heavier and more sour than Bakken crude (43 API and 0.07% sulfur). To account for this grade difference, we compute the average price difference between Light Louisiana Sweet and Heavy Louisiana Sweet using their full price history from 1988–2019, and we then add this value ($0.62/bbl) to the ANS price series.

The S&P Global (2021) "Bakken local" price series begins in April, 2014, so we impute earlier Bakken local prices using prices at the Clearbrook, MN hub, also obtained from S&P Global (2021). We compute an average Clearbrook premium of $2.21/bbl over the Bakken local price from April, 2014 through 2019, when both price series were available. To impute Bakken local prices before April, 2014, we subtract this value from the Clearbrook price series (which goes back to May, 2010).

To deflate all price data, we use the Bureau of Labor Statistics' Consumer Price Index for all goods less energy, all urban consumers, and not seasonally adjusted (Bureau of Labor Statistics, 2023). The CPI series ID is CUUR0000SA0LE.

The destination-specific crude-by-rail flow data that we obtained from U.S. Energy Information Administration (2021) define regions using Petroleum Administration for Defense Districts (PADDs). The Midwest is PADD 2, and we use data on rail flows from PADD 2 to three destination PADDs: we define East Coast destinations as PADD 1, Gulf Coast destinations as PADD 3, and West Coast destinations as PADD 5. A map of PADDs is presented in appendix figure A.1. Note that the EIA reports that small volumes of crude-by-rail stay within the Midwest; we allocate these volumes proportionally to PADDs 1, 3, and 5 rather than include the Midwest as a fourth crude-by-rail destination. Shipping to the Midwest is dominated by shipping to the coasts due to both the depressed West Texas Intermediate crude oil price at Cushing, OK during most of 2010–2015 and the presence of pre-existing pipeline connections to Midwest destinations.

Due to sampling error in the STB waybill sample (which the EIA uses to estimate destination-specific rail flows), the total rail volumes reported by the EIA do not perfectly match what we obtain from the NDPA. We therefore estimate destination-by-month rail volumes by: (1) using the EIA data to compute the fraction of rail shipments going to each destination each month; and (2) multiplying these fractions by total Bakken rail exports per the NDPA.

B.2 Crude-by-rail costs

Per the discussion in section 4 of the main text, our model does a satisfactory job of matching aggregate crude-by-rail flows in our validation exercise but less well at matching destination-specific flows. Increasing the cost of rail shipments to the East Coast (r_E) by \$1.50/bbl and decreasing the cost of rail shipments to the West Coast (r_W) by \$1.50/bbl yields a better match to East Coast versus West Coast flows while holding total simulated rail transport roughly constant. In this alternative specification, we compute that the expected return to DAPL shippers is \$6.18/bbl, essentially unchanged from the value of \$6.17/bbl reported in section 4. When we use this alternative specification to evaluate the counterfactual in which DAPL's construction is foreclosed, we find that 81% of the decrease in expected pipeline flow would move by rail instead; this value is again essentially unchanged from its corresponding value in the main text (section 5.1). Finally, the total (private surplus + local pollution damages) cost per tonne of CO_2 abated that we obtain with this model is \$43. This value is slightly smaller than the value of \$45 reported in table 5 in the main text; the difference arises because crude-by-rail shipments to the West Coast are associated with less local pollution damage than shipments to the East Coast (recall table 4).

B.3 Construction of data series on production from new wells

In each month t, we compute the share of production from new wells, s_{nt}, as the DPR's estimate of new Bakken production per rig, times the DPR's estimate of the number of active Bakken rigs, divided by the DPR's estimate of total Bakken production. The DPR estimates of total Bakken production differ from the "Tight oil production estimates by play" data series (which we use for Q_t) because the former are short-term production estimates, while the latter use state administrative data and are published with a lag of several months. Finally, we compute $Q_{nt} = s_{nt} * Q_t$.

B.4 Shippers' beliefs about future oil prices and upstream supply

Beliefs about future Gulf Coast prices

We estimate the parameters ϕ_0, ϕ_1, and σ_G^2 from equation 14 to fit measures of the long-run expected Gulf Coast price $\mathbb{E}[P_G]$ and the evolution of its variance over time, conditional on P_{G0}. For $\mathbb{E}[P_G]$, we use the June, 2014 three-year Brent crude future price of \$93.47/bbl. This value is based on a nominal futures price of \$99.19/bbl from Quandl (2017), adjusted for 2.0% inflation per Federal Reserve Bank of Atlanta (2014). Setting the long-run expected

price equal to \$93.47/bbl pins down the parameter ϕ_0 in equation 14 as a function of ϕ_1 and σ_G^2 via the formula $\phi_0 = (1 - \phi_1) \log \mathbb{E}[P_G] - \sigma_G^2/2$. Given $\mathbb{E}[P_G] = \$93.47/\text{bbl}$ and our AR(1) estimates of ϕ_1 and σ_G^2, this calculation yields $\phi_0 = 0.0293$. For the first 36 months of our simulations, we use a slightly smaller value of $\phi_0 = 0.0264$ so that the expected price three years from June, 2014 (when the LLS spot price was \$108.21/bbl) is \$93.47/bbl.

Given the known initial price at $t = 0$, the variance of P_{Gt} in the model given by equation (14) is equal to $\sigma_G^2(1 - \phi_1^{2t})/(1 - \phi_1^2)$. The variance of the future oil price therefore increases with the time horizon for $\phi_1 > 0$. To estimate ϕ_1 and σ_G^2, we fit this variance formula to the historic variance of log oil prices over time horizons ranging from 1 month to 13 years. We calculate historic variance at each horizon t by taking the variance of t-month differences in historic logged Brent prices.

Appendix figure A.3 shows the price volatilities that we compute via this long differences approach, using Brent price data from May, 1996 (the first observation for which a 13-year lag is available in the data) through 2019.[1] We find that uncertainty over the future price of Brent increases substantially over the 1-month to 13-year horizon, from a volatility of 9.4% at one month, to 64.8% at 37 months, and to 126% at 13 years.

We fit our AR(1) model given by equation 14 to these volatilities using a minimum distance estimator, obtaining estimates of ϕ_1 and σ_G equal to 0.9925 and 0.098, respectively. This estimator minimizes the sum, over $t \in [1, 156]$, of the squared differences between the AR(1) model's log variance at horizon t and the log long-differenced variance at horizon t. Figure A.3 shows how this estimated AR(1) process smoothly fits the estimated long-differenced volatilities. The figure also shows an alternative model that assumes a random walk ($\phi_1 = 1$) and sets σ_G equal to the historic one-month volatility of 9.4%. This random walk assumption produces a much greater price variance over long time horizons.

Expected upstream production

The NDPA monthly production forecast (North Dakota Pipeline Authority, 2014) provides monthly expected Bakken production volumes throughout the ten-year pipeline contract (and beyond), under the prevailing EIA oil price forecast (personal communication from Justin J. Kringstad at NDPA (June, 2018)). The NDPA publishes a figure showing the production forecast time series but not the underlying data. We digitized the NDPA figure and then evaluated production for each month $t \in [1, 156]$ by applying a local linear smoother to the digitized points (Calonico et al., 2019). Our digitized forecast starts in January, 2016; we linearly interpolate this forecast back to July, 2014 using realized production from June,

---

[1]The volatility estimates are similar if we use only data through May, 2014, the last month prior to DAPL's open season.

2014.

In an alternative specification, we allow for uncertainty in upstream supply by letting the supply intercept be stochastic, using a conservative ("case 2") production forecast from the NDPA that is, on average, 188 mbbl/d lower than the expected production path. We also construct an "optimistic" forecast that is symmetric to this conservative forecast. That is, in each month our optimistic forecast exceeds the baseline forecast by the same amount that the conservative forecast falls short. Then, in alternative specifications of our model, we model prospective pipeline shippers as having beliefs that the supply intercept each period is stochastic, with a probability of 1/3 assigned to each production path. These probabilities are arbitrary; the NDPA does not assign probabilities to its production cases.

DAPL capacity commitment

There are two reasons why it is difficult to be certain of the capacity to which Bakken pipeline shippers had committed in June, 2014. First, the official June, 2014 announcement of DAPL shippers' commitments stated a volume of 320 mbbl/d (Energy Transfer Partners LP, 2014a). However, by September, 2014 DAPL announced executed precedent agreements with shippers supporting a capacity of 450 mbbl/d (Energy Transfer Partners LP, 2014b), and then a supplemental open season in early 2017 supported an increase in capacity to the constructed value of 520 mbbl/d (Energy Transfer Partners LP, 2017). Second, back in 2013 a competing project, the Sandpiper Pipeline, had secured shipper commitments for a 225 mbbl/d line from the Bakken to Lake Superior (Enbridge Energy Partners LP, 2015). This project was beset by environmental permitting delays in Minnesota and was postponed indefinitely in September, 2016 after Enbridge (Sandpiper's main owner) and Marathon (Sandpiper's "anchor shipper") invested in DAPL and canceled their Sandpiper shipping agreement (Shaffer, 2014; Marathon Petroleum Corporation, 2016). It is not clear to what extent Sandpiper's demise was foreseen in June, 2014, when shippers initially committed to DAPL. The base case specification of our model assumes a committed capacity of 520 mbbl/d, equal to the DAPL capacity actually constructed and approximately equal to the total DAPL plus Sandpiper capacity that had been announced by June, 2014 (320 mbbl/d for DAPL and 225 mbbl/d for Sandpiper). As sensitivities, we also examine results based on assumed capacities of 320, 450, and 570 mbbl/d (DAPL was built with the ability to expand to 570 mbbl/d, and it initiated an open season on the incremental 50 mbbl/d in March, 2018 (Energy Transfer Partners LP 2018)).

B.5   Pollution emission factors and costs per barrel

To adjust Clay et al.'s (2019) pipeline transportation emissions factors for changes in the electric generation fleet, we use data from eGrid (U.S. Environmental Protection Agency, 2023). These data provide average emissions factors (for CO_2, NO_x, and SO_x) by eGrid subregion for each year in 2018–2021, and for even-numbered years before 2018. For each year, we compute weighted average emissions factors over the four eGrid regions that DAPL passes through. eGrid subregions roughly correspond to North American Electric Reliability Corporation reliability assessment areas. We assign weights of 1/2 to MROW (upper Midwest), 1/6 to SRMW (IL and MO), 1/6 to SRTV (Tennessee Valley), and 1/6 to SRMV (Arkansas and Louisiana). We interpolate emissions factors for odd-numbered years before 2018, and then we compute per-barrel emissions from pipeline transport for each year in 2014–2021 by multiplying the values from Clay et al. (2019) by the ratio of the target year's eGrid emissions factors to eGrid's 2011 emissions factors. When modeling 2022 onward, we hold the emissions factors fixed at their 2021 values. As shown in appendix figure A.7, emissions factors for these eGrid regions have been roughly constant since 2017. Were emissions factors to fall after 2021, our analysis would be an under-estimate of the increase in local pollution damage from foreclosing DAPL's construction.

For rail emissions, we re-scale the damage estimates from Clay et al. (2019) using locomotive emissions factors from U.S. Department of Transportation (2018), which projects the locomotive fleet-weighted emissions of NO_x, VOCs, and PM per gallon of fuel out to 2040. For each year and route of our analysis, we multiply the per-bbl-mile rail damages from Clay et al. (2019) by the U.S. Department of Transportation's (2018) emissions factors for the target year and divide by the 2014 emissions factors. We assume that CO_2 emissions per bbl-mile are constant over time.

Finally, we note that Clay et al. (2019) provides damage valuations for 2014 using the Environmental Protection Agency's 2014 VSL of \$8.5 million. We therefore treat their estimates as being in real 2014 dollars and do not adjust them for inflation.

C   Existence and uniqueness of each period's equilibrium oil flows and upstream price

This appendix proves that the equilibrium defined by the four conditions enumerated in section 3.4 exists and is unique.

C.1 Existence

The equilibrium involves only a single endogenous price, P_{ut}. Given P_{ut}, upstream production Q_t and rail flows to each destination Q_{rit} are determined by the functions in conditions 1 and 3. Pipeline flows Q_{pt} are determined by the correspondence in condition 2. Define the net supply correspondence $NS(P_{ut})$ as the left-hand-side of condition 4 (upstream supply) minus the right-hand-side of condition 4 (local and transportation demand).

For $P_{ut} \neq P_{Gt}$, $NS(P_{ut})$ is a continuous and increasing function, since: (1) Q_t is a continuous and increasing function of P_{ut}, and strictly increasing for $P_{ut} > 0$;[2] (2) the Q_{rit} are continuous and weakly decreasing functions of P_{ut}, and strictly decreasing whenever $Q_{rit} > 0$; and (3) local consumption Q_L and pipeline flows Q_{pt} are constant and single-valued. It is always possible to find a P_{ut} sufficiently large that $NS(P_{ut}) > 0$, since upstream supply $Q(P_{ut})$ is unbounded above. Likewise, it is always possible to find a P_{ut} sufficiently small that $NS(P_{ut}) < 0$, since the rail flows $Q_{ri}(P_{ut})$ are unbounded above. Finally, at $P_{ut} = P_{Gt}$, the correspondence $NS(P_{ut})$ is upper hemicontinuous, with $NS(P_{ut})$ forming a connected set.

A candidate value for P_{ut} is an equilibrium if $0 \in NS(P_{ut})$. We now prove constructively that such a P_{ut} exists. We consider three exhaustive cases based on the elements of $NS(P_{ut})$ at $P_{ut} = P_{Gt}$.

First, if all elements of $NS(P_{Gt})$ are strictly less than zero, then by the upper hemicontinuity of $NS(P_{ut})$ at P_{Gt}, the continuity of $NS(P_{ut})$ for $P_{ut} > P_{Gt}$, and the existence of $\overline{P_{ut}} > P_{Gt}$ such that $NS(\overline{P_{ut}}) > 0$, the intermediate value theorem implies that there must exist a $P_{ut} > P_{Gt}$ such that $NS(P_{ut}) = 0$.

Second, if all elements of $NS(P_{Gt})$ are strictly greater than zero, then by the upper hemicontinuity of $NS(P_{ut})$ at P_{Gt}, the continuity of $NS(P_{ut})$ for $P_{ut} < P_{Gt}$, and the existence of $\underline{P_{ut}} < P_{Gt}$ such that $NS(\underline{P_{ut}}) < 0$, the intermediate value theorem implies that there must exist a $P_{ut} < P_{Gt}$ such that $NS(P_{ut}) = 0$.

Finally, if neither of the first two conditions hold, then the connectedness of $NS(P_{Gt})$ implies that $0 \in NS(P_{Gt})$, which implies that $P_{ut} = P_{Gt}$ is an equilibrium. Then, because the three conditions we have considered are exhaustive, an equilibrium must exist.

C.2 Uniqueness

Suppose there are two values of P_{ut}, denoted P_{1t} and P_{2t}, that are equilibria (without loss of generality, we assume $P_{1t} < P_{2t}$). We show that this conjecture leads to a contradiction. First, suppose that $P_{1t} > P_{Gt}$ and $P_{2t} > P_{Gt}$. In this case, the strict monotonicity of $NS(P_{ut})$

---

[2]P_{ut} may be negative, in which case we define new upstream supply $Q_n(P_{ut}) = 0$ so that $Q_t = \beta Q_{t-1}$.

for $P_{ut} > P_{Gt}$ implies that we cannot have both $NS(P_{1t}) = 0$ and $NS(P_{2t}) = 0$. A similar contradiction holds, with one knife-edge exception, in the case that $P_{1t} < P_{Gt}$ and $P_{2t} < P_{Gt}$.[3]

Suppose $P_{1t} < P_{Gt}$ and $P_{2t} > P_{Gt}$. With $NS(P_{1t}) = 0$, strict monotonicity of $NS(P_{ut})$ for $P_{ut} < P_{Gt}$ and the upper hemicontinuity of $NS(P_{Gt})$ imply that all elements of $NS(P_{Gt})$ are strictly greater than zero. Then, the upper hemicontinuity of $NS(P_{Gt})$ and the strict monotonicity of $NS(P_{ut})$ for $P_{ut} > P_{Gt}$ imply that $NS(P_{2t}) > 0$, so that P_{2t} is not an equilibrium.

Finally, suppose $P_{1t} < P_{Gt}$ and $P_{2t} = P_{Gt}$ (the case of $P_{1t} = P_{Gt}$ and $P_{2t} > P_{Gt}$ is similar). With $NS(P_{1t}) = 0$, strict monotonicity of $NS(P_{ut})$ for $P_{ut} < P_{Gt}$ and the upper hemicontinuity of $NS(P_{Gt})$ imply that all elements of $NS(P_{Gt})$ are strictly greater than zero. Thus, P_{2t} is not an equilibrium. By contradiction, it must therefore be the case that the equilibrium is unique.

References

Bureau of Labor Statistics, "Consumer Price Index, all urban, all items less energy, not seasonally adjusted (CUUR0000SA0LE)," https://www.bls.gov/cpi/data.htm; accessed 9 August, 2023.

Calonico, Sebastian, Matias D. Cattaneo, and Max H. Farrell, "nprobust: Nonparametric Kernel-Based Estimation and Robust Bias-Corrected Inference," *Journal of Statistical Software*, 2019, *91* (8), 1–33.

Clay, Karen, Akshaya Jha, Nicholas Muller, and Randall Walsh, "The External Costs Of Transporting Petroleum Products: Evidence From Shipments Of Crude Oil From North Dakota by Pipelines and Rail," *Energy Journal*, 2019, *40* (1), 55–72.

Enbridge Energy Partners LP, "SEC 10-K Filing," February 2015. [Online; `https://www.enbridgepartners.com/Investor-Relations/EEP/Financial-Information/Filing-Details.aspx?filingId=10082267&docId=222781`; accessed 18 March, 2018].

---

[3]The exception concerns the possibility that $P_{1t} < 0$ and $P_{2t} \leq 0$, in which case $Q_t = \beta Q_{t-1}$ (since new well production is zero). If the downstream prices P_{it} are sufficiently low, then it is possible to have a situation in which $Q_{rit} = 0 \ \forall i$. Then if $Q_t = Q_L + K_t$, there can be multiple strictly negative downstream prices that satisfy the equilibrium conditions (though oil production and flows are unique, with the pipeline at capacity and no rail flow). This situation requires that $\beta Q_{t-1} = Q_L + K_t$, which is a measure zero event. If $\beta Q_{t-1} > Q_L + K_t$, then rail flows must be strictly positive, restoring the strict monotonicity of $NS(P_{ut})$ and the uniqueness of equilibrium. If $\beta Q_{t-1} < Q_L + K_t$ while rail flows are zero, then equilibrium condition 2 requires the unique $P_{ut} = P_{Gt} > 0$.

Energy Transfer Partners LP, "Energy Transfer Announces Crude Oil Pipeline Project Connecting Bakken Supplies to Patoka, Illinois and to Gulf Coast Markets," 2014. [Online; `http://ir.energytransfer.com/phoenix.zhtml?c=106094&p=irol-newsArticle&ID=1942689`; accessed 13 September, 2017].

_ , "Energy Transfer Commences Binding Expansion Open Season for Bakken Pipeline Transport," 2014. [Online; `http://ir.energytransfer.com/phoenix.zhtml?c=106094&p=irol-newsArticle&ID=1969920`; accessed 18 March, 2018].

_ , "Energy Transfer Announces the Bakken Pipeline is in Service Transporting Domestic Crude Oil from the Bakken/Three Forks Production Areas," 2017. [Online; `http://ir.energytransfer.com/phoenix.zhtml?c=106094&p=irol-newsArticle&ID=2278014`; accessed 13 March, 2018].

_ , "Announcement of Expansion Open Season," 2018. [Online; `https://www.energytransfer.com/ops_bakken.aspx`; accessed 13 March, 2018].

Federal Reserve Bank of Atlanta, "Atlanta Fed Survey of Business Inflation Expectations," June 2014. https://www.atlantafed.org/-/media/Documents/news/pressreleases/BIEsurvey/2014/1406.pdf.

Gordon, Meghan, "Dakota Access, ETCO Oil Pipelines to Start Interstate Servce May 14," *Platts News*, April 2017.

Marathon Petroleum Corporation, "Marathon Petroleum Corporation agrees to equity participation in Bakken Pipeline system," 2016. [Online; `http://ir.marathonpetroleum.com/phoenix.zhtml?c=246631&p=irol-newsArticle_pf&ID=2192194`; accessed 18 March, 2018].

North Dakota Pipeline Authority, "Governor's Pipeline Summit," June 2014. Online; `https://ndpipelines.files.wordpress.com/2012/04/kringstad-pipeline-summit-6-24-2014.pdf`; accessed 13 June, 2018.

Quandl, "Brent Futures Prices," Online; `https://www.quandl.com/c/futures`; accessed on 16 May, 2017.

Shaffer, David, "Minnesota Regulators Order a Broader Look at Pipeline Route," *Minneapolis Star Tribune*, September 2014.

S&P Global, "Platts Methodology and Specifications Guide: Crude Oil," 2017. [Online; `https://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/Crude-oil-methodology.pdf`; accessed 18 March, 2018].

___ , "Platts Oilgram Price Report," Accessed via Factiva on 6 Feb, 2018 and 6 April, 2021.

U.S. Department of Transportation, "Railroad Energy Intensity and Criteria Air Pollutant Emissions," report DOT/FRA/ORD-18/34 2018.

U.S. Energy Information Administration, "Petroleum and Other Liquids Data," https://www.eia.gov/energyexplained/oil-and-petroleum-products/data/US-tight-oil-production.xlsx and https://www.eia.gov/dnav/pet/PET_MOVE_RAILNA_A_EPC0_RAIL_MBBL_M.htm; accessed 26 April, 2021 and 17 March, 2021.

___ , "Petroleum Administration for Defense Districts," https://www.eia.gov/petroleum/marketing/monthly/pdf/paddmap.pdf; accessed 6 September, 2023.

U.S. Environmental Protection Agency, "Emissions and Generation Resource Integrated Database," https://www.epa.gov/egrid/summary-data and https://www.epa.gov/egrid/historical-egrid-data; accessed 9 April, 2023.